Volume Title: The Measurement of Labor Cost

Volume Author/Editor: Jack E. Triplett, ed.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-81256-1

Volume URL: http://www.nber.org/books/trip83-1

Publication Date: 1983

Chapter Title: The Fixed Employment Costs of Specialized Labor

Chapter Author: Walter Oi

Chapter URL: http://www.nber.org/chapters/c7374

Chapter pages in book: (p. 63 - 122)

2 The Fixed Employment Costs of Specialized Labor

Walter Y. Oi

The discipline of labor economics has now accepted the proposition that labor is a quasi-fixed factor of production. The basic idea can be traced to J. M. Clark (1923) in *Studies in the Economics of Overhead Costs*. The fixed cost hypothesis was developed to explain the occupational differences in employment and wage rate responses to cyclical fluctuations in the aggregate level of output and employment. The early models did not provide satisfactory explanations for the macroeconomic behavior of unemployment and money wages. Search and contract theories were constructed to fill this void. Hall emphasized the importance of long-term "permanent" jobs which provide the support for contract theories of the labor market. But contract and search theories must ultimately rest on a foundation of fixed costs of one sort or another. Prudent research strategy calls for us to inquire about the factors that can explain why firms and individuals choose to invest in specialized resources which generate these fixed costs.

I shall advance the hypothesis that specialized labor and fixed employment costs are derivatives of an organization of production that reflects the heterogeneity of firms. A comparatively small number of firms grow to extraordinarily large sizes because they are controlled by exceptionally talented and able entrepreneurs. They assemble large production teams by adopting rigid, batch assembly line production processes that are most efficient for the volume production of standardized goods. Each giant

Walter Y. Oi is professor of economics at the University of Rochester, New York, and the Hoover Institution.

The author wishes to thank Jack Triplett, whose encouragement persuaded him to undertake this project. Conversations with Barbara Mann and Stanley Engerman helped to clarify his thoughts. The tabulations from the May CPS were carried out by Barbara Mann. The author alone is responsible for remaining flaws.

firm is characterized by a sufficiently unique organization that can justify and sustain the substantial investments in recruiting and firm-specific training which are optimal only for the largest firms. Employee compensation packages are designed to attract and retain specifically trained workers. Those employees who join and remain with the giant firms are rewarded by implicit contracts that guarantee stable wages and tenured employment. Other individuals who prefer the flexibility and adaptability of small establishments must accept the risks of employment instability due to the shorter life expectancy of small firms. Rigid production techniques and intertemporal substitutions of maintenance activities provide an explanation for Okun's law which deals with the procyclical movements of labor productivity. The empirical evidence and theory developed in this paper support the conclusion that the concept of firm-specific human capital is only applicable to one sector of the economy, namely, the large firms with one thousand or more employees.

2.1 The Quasi-Fixity of Labor

> Once the holding of unused productive capacity was conceived as "idle overhead," it was inevitable that the idea should be extended to human powers as well as to the powers of physical plant and machinery. . . . Wherever a laborer has invested time and money in specialized training, the result is in a sense *fixed capital* which is useful in one occupation and in no other and which must earn whatever return it can because the investment cannot be withdrawn and moved into some other line of business. In such a case, it seems fairly clear that labor involves an overhead cost. (J. M. Clark, 1923, p. 15)

This excerpt contains the ideas of labor as a quasi-fixed factor and of firm-specific human capital.

The cyclical behavior of labor markets exhibits an uneven incidence of unemployment, a compression of occupational wage differentials in the upswing, persistent differences in labor turnover rates, and hiring/firing practices that smack of discrimination. If the partial elasticity of substitution of skilled labor A for the fixed factor capital K is less than that of unskilled labor B for K ($\sigma_{AK} < \sigma_{BK}$), the larger cyclical shifts in demand for unskilled labor can be explained by a neoclassical theory of factor demands. However, that theory cannot explain lags in turning points or differences in labor turnover. These phenomena can better be understood by introducing the concept of labor as a quasi-fixed factor.

Labor cost contains two components—a variable wage that must be paid to obtain a worker's services and a fixed employment cost that a firm incurs to acquire and train a specific stock of employees. In equilibrium, labor's marginal value product (MVP) is equated to the sum of the

expected wage plus the amortization of the fixed employment cost, $q = F/\Sigma(1 + r)^{-t}$; MVP $= W + q$. Since q is a sunk cost, workers in a particular skill will be retained as long as MVP exceeds the variable wage. Workers with high degrees of fixity, meaning high ratios $[q/(W + q)]$, thus experience smaller relative fluctuations in factor demands. Further, the returns to an investment in fixed employment costs can be increased by adopting policies that reduce labor turnover. One should therefore find a negative correlation between an occupation's degree of fixity and its labor turnover rate. The empirical tests generally confirmed the implications of a theory of labor as a quasi-fixed factor.[1]

2.2 Unemployment and Rigid Wages

Quasi-fixity cannot explain the persistence of involuntary unemployment and the sluggish response of wages to changes in aggregate demand. I shall not try to survey the substantial literature that has been motivated by these phenomena. I shall instead provide a brief review of the salient facts and summarize my assessment of search and contract theories. The value of search and contracts obviously depends on the duration of employment relations, which is discussed in section 2.2.4.

2.2.1 The Macrobehavior of the Private Business Sector

In the postwar period, 1947–79, output of the private business sector almost trebled from an index of 48.7 to 144.0 (table 2.1). Fluctuations in output and labor input (measured by man-hours or unemployment rates) were only weakly correlated. Procyclical movements in labor productivity and hours per employee accounted for much of the year-to-year changes in output. Hall (1980) emphasized the absence of market clearing adjustments in money wages; the simple correlation between annual rates of changes in man-hours and wages was $-.123$ for the decade of the 1970s. Data for the longer time series, 1947–79, suggest that the economy may have undergone a structural change, but the two key puzzles which were emphasized by Hall still remain. We must still explain why wages fail to clear the labor market and why labor productivity follows a strong procyclical pattern.

2.2.2 Search and the Pseudoidleness of Prospectors

Idleness can be efficient. Stigler (1962) and Alchian (1969) recognized that search was costly, but unemployed workers willingly incurred these costs to find and secure better paying jobs. In Hutt's (1977) terminology, an unemployed worker is in *pseudoidleness* while he is searching for work or serving the productive function of *availability*—awaiting a call to fill a specialized job.[2]

Table 2.1 **Productivity and Related Variables**[a]

Year	Output X	Total Hours H	Hourly Compen-sation W	Price Deflator P	Adult Pop. (20–64) N	Labor Force LF	Employed Persons E
1947	48.7	90.9	36.0	65.1	84,969	59,350	57,038
1948	50.9	91.5	39.0	70.6	86,013	60,621	58,343
1949	50.0	88.5	39.7	69.8	87,021	61,286	57,651
1950	54.6	89.5	42.4	70.8	88,201	62,208	58,918
1951	57.8	92.1	46.6	76.0	89,017	62,017	59,961
1952	59.5	92.2	49.6	77.1	89,729	62,138	60,250
1953	62.0	93.2	52.8	77.9	90,242	63,015	61,179
1954	60.9	90.1	54.5	78.6	90,775	63,643	60,109
1955	65.7	93.5	55.8	79.8	91,414	65,023	62,170
1956	67.6	94.9	59.5	82.2	92,052	66,552	63,799
1957	68.5	93.5	63.4	84.8	92,634	66,929	64,071
1958	67.0	89.3	66.2	86.4	93,202	67,639	63,036
1959	71.9	92.8	69.0	88.1	93,824	68,369	64,630
1960	73.2	92.9	71.9	89.3	94,477	69,628	65,778

1961	74.2	91.5	74.6	89.8	95,289	70,459	65,746
1962	78.8	92.9	78.1	90.6	96,227	70,614	66,702
1963	82.3	93.4	81.0	91.4	97,490	71,833	67,762
1964	86.9	94.9	85.3	92.7	98,565	73,091	69,305
1965	92.9	97.8	88.7	94.2	99,574	74,455	71,088
1966	98.1	100.0	94.9	97.2	100,585	75,770	72,895
1967	100.0	100.0	100.0	100.0	102,635	77,347	74,372
1968	105.1	101.8	107.6	103.9	104,353	78,737	75,920
1969	108.3	104.6	114.9	108.8	105,981	80,734	77,902
1970	107.3	103.0	123.1	113.9	107,594	82,715	78,627
1971	110.3	102.4	131.4	118.9	109,313	84,113	79,120
1972	117.5	105.5	139.7	123.2	111,071	86,542	81,720
1973	124.4	109.6	151.2	130.3	112,833	88,714	84,409
1974	121.4	110.3	164.9	143.1	114,653	91,011	85,934
1975	118.7	105.6	181.3	157.5	116,510	92,613	84,783
1976	126.4	108.6	197.2	165.5	118,466	94,773	87,485
1977	133.8	112.8	213.0	174.8	120,578	97,401	90,546
1978	140.7	118.1	231.2	187.2	122,717	100,420	94,373
1979	144.0	122.0	252.8	203.8	124,797	102,908	96,945

Source: Economic Report of the President, 1980, table B-37, p. 246.

[a]The first four columns are indexes with 1967 as the base. The last three columns are in thousands.

> When actively searching for work, the situation is that he is really investing in himself by working on his own account without immediate remuneration. He is *prospecting*. He is doing what he would pay an efficient employment agency to do if the course of politics had allowed this sort of institution to emerge in modern society. He judges that the search for a better opening was worth the risk of immediately forgone income. If his relatives or his friends or the state is keeping him, then in a sense they also may sometimes be regarded as investing in him, and it may still be wrong to think of him as idle. But this condition is very difficult to distinguish *in practice* from the various types of "preferred idleness." Thus, unemployment insurance may lessen his incentive to find work, and an apparent or supposed search for the best employment opportunity may be a mask for what is known as loafing. (Hutt 1977, pp. 83–84)

Search and turnover models have been severely criticized by K. Clark and Summers (1979). These models imply large flows into and out of unemployment with only a thin tail of long unemployment spells. Hall (1972) figured that 3.3 points of the unemployment rate could be attributed to normal turnover and search. Estimates for completed spells of unemployment by Clark and Summers sharply reduce this figure to only 0.25 percent. Further, 64 percent of job changes were made with no intervening spell of unemployment. According to Rosenfeld (1977), an individual who was unemployed for four weeks or more devoted only seventeen hours a month to search. Only 35 percent of successful job seekers found their jobs through direct applications to employers. Finally, most workers take the first job offer they receive, and the jobs which they take are held for only short periods lasting less than two years. In the light of these facts, Clark and Summers conclude that it is irrational for an unemployed worker to remain idle while he is searching for a better job.[3]

I see at least three problems that may limit the applicability of traditional search models. First, these models posit an underlying distribution of wage offers and assume that workers search for better paying jobs. The wage rate is, however, simply a proxy for the total utility of employment at different firms. Wage information can be cheaply communicated, but prospectors must visit heterogeneous firms to ascertain the quality of employment. Second, most models assume, for analytic ease, that individuals are alike. Search costs and the returns of search will obviously vary across individuals. Third, the models neglect firm heterogeneity. Search is surely a reciprocal process in which the unemployed seek jobs, and firms search for qualified applicants for vacant jobs.

2.2.3 Risk Sharing and Implicit Contracts

I can identify at least four reasons that can explain the existence of long-term employment agreements: (1) sharing the risks of uncertain

product demands, (2) sharing the risks of uncertain labor productivity, (3) agency costs, and (4) transaction costs of which the most important are the fixed costs of recruiting and training. If workers are risk averse and firms have a comparative advantage in risk bearing, a mutually advantageous agreement can be struck wherein part of the risk of employment instability is shifted to the firm. Gordon (1974) and Baily (1974) appealed to this principle to rationalize implicit long-term contracts. But what is to prevent postcontractual opportunistic behavior? Compliance is hopefully assured by reputation. In short, contract theories must presume that each firm is sufficiently long-lived to have a reputation that is worth protecting.

If productivity is uncertain, payments by results and spot contracts result in uncertain labor incomes. Given risk aversion, F. Smith (1977) has shown that a long-term contract with rigid wages and tenured employment will dominate a contract with fluctuating wages. An implication of this model is that, in competitive equilibrium, junior workers are underpaid, while senior (unsuccessful) workers receive a wage that exceeds their marginal value product.[4]

The delegation of authority is unavoidably accompanied by agency costs of the type analyzed by Jensen and Meckling (1976). The costs of monitoring and malfeasance in a principal-agent relation can be reduced by negotiating long-term contracts that contain substantial elements of deferred compensation.[5] The presence of long-term contracts and deferred pay confound the econometric task of estimating the empirical relation between pay and marginal value products.

A common thread running through these contract theories is that there is more than one dimension to "work." Risk sharing involves a package in which the worker accepts a lower stable wage in return for an implicit insurance policy that yields income smoothing in an uncertain world. In the Becker-Stigler (1974) model, one can imagine that the firm demands a tied bundle consisting of an agent's work effort plus the tied risks of potential losses due to malfeasance or shirking. Specific human capital also involves a tie linking a trained worker to his unique work setting. The forging of these ties in implicit or explicit long-term employment arrangements is presumably advantageous to both parties. The firm's reputation is allegedly the support which persuades workers to believe that the promises will be kept, while deferred benefits induce workers into keeping their part of the bargain. These implications suggest contract theories should only apply to those firms with credible reputations.

2.2.4 Job Tenure and Turnover

A search model examines the behavior of an individual seeking a permanent job, while contract theory describes the behavior of a firm that tries to design pay and employment policies that will attract and

retain "permanent" workers. Employment relations are established and broken, and these labor flows determine a distribution of job tenures. If risk sharing and mobility costs are important to workers, and if fixed employment costs are significant, we should observe long, mean durations of job tenure. The gross flow statistics reveal high labor turnover rates which seem to contradict the underpinnings of search and contract theories.

Labor turnover is costly, and many economists have argued that much can be learned by studying the organization of Japanese firms which purportedly promise their employees guaranteed "lifetime" employment contracts. The myth of the protected and coddled Japanese worker was so pervasive that I never questioned its validity. A very different picture is painted by Koike (1978):[6]

> Those who deserve to be regarded as having "lifetime employment" are *not* Japanese workers, but those in the organized sector of the United States with five or more years of continuous service. (p. 46)

The data assembled by Koike (pp. 64–65) reveal that the percentage of employed persons with fifteen or more years of continuous service is larger in the United States. The percentage with ten or more years was around 34 percent in both countries, while the percentage with less than one year was larger in the United States. I suspect that the wider dispersion of job tenures in the United States can partially be explained by differences in the size distribution of firms.

Using estimated marginal retention rates, Hall (1982) constructed distributions of "eventual" job tenures. A representative worker can be expected to hold ten jobs over a lifetime. Most jobs are of short duration, but by the age of thirty, 40 percent of employed persons will be at a job that they will hold for twenty or more years. There are obvious sex differences; 50 percent of thirty-year-old men will find permanent job attachments compared to only 25 percent of thirty-year-old women. Director and Doctors (1976) found that among blue-collar workers at three large factories, blacks had slightly longer job tenures than whites. Hall's distributions, which pertain to random samples of employed persons, confirm this finding: namely, race is unrelated to job tenure. The picture of the labor market implicit in these job tenure distributions is one of turbulence during the first five to ten years in the labor force. Young persons move from job to job as they look for a "permanent" job. With increasing age, larger fractions eventually settle into a job that will last for twenty or more years. Data from the May 1979 Current Population Survey (CPS) reveal some interesting interactions between job tenure and firm size which are explored in section 2.3.

Our concerns over unemployment and rigid wages have promoted the development of search and contract theories that rest on the presence of

fixed costs. The magnitude of these fixed costs obviously varies across individuals and firms. A marginal firm, whose survival probabilities are slim, is unlikely to make large investments in recruiting and training. An individual with few assets and general talents will not incur high search costs. Search and contract theories are only applicable to a sector of the economy for which these fixed costs are significant. Attention is directed in the next section to the characteristics of firms and workers that put them into this sector.

2.3 The Production and Compensation of Specialized Employees

Training and goods are joint products. Firms "produce" specifically trained employees whose internal value to the firm exceeds their external value in an outside labor market. Firm-specific human capital and the discrepancy between internal and external values can only be sustained when the host firms are sufficiently differentiated from one another. Variations in entrepreneurial ability can generate a distribution of firms that differ in size and organization. The very large firms achieve the economies of volume production by installing rigid, specialized production plants. The resulting organizations yield short-run factor demands that produce procyclical patterns in labor productivity. Less able entrepreneurs, who control smaller production teams, occupy a different segment of the "product line." They survive by assembling adaptable production teams that utilize general-purpose equipment and employ workers with general human capital. Firm-specific human capital is a phenomenon that is only observed in that sector of the economy consisting of very large firms. The labor market in this part of the economy does *not* conform to the neoclassical theory of factor markets.

2.3.1 Fixed Employment Costs and the Joint Production of Training and Goods

The full costs of quasi-fixed labor inputs are the sum of variable and fixed components. Wages and fringe benefits that make up total employee compensation are usually included in the first component.[7] The fixed employment costs, which represent outlays for recruiting and training, are likely to be higher, the greater the specificity of the firm's labor force.

If a resource is *specialized* and *specific* to a firm, it earns an equilibrium return that contains an element of economic rent. Its internal value will exceed its external value. Firms may purchase specialized resources from outside vendors, but the usual arrangement involves vertical integration wherein specific factors are "internally produced."[8] The firm that demands specially trained labor input will ordinarily find that specific training can be most economically provided through internal or in-house

production. It thus becomes a multiproduct firm that jointly produces goods and specific human capital.

The costs of producing specific human capital will depend on the technology of production and the input prices which include the wages paid to apprentice-trainees and instructors, materials costs, and the opportunity costs of forgone output. If S units of specific human capital are produced and supplied to each new hire, the gross return realized by the firm will be equal to the present value of increments to labor productivity (in value terms) attributable to S. The gross returns will be larger, the lower the interest rate, and the greater the durability of S.[9] The optimum investment in firm-specific human capital is attained by equating marginal costs to marginal returns. This investment is likely to vary across individuals and firms. More specific human capital will be invested in those individuals who have longer expected employment tenures and larger increments in productivity. If apprentice-trainees agree to share the costs by accepting a lower wage during the training period (which translates into a fall in the price of an input entering the production function), the marginal cost of specific training falls thereby increasing the equilibrium investment in S. Other implications could be derived by placing more structure on the model.[10]

A firm hires individuals of varying abilities to perform different tasks, and the nature of the tasks will surely affect the returns to specific training. Orientation costs are, for example, small when the job involves simple tasks that are performed in more or less the same way in many firms. Other jobs that require the use of specialized equipment or close cooperation with team members may demand extensive orientation and training. We have read about the highly valued clerk who through formal training and informal on-the-job experience has learned how to deal with the firm's best customers. Substantially more training must be supplied to those managers and supervisors who are asked to monitor performance, to train new employees, and to handle unanticipated departures from normal work routines. These examples suggest that specific human capital is largely concentrated among highly paid, skilled workers.

Individuals who have a greater capacity to learn are likely to acquire larger stocks of both types of human capital. The marginal returns to specific human capital S are likely to be greater, the larger the supply of the cooperating input of general human capital G. Equalization of marginal returns across individuals will thus yield a positive correlation between S and G. Further, general human capital G determines an individual's external market wage W.[11] The usual principles of production can thus generate a positive association between the wage rate and the degree of fixity. Moreover, the returns to a given investment in specific human capital are likely to vary across firms. Those firms that enjoy higher returns will realize greater gains by demanding more able workers

whose general human capital will have a complementary effect on the productivity of specific human capital. This heterogeneity of firms thus reinforces the tendency for wages to be positively correlated with investments in firm-specific human capital.

2.3.2 The Planning of Production by Heterogeneous Firms

The analysis of labor markets has explicitly recognized the presence of individual worker differences in both observable variables (education, job experience, race, sex, etc.) and unobservable traits (intelligence, honesty, perseverence, etc.) These differences operating through the supply side play important roles in explaining the dispersion in earnings, the differential incidence of unemployment, job mobility, and so on. On the demand side, there are obvious differences among firms. Industrial differences are usually explained by arguing that there are different technical substitution opportunities embedded in the production functions applicable to different industries. However, firms in a given industry also differ in behavior and in the organization of production. These firm differences cannot be adequately explained by our received theory of value.

Economic theory only provides a loose definition for the concept of the *firm*. Coase (1937) persuasively argued that the firm is a viable organization because some resource allocations are more economically made by command rather than by market transactions. Alchian and Demsetz (1972, p. 778) emphasized the role of coordinating team production when they wrote, "It [the firm] is the centralized contractual agent in a team productive process, not some superior authoritarian directive or disciplinary power." Following Kaldor (1934), I shall assume that the "centralized contractual agent" can be equated to a *single*, firm-specific entrepreneurial input.[12] The quality and quantity of this input will surely vary, and it is this variation which can explain the size distribution of firms.

The Entrepreneurial Input in a Neoclassical Model

Before turning to those aspects of production which deal with adaptability, specific training, and recruiting, attention is directed to a neoclassical model in which output Q is produced by combining three inputs:

$$(1) \qquad Q = f(N, K, T); \qquad \frac{dQ}{dK} = f_K, \frac{dQ}{dN} = f_N, \frac{dQ}{dT} = f_T.$$

Capital K is purchased at a price R. A firm that hires M workers of type μ obtains a labor input measured in efficiency units of $N = \mu M$. All entrepreneurs are endowed with the same fixed supply of calendar time $\bar{H}$ which can be allocated to coordinating production or monitoring worker performance. Monitoring is an essential joint input that must be supplied

by the entrepreneur to assure that each of the M workers contributes μ efficiency units of labor services.[13] Entrepreneurs are assumed to be alike as monitors, and they must devote h hours to monitor each worker. They do, however, differ in their capacity to coordinate production. These differences will be described by an entrepreneurial ability parameter λ which transforms the time allocated to coordinating and decision making into efficiency units of managerial effort T. The supply of managerial effort which cooperates with the hired input is thus determined by entrepreneurial ability λ and the *number* of employees M:

$$T = \lambda H = \lambda(\bar{H} - hM). \tag{2}$$

More productive workers can command higher wages along a market wage structure $W(\mu)$ with $W'(\mu) > 0$. The profits of a competitive firm are thus given by

$$\pi = PQ - RK - W(\mu)M. \tag{3}$$

Inputs of capital K and employees M, as well as worker quality μ, are chosen to maximize profits. The first-order conditions are

$$Pf_K = R, \tag{3a}$$

$$P(\mu f_N - \lambda h f_T) = W(\mu) \text{ or } P\mu f_N = W(\mu) + \delta \qquad (\delta = P\lambda h f_T), \tag{3b}$$

$$Pf_N = W'(\mu). \tag{3c}$$

This system of three equations determines the profit-maximizing values of (K, M, μ). The properties of this model can be more easily understood by temporarily assuming that worker quality μ is held constant.

In equilibrium, the MVP of capital is equated to its price R. However, the MVP of workers exceeds their wages by an amount equal to the implicit monitoring cost δ which represents the opportunity cost of diverting h hours of entrepreneurial time away from coordination to the supervision of worker performance. The marginal rate of substitution of capital for workers is equated to relative factor prices where the pertinent "price" of labor is its full cost, defined as the sum of the wage needed to obtain a worker's services plus the opportunity cost of the time required to guarantee that the worker will contribute μ efficiency units of work effort:[14]

$$-\frac{dK}{dM} = \frac{\mu f_N}{f_K} = \frac{W(\mu) + \delta}{R}. \tag{4}$$

Firms that incur higher monitoring costs face a higher full price of labor. Consequently, they adopt more capital-intensive production techniques.

Profits, in this model, are the returns to the quasi-fixed entrepreneurial input. If the production function, equation (1), is homogeneous of the

first degree, profits in a competitive industry will be directly proportional to the shadow price of entrepreneurial time:[15]

$$\pi = P\lambda f_T \bar{H}. \tag{5}$$

Profits are positively related to λ and P, and inversely related to h, R, and $W(\mu)$. A marginal entrepreneur is one whose ability level λ_0 is such that he realizes the same income from either pursuit, entrepreneur or worker; i.e., $\pi_0 = W$. Those whose entrepreneurial ability exceeds this threshold level ($\lambda > \lambda_0$) will make up the supply of entrepreneurs, while the remaining individuals constitute the supply of hired workers. The production function and the right tail of the frequency distribution of entrepreneurial abilities, $\phi(\lambda)$ to the right of λ_0, jointly determine an equilibrium size distribution of firms. The output supplied by each firm is thus a function of real input prices, entrepreneurial ability, and the monitoring loss parameter, $Q = Q[(R/P), (W/P), \lambda, h]$. In competitive market equilibrium, the product price equates the industry supply, Q^s, to the market demand, $Q^d = D(P)$:

$$Q^s = \int_{\lambda_0}^{\infty} Q\left(\frac{R}{P}, \frac{W}{P}, \lambda, h\right)\phi(\lambda)d\lambda = D(P). \tag{6}$$

Inframarginal entrepreneurs with high values of λ earn economic rents ($\pi - W$) that are *not* eliminated by competitive market forces. Although higher entrepreneurial abilities entail higher monitoring costs, output, employment, and profits are positively related to ability λ. As a consequence, more able entrepreneurs control larger firms.

For a given worker type, say μ_1, inputs of K and M are demanded so that the constrained marginal cost $\bar{\gamma}$ is equated to the product price:

$$\bar{\gamma} = \frac{R}{f_K} = \frac{W_1}{\mu_1 f_N - \lambda h f_T} = P \qquad [W_1 = W(\mu_1)].$$

The opportunities to vary worker quality introduces a new degree of freedom. The input mix for $\bar{\gamma}$ need not correspond to a global minimum of costs. If the quality margin is equated to the other two margins, the firm attains a global maximum of profits described by the equality of the unconstrained marginal cost to price:

$$\gamma = \frac{R}{f_K} = \frac{W(\mu)}{\mu f_N - \lambda h f_T} = \frac{W'(\mu)}{f_N} = P.$$

The nature of the full equilibrium is clarified by examining the way in which the choice of worker quality affects the costs of the labor input. A given input of labor services measured in efficiency units can be produced by various combinations of numbers M and qualities μ. The full cost of labor is the sum of wages and implicit monitoring costs:

$$C = [W(\mu) + \delta]M = [W(\mu) + \delta]\left(\frac{N}{\mu}\right).$$

We can derive the marginal cost of a move to higher quality accompanied by a decrease in numbers, holding N constant:

$$\frac{dC}{d\mu} = \frac{N}{\mu^2}[\mu W'(\mu) - W(\mu) - \delta].$$

The sign of $(dC/d\mu)$ will depend on μ and the implicit monitoring cost δ which is higher for larger firms. There is an ability level λ_1 for which μ_1 would have been that firm's optimal choice of worker quality. A firm with a higher ability entrepreneur ($\lambda_A > \lambda_1$) incurs a higher monitoring cost so that $\mu_1 W'(\mu_1) < W(\mu_1) + \delta_A$. The λ_A firm can reduce the full costs of labor by substituting quality for quantity. In equilibrium, the worker productivity which minimizes full labor cost satisfies the condition that the marginal cost of quality equals the full cost of an additional worker:

$$\mu W'(\mu) = W(\mu) + \delta. \tag{7}$$

Entrepreneurs will locate along the wage structure $W(\mu)$ in a manner analogous to the hedonic price model of Rosen (1974). Large firms that incur higher monitoring costs will demand more productive workers who command higher wages, requiring less monitoring per efficiency unit of labor services. The equilibrium market wage structure must equilibrate the relative demands and supplies for workers of varying productivities. Moreover, it must exhibit increasing returns to quality, meaning that if individual A is twice as productive as B, A must receive a wage that is more than twice B's wage.[16] More productive workers are matched with more able entrepreneurs, thereby generating a positive relation between wages and firm size. This assignment of workers to firms is socially optimal in the sense that it minimizes the full social cost of producing monitored labor services.

A displacement of equilibrium results in distributional effects among the firms in a given industry as well as allocative effects across industries. The nature of these adjustments can be described with the aid of an illustration. Suppose that a wage tax is placed on workers in a particular industry. In a Marshallian analysis, the "representative firm" will contract output, profits will fall, and capital will be substituted for labor. The industry demand for capital will decline if the elasticity of substitution is less than the price elasticity of demand. In the presence of heterogeneous firms, these conclusions have to be qualified. Marginal entrepreneurs, whose abilities are only slightly above the threshold level λ_0, are driven out of the market as profits fall below the alternative wage that they could have earned as workers. The tax imposes a greater burden on the smaller

surviving firms whose higher labor/capital ratios cause larger upward shifts in their marginal cost curves. For the industry as a whole, the adjustment in product price due to the tax, will depend on four factors: (a) production technology, (b) the price elasticity of the product demand curve, (c) the increase in the threshold ability level λ_0 due to the tax, and (d) the frequency distribution of abilities $\phi(\lambda)$. A dispersion of entrepreneurial abilities could thus result in a situation where the smaller firms cease or contract production, while the larger firms actually expand their output, employment, and profits.[17]

Differences in entrepreneurial ability can generate an equilibrium size distribution of firms even though all entrepreneurs have access to the same production technology and to common, perfectly competitive factor markets. Ability in my model is not "Hicks neutral." Specifically, talented entrepreneurs have a comparative disadvantage at monitoring. As a consequence, they try to economize on monitoring by adopting capital-intensive production techniques and hiring more productive workers. These are implications that can be empirically tested.

Monitoring, Training, and Productivity

Monitoring costs could be reduced if production could be organized so that workers are paid by "results." The production methods that allow for piece-rate compensation may exclude techniques that can realize the economies of specialization and team production.[18] Team production requires the joint input of monitoring to prevent shirking. The costs of monitoring teams cannot be easily allocated to individual team members. The determination of an optimal level of monitoring thus involves elements of the problem of public goods.

The production of specialized teams ordinarily requires specific training that raises a worker's productivity in only one particular firm. If some minimal amount of firm-specific human capital is required to become a team member, what determines the manner and timing of its production? With respect to general human capital, Ben-Porath (1967) showed that if an individual maximizes utility, a rising marginal cost curve will generate a time path of investments in human capital that are spread out over time. The stock of capital will increase at a decreasing rate, resulting in the familiar concave age-earnings profile. A firm's investments in specific human capital ought to follow a similar path, increasing at a diminishing rate as a function of length of service. Further, if the returns to specific human capital are shared, productivity and wages should both increase with job tenure. Medoff and Abraham (1981) have challenged the validity of this model. They reported that wages within a job grade increase with tenure, but productivity measured by supervisory ratings or physical output rates is unrelated to job experience. Wages are evidently *not*

determined by the equilibrium returns to human capital and, by implication, they have to be explained in some other way, such as risk sharing or paternalism.

The Medoff-Abraham conclusion rests on the tacit assumption that workers are like single product firms; i.e., secretaries type, roofers lay shingles, and scientists publish articles. In reality, most workers resemble multiproduct firms that jointly supply several products. In addition to typing, a secretary may be responsible for organizing the office, training new employees, and being available for service as a temporary replacement for an absentee. Firm-specific training is not intended to increase typing speed. It is designed to improve performance in those aspects of the job that are unique to the firm. The option value of this backup capability (which is similar to Hutt's example of an idle worker performing the productive function of "availability") should be included in measuring a worker's total productivity. Additionally, a worker who is reliable and requires less monitoring has a higher *net product* which cannot be measured by simply observing his *gross product*. These related, firm-specific dimensions of workers' value to their employers are largely neglected in conventional measures of labor productivity. Reliance on conventional measures thus tends to understate the impact of firm-specific training on total labor productivity.[19]

Specialization and Team Production

A firm can realize the gains from specialization by organizing production around units and teams. The output of the entire team can be observed, but the marginal contribution of a particular worker is not easily ascertained because of (a) interdependence in the production function and (b) variations in the supply of work effort. If effort and performance are to be properly compensated, someone has to monitor and meter worker performance. In the Alchian-Demsetz model, the entrepreneur is the specialist who detects shirking and metes out rewards. The delegation of authority in a principal-agent relation is unavoidably accompanied by shirking and incompatible incentives that produce "agency costs." Jensen and Meckling (1976) point out that agency costs (monitoring, bonding, and the residual loss) constitute an efficient allocation of resources. Without them, a firm could not obtain and retain the services of agents and employees who have comparative advantages in performing certain tasks. In short, agency costs have to be incurred if a firm is to achieve the requisite size and organizational structure that are needed to exploit the economies of specialization and volume production.

In the neoclassical model discussed earlier, all firms used the *same* production function to produce a homogeneous good Q. However, the

firms in a given industry appear to produce slightly differentiated products. Large firms specialize in the production of standardized goods, while small firms supply customized goods that are produced in small batches.

Standardized goods are most efficiently produced by combining specialized capital with a disciplined labor force that will conform to prescribed work schedules. Production is characterized by a putty clay technology in which capital can be substituted for labor in the ex ante planning stage, but once in place, machines and men are employed in virtually fixed proportions which simplifies the monitoring task. The rigidity of this organization seems well suited to exploit the volume economies emphasized by Alchian (1959).[20]

If a firm expects to sell only fifty units of a good, it will choose an adaptable production technique. If, however, the planned volume is one hundred thousand units, production will be organized around an assembly line.[21] Specialized, durable machines will be designed for batch production, and these machines will be operated by fixed complements of workers. Monitoring costs can be reduced when workers are compelled to adhere to the same inflexible work schedules and when the opportunities for discretionary behavior are limited by a fixed-proportions technology. To the extent that worker preferences vary, the lower monitoring costs are not a windfall. Large firms must pay a compensating wage difference to attract marginal team members.[22]

The coordination of very large production teams is facilitated by developing detailed job descriptions. Job applicants are tested and interviewed, and prior work histories are carefully reviewed to determine if the applicant's qualifications meet the prescribed job specifications. Applicants are passed over and job vacancies are kept open until a suitable match is found. Small firms which have lower monitoring costs and more adaptable production teams are able to fill job vacancies more quickly because the requirements of the job can be more easily modified to fit the applicant's qualifications. The relative rigidities of production plans thus predict that the ratio of applicants to job vacancies will be higher at larger firms and for those positions that have tighter, inflexible job specifications.[23]

The new employees at very large firms are likely to receive more firm-specific training. But training here must be broadly defined to include the acquisition of new skills (e.g., learning how to operate a word processor or a forklift truck) as well as the adaptation to a particular production process (meaning compliance with prescribed working practices or learning preferred ways to perform certain jobs). When all of a team's members are more or less alike and follow standardized work routines, monitoring costs are reduced, and some monitoring could even

be delegated to subordinates.[24] The fixed costs of assembling and training a homogeneous labor force may not be warranted for small firms that have shorter, uncertain lives.

Small firms are drawn to a segment of the industry product line containing what can be called customized goods. Products are supplied in small batches and are often differentiated to meet idiosyncratic demands. A firm in this market should assemble an adaptable production team which can easily adjust to changing demands. I have argued earlier that the capital to labor ratio will be smaller because the lower monitoring cost reduces the full cost of labor. The lower capital intensity of smaller firms is further reinforced if substitutions across differentiated products are more easily made by moving men rather than machines. Part-time workers can be more efficiently employed by small firms which are continually adjusting to changes in the level and composition of productive activities. The virtual absence of firm-specific training at small firms can be explained by the adaptability and generality of production and by the shorter expected lives of these firms.

Every industry contains firms of varying sizes and types. At one end of the spectrum, little companies produce goods in small batches by using labor-intensive adaptable techniques and by hiring low-wage workers with general human capital. At the other extreme, very large firms specialize in the volume production of standardized goods. Production is organized around assembly lines (or fixed plants designed for large batches) that are characterized by (a) ex post fixed factor proportions, (b) rigid work schedules, (c) detailed job descriptions, (d) homogeneous inputs, (e) high capital utilization rates achieved through multiple shifts, and (f) quantum adjustments to changing demand conditions. The composition of a firm's product line and the organization of production are thus determined in a manner that maximizes the returns to the scarce entrepreneurial input.

2.3.3 Short-Run Cyclical Adjustments and Okun's Law

The fixed cost hypothesis was originally advanced to explain different employment and wage responses to cyclical changes in aggregate demand. The literature in the last two decades has examined a broader range of issues including search and contract theories, labor turnover, equilibrium unemployment rates, the duration of job tenures, and, last but not least, Okun's law which dealt with the cyclical behavior of labor productivity. Specifically, Arthur Okun found that a 1 percent decrease in the unemployment rate (which is approximately equal to a 1 percent increase in the aggregate labor input) was accompanied by a 3 percent increase in aggregate output. This empirical regularity is contrary to a naive production function model in which capital is fixed in the short run, resulting in diminishing returns to the variable labor input. Hall (1980)

offered a conjectural explanation wherein labor markets are dominated by workers on long-term contracts who felicitously vary their work efforts in a procyclical fashion. The background model is one where the effective labor input is a function of work effort which, in turn, responds positively to changes in product demand.[25] A durable labor input may play an important role in Okun's law, but the elusive concept of "work effort" can, I believe, be replaced by a more plausible argument that appeals to an efficient intertemporal allocation of specifically trained workers.

Short-run adjustments to demand shifts will be determined by the production technology, factor supply conditions, and opportunities for product substitutions. I shall direct attention to two themes that have not been adequately developed in the literature and that go a long way in explaining Okun's law. The first is concerned with the organization of volume production in which zero ex post factor substitutions (fixed factor proportions) are a consequence of volume production. The second theme extends the idea of joint production to the intertemporal allocation of quasi-fixed factors to market and nonmarket activities.

The essence of the first theme can be found in the peak load pricing model and is implicit in the regenerative growth model of Gordon and Walton (1982).[26] In the Steiner (1957) model, outputs in peak and slack periods are produced by combining capital and labor in fixed (one-to-one) proportions; $X_j = K_j = L_j$ for $j = 1, 2$. Capital is specialized with no alternative use, while the variable labor input is general. If demands produce a firm peak ($X_1 = K$ and $X_2 < K$), the output/capital ratio is less than unity in the slack period. As the firm moves from slack to peak periods, the output/capital ratio rises, but no one would point to this as an example of increasing returns.

The application of the peak load model to Okun's law is clarified by assuming that there are three inputs: specific labor A, variable general labor B, and specific capital K. If one unit of A is required for the maintenance of one unit of K whether it is or is not in use, the labor input in the peak period is $L_1 = (A + B_1)$ where $B_1 = X_1$. In the slack period, the labor input falls to $L_2 = (A + B_2)$ where $B_2 = X_2 < X_1$. The A labor is specialized and is retained during slack periods. As the firm moves from slack to peak periods, the output/labor ratio climbs because the A labor is, in a sense, more efficiently utilized during peak periods.

The peak load model has been extended by Turvey (1968), Wenders (1976), and others to allow for a portfolio of diverse technologies. Capital-intensive plants that yield the lowest full-cycle costs are fully utilized in all periods, while standby plants with lower capital/labor ratios are idle in slack periods but activated to meet peak demands. Large manufacturing firms seem to embrace a similar strategy. Those plants (or parts of a plant) that are operated on multiple shifts tend to be highly capital intensive. Standby plants and shifts that are added to meet peak demands

have lower ratios of capital to variable labor inputs. In a cyclical downturn, entire plants and shifts are closed down, thereby producing a countercyclical movement in the ratio of output to *variable* labor inputs. However, the diminishing returns to variable labor have to be combined with increasing returns to the quasi-fixed labor inputs in determining the cyclical behavior of *total* labor productivity.[27] If production entails fixed proportions in which quasi-fixed and variable labor inputs are combined with capital, short-run adjustments to demand shifts could generate time paths for output and employment that conform to Okun's law. Further, since large firms are more likely to adopt putty clay technologies, the procyclical movements in labor productivity should be stronger for larger firms.

The second theme acknowledges that all large firms are vertically integrated, multiproduct enterprises. Resources at the command of a firm can be allocated in at least three directions, namely, to the production of (1) final goods Q, (2) firm-specific human capital, or (3) increments to the value of existing physical plant and equipment. The joint production of training and goods has already been discussed, and the same idea can be extended to the joint internal production of capital values.

Machines do not run like the "one-hoss shays" of some economic models. Depreciation is *not* exogenous, and investments (additions to the capital stock) do not always take the form of new capital goods purchased from outside vendors. Further, machines are not homogeneous, and one of the important quality features is the probability of breakdown. Resources are allocated to maintenance and repairs to sustain the service flows from capital and to raise the market value of the existing capital stock. Each firm will choose a quality of capital (I shall emphasize age as a proxy for quality) and a level of maintenance inputs that maximize total profits.

A firm can control the age distribution of its capital stock through its choice of age of additions (new vs. used, and if used, the age of used equipment) and the age at which equipment is scrapped or sold in the used market. These decisions are made by comparing three components of capital costs—amortization, maintenance, and disruption. Since depreciation and obsolescence rates decline with age, newer machines entail higher amortization charges. These may be offset, in part or in whole, by lower maintenance and disruption costs, where the latter include the opportunity costs of forgone output and the costs of any tied inputs that must be retained during any downtime. Small firms are more likely to purchase used machines and to discard them at older ages for two reasons: First, maintenance is a labor-intensive activity that must be closely supervised. The lower "price" of maintenance due to lower monitoring costs leads to a substitution of more maintenance for lower

amortization charges. Although the probability of a breakdown rises with increasing machine age, each disruption entails lower costs when the firm uses an adaptable production technique that allows for variable factor proportions. Thus, small firms that have a comparative advantage in maintenance and in coping with disruptions, tend to own older machines. The giant firms that have large, inflexible production lines face higher maintenance prices. They willingly accept the higher amortization costs of new equipment in order to economize on monitoring and disruption costs. The optimal trade-offs of amortization versus maintenance/disruption costs are determined by relative "prices" which happen to be related to firm size. The age structure of a firm's capital assets can thus be explained by the minimization of the sum of amortization, maintenance, and disruption costs.[28]

The principle that maintenance can be substituted for investment is well known.[29] Durable capital is scrapped at older ages in those countries where the wage rate is low in relation to the price of new equipment. This same principle applies to the intertemporal substitutions that a firm can make in response to cyclical fluctuations in relative factor prices and product demands. In a recession, the shadow prices of quasi-fixed specialized resources fall, but firms continue to retain them. Product prices decline in relation to the implicit value of internal investments. As a consequence, quasi-fixed inputs are shifted away from the production of goods toward internal investments in specific human capital and in maintenance of physical capital.[30] The market value of capital assets increases (or falls by less than it otherwise would in the absence of increased maintenance), thereby reducing the effective depreciation rate. These increments to capital values are all implicit and never appear on the company's books.

The process is reversed when the firm experiences an upturn in product demand. The demand for variable inputs increases, and specialized workers are reassigned from maintenance to the production of goods. Physical capital is more intensively utilized, and with less maintenance, it depreciates at a faster rate. To the extent that Okun's law only deals with cyclical movements in the value of final goods Q (and ignores the countercyclical movements of increments to implicit capital values), we get a biased picture of the cyclical behavior of "total" labor productivity. The magnitude of the output response is muted if we followed the correct procedure and related changes in the value of "total product" (including internally produced increments to capital values) to changes in labor inputs.[31]

The force of these two themes—fixed factor proportions in producing goods and intertemporal product substitutions—obviously varies across firms and industries. The high shadow price of entrepreneurial time in large manufacturing firms raises the costs of monitoring workers as well as the costs of enforcing compliance with frequent changes in prices

and wages. Hence, large firms should exhibit greater rigidities in prices and wages accompanied by discrete quantum adjustments in output and employment. The firm adjusts to a downturn by closing down entire plants and shifts, releasing variable labor inputs, and placing some trained workers on temporary layoffs. The specialized workers who are retained are diverted to the nonmarket activities of rebuilding human and nonhuman capital. At the other end of the size distribution, small entrepreneurs confront a lower opportunity cost of time. They are better suited to monitor workers in adaptable production teams, and more importantly, they can negotiate frequent changes in prices and wages. Although a recession may drive many small firms into bankruptcy, the surviving small firms may, as a result of their flexibility, experience less volatility in output and employment. King (1923) found, for example, that the reductions in employment in the recession of 1920–22 were relatively greater in large firms, as revealed by the data of table 2.2.

Employment in the volatile manufacturing sector is heavily concentrated in large firms. In this sector, fixed factor proportions and intertemporal substitutions in the production of goods versus internal investments are likely to produce strong procyclical movements in labor productivity. In other sectors, small firms with flexible prices and adaptable production teams should exhibit weak procyclical or even countercyclical patterns in the ratio of output to employment.

2.3.4 Impact of Firm Size on Compensation and the Composition of Employment

Fixed employment costs are incurred to recruit and train a *firm-specific* labor force. The amount invested in and the returns to specific human capital will be larger, the greater the durability of the asset measured by the worker's expected job tenure. The retention rate is a function of the level and structure of compensation, the quality of working conditions, and the composition of the firm's labor force. The design of a compensation package and the selection of new employees are clearly more important to those firms that make large investments in specific human capital.

Total employee compensation (TEC) is the sum of gross wages plus employer contributions for fringe benefits; TEC = $WH + BH$, where W is the hourly wage rate, H denotes total paid hours (the sum of hours actually worked H_W plus paid leisure hours H_L for vacations, holidays, sick leave, etc.), and B is the employer outlay for fringes converted to an hourly rate. The total hourly compensation of employees which appears in government publications[32] is given by $C = (\text{TEC}/H) = W + B$. But C is not the right measure for the compensation component of full labor costs.[33] A better measure is provided by compensation per hour actually worked, which I shall call *pay*, $C^* = (TEC/H_W) = C/(1 - \phi)$, where

Table 2.2 **Employment, Hours, and Earnings for Selected Quarters, 1920–1922**

	All Industries			Factories		
Firm Size	1920:3	1921:3	1922:1	1920:3	1921:3	1922:1
	Number of Employees					
1–21	10,110	9,843	8,739	1,360	1,251	1,121
22–99	4,630	4,084	3,956	1,950	1,541	1,573
100+	14,440	11,151	11,452	8,060	5,668	5,927
Total	29,180	25,078	24,147	11,370	8,460	8,621
	Scheduled Full-Time Weekly Hours					
1–21	54.1	53.9	53.0	52.1	51.9	51.0
22–99	52.0	51.4	51.3	51.8	50.4	49.9
100+	49.1	48.3	47.8	50.1	49.1	49.3
Total	51.3	51.0	50.3	50.7	49.7	49.6
	Actual Weekly Hours					
1–21	52.9	52.7	51.7	51.0	50.8	49.7
22–99	48.6	48.8	49.0	46.2	47.2	46.3
100+	48.9	45.5	45.6	49.0	44.4	45.7
Total	50.3	48.8	48.1	48.7	45.9	46.3
	Average Hourly Earnings					
1–21	0.44	0.43	0.44	0.50	0.52	0.54
22–99	0.56	0.54	0.52	0.58	0.56	0.54
100+	0.63	0.57	0.56	0.64	0.56	0.53
Total	0.55	0.51	0.51	0.61	0.56	0.53

Source: W. I. King, *Employment, Hours, and Earnings in Prosperity and Depression* (New York: National Bureau of Economic Research, 1923), pp. 30, 82, 87, 113.

$\phi = H_L/H$ is the paid leisure ratio describing the fraction of total paid hours that the worker can take in leisure. The components of employee compensation from two surveys are summarized in table 2.3. In the BLS survey of all establishments in the private business sector, the hourly wage W was only 78.1 percent of pay per workhour C^*; it was only 70.3 percent in the Chamber of Commerce survey of large firms. Establishments in the BLS survey reported giving an average of four weeks of paid leisure, while the large firms in the CC survey gave nearly seven weeks.[34] Fringe benefits B accounted for 15.5 and 19.1 percent of total compensation in the two surveys. Roughly 40 percent of these fringes were legally required for social security, worker compensation, and unemployment

Table 2.3 Components of Total Employee Compensation

Item	Bureau of Labor Statistics 1977	Ratio[a]	Chamber of Commerce 1979	Ratio[b]
Gross hourly wage, W	6.28	2.18	7.311	2.37
Compensation per paid hour, C	7.43	2.32	9.037	2.52
Compensation per workhour, C^*	8.04	2.37	10.387	2.59
Paid leisure percentage, ϕ	7.59	1.29	13.00	1.23
Legally required benefits, B_{LR}	0.51	3.00	0.658	3.60
Employer benefits, B	1.15	3.59	1.730	3.49
Total benefits, $(B + \phi W)$	1.63	3.33	2.680	3.26

Sources: U.S. Bureau of Labor Statistics. 1980. "Employee Compensation in the Private Nonfarm Sector, 1977." Washington, D.C.: Government Printing Office. U.S. Chamber of Commerce. 1981. "Employee Benefits: Historical Data, 1951–79." Washington, D.C.

NOTE: All figures except for C^* and ϕ are in current dollars per paid hour.

[a]Ratio of the 1977 value to the 1967 value.

[b]Ratio of the 1979 value to the 1969 value.

insurance. Differences by occupation and industry from the BLS survey are shown in table 2.4, and the CC data for some twenty industries appear in table 2.5.

The relation of wage to pay can be described by the ratio W/C^*. Recall that $C = W + B$, and $C^* = C/(1 - \phi)$, so that

$$\frac{W}{C^*} = \frac{1 - \phi}{1 + \frac{B}{W}}.$$

Reference to table 2.6 reveals that the wage to pay ratio, W/C^*, fell from 83.1 percent in 1951 to 70.4 percent in 1979. This secular trend resulted from increases in both the benefits rate B/W and the paid leisure ratio ϕ. Legally required supplements B_{LR} accounted for 18.4 percent of total benefits $(B + \phi W)$ in 1951 and 24.6 percent in 1979. I conjectured that as B_{LR} rose, private fringes would fall so that the sum (as a percentage of total compensation) would remain roughly stable. Although private pension contributions grew more slowly than social security, it increased in relation to total compensation.

The cross-sectional industrial dispersion in the W/C^* ratio is substantial (table 2.5), ranging from 83.5 percent in textile mills to 65.4 percent in petroleum. The relative importance of fringe benefits is greater in the regulated public utilities, banking, and the public sector. The variance in W/C^* is largely attributable to variations in the paid leisure ratio. Employees in chemicals and public utilities get more than eight weeks of paid leisure, while workers in retail trade got only four weeks. The data of table 2.7, showing the percentage of workers receiving three selected

Table 2.4 **Employee Compensation in the Private Nonfarm Sector, 1966 and 1977**

	All Employees		Office Workers		Nonoffice Workers	
	1966	1977	1966	1977	1966	1977
C^* = Compensation per work hour						
All industries	3.40	8.04	4.51	9.96	2.92	6.96
Manufacturing	3.67	8.82	5.28	11.80	3.17	7.77
Nonmanufacturing	3.23	7.68	4.15	9.42	2.75	6.49
W = Wage (average hourly earnings per paid hour)						
All industries	2.88	6.28	3.79	7.74	2.48	5.44
Manufacturing	3.05	6.64	4.36	8.80	2.64	5.86
Nonmanufacturing	2.77	6.11	3.52	7.42	2.36	5.18
ϕ = Paid leisure percentage (paid nonworkhours/paid hours)						
All industries	5.9	7.3	7.5	8.2	5.1	6.7
Manufacturing	6.8	8.9	8.3	9.6	6.3	8.4
Nonmanufacturing	5.3	6.5	7.2	7.7	4.4	5.4
W/C^* = Wage as percentage of compensation per workhour						
All industries	84.7	78.1	84.0	77.7	84.9	78.2
Manufacturing	83.1	75.3	82.6	74.6	83.3	75.4
Nonmanufacturing	85.4	79.6	84.8	78.8	85.8	79.8
B_{LR}/C = Legally required benefits to compensation per paid hour						
All industries	5.31	6.86	3.84	5.42	6.14	8.01
Manufacturing	4.97	6.61	3.51	5.11	5.72	7.31
Nonmanufacturing	5.23	6.99	3.90	5.59	6.46	9.80

Source: U.S. Bureau of Labor Statistics, 1980. "Employee Compensation in the Private Nonfarm Economy, 1977." Summary 80-5.

Table 2.5 **Employee Benefits in Large Companies by Industry**

Industry	Compensation per Workhour, C^*			Wage to Compensation Ratio, (W/C^*)			Paid Leisure Ratio ϕ		
	1967	1973	1979	1967	1973	1979	1967	1973	1979
Total, all industries	4.009	6.479	10.386	77.1	72.7	70.4	10.6	12.7	13.0
Total manufacturing	4.013	6.231	10.231	76.2	73.3	70.2	10.3	12.0	12.4
Food	3.803	6.017	9.520	75.9	72.7	70.5	11.9	13.2	11.7
Textiles/apparel	2.502	3.989	6.172	88.3	85.8	83.5	6.9	8.7	9.2
Paper/lumber	3.662	5.668	9.494	80.1	75.5	70.8	8.9	10.2	12.3
Printing	4.466	6.470	10.264	78.2	75.8	71.6	10.4	11.3	11.8
Chemicals	4.337	6.869	11.951	73.7	69.4	66.0	13.6	15.9	16.2
Petroleum	5.231	7.692	14.867	73.4	68.3	65.4	15.1	17.3	15.9
Rubber/leather/plastic	3.740	6.112	8.906	77.2	72.9	71.6	10.8	12.4	11.9
Stone/clay/glass	3.611	5.960	9.592	78.2	72.0	70.7	10.3	13.2	12.1
Primary metals	4.344	6.617	10.931	75.8	72.5	66.9	10.3	10.9	13.1
Fabricated metals	3.913	6.341	9.637	78.9	73.1	70.8	9.3	11.6	11.8
Machinery	4.028	6.471	10.619	77.9	73.4	70.4	10.2	11.9	12.1
Electrical machinery	3.746	6.112	10.454	78.2	74.2	70.2	10.9	12.1	13.6
Transportation equip.	4.247	6.676	11.751	77.5	71.8	68.9	10.9	13.3	13.6
Instruments	3.694	6.103	9.738	77.9	73.1	70.8	10.9	12.6	11.9
Total nonmanufacturing	4.141	6.796	10.619	75.9	71.8	70.8	11.3	13.7	13.6
Public utilities	4.650	7.530	13.012	65.3	70.9	67.5	12.4	15.0	15.8
Department stores	2.703	4.547	6.679	80.0	76.2	73.7	8.6	10.3	11.2
Trade (wholesale/retail)	3.253	5.202	8.460	79.7	75.7	75.8	8.6	10.3	9.8
Banks/finance	3.899	6.447	9.553	73.1	69.8	68.5	11.8	14.0	14.2
Insurance	4.136	6.494	10.078	75.2	71.0	69.2	12.1	14.0	14.1
Hospitals	x	x	7.622	x	x	75.6	x	x	12.4
Miscellaneous industries	4.473	7.208	11.628	79.3	76.2	71.4	9.6	11.8	13.9
Mean (20 industries)				77.19	73.52	70.32	10.68	12.50	17.81
Std. dev.				4.29	3.69	2.76	1.86	2.08	23.15

Source: U.S. Chamber of Commerce. 1981. "Employee Benefits: Historical Data, 1951–79." Washington, D.C.

Table 2.6 Employee Benefits in Large Companies, 1951–1979 (in current dollars)

Year	Comp. per Workhour C^*	Wages to Comp. (percent) W/C^*	Paid Leisure (percent) ϕ	Legal Req. Benefits (percent) B_{LR}/B	Wage Index W_{LC}/W_{PS}
1951	2.024	83.1	8.0	18.4	115.9
1953	2.179	82.7	8.2	16.8	111.9
1955	2.359	81.9	8.5	17.6	112.9
1957	2.696	80.6	9.0	17.7	115.0
1959	3.010	79.9	9.6	18.2	119.0
1961	3.159	78.3	10.2	20.1	115.6
1963	3.456	77.8	10.5	22.1	117.9
1965	3.691	78.4	10.2	19.9	117.7
1967	4.010	77.1	10.6	22.3	115.3
1969	4.630	76.1	11.2	22.7	115.9
1971	5.369	74.0	12.5	20.6	119.6
1973	6.479	72.7	12.7	23.1	119.6
1975	7.697	71.0	13.7	22.7	120.5
1977	8.779	70.3	13.2	23.1	117.5
1978	9.550	70.1	13.4	24.3	117.7
1979	10.387	70.4	13.0	24.6	118.7

Source: U.S. Chamber of Commerce. 1981. "Employee Benefits: Historical Data, 1951–79." Washington, D.C.

fringe benefits, reveal that males in large firms are more likely to get these fringes, implying lower W/C^* ratios.

Several explanations can be offered for the secular growth in fringe benefits: (a) Leisure is a normal good, and as real incomes grew, individuals demanded more leisure in the form of paid vacations rather than in shorter workdays. (b) The higher marginal tax rates which accompany higher wages reduce the net "after-tax" prices of certain fringes. (c) Pensions and deferred pay account for a larger share of the compensation of salaried employees and of workers in large firms who are likely to be more firm specific. More generous but less portable fringe benefits reduce labor turnover, thereby increasing the returns to specific human capital. In addition, deferred pay discourages malfeasance and shirking, thereby reducing monitoring costs.[35] The cross-sectional differences in the wage to pay ratio are broadly consistent with the fixed cost hypothesis. Specifically trained workers in large firms and high-wage industries are provided with compensation packages that put more pay in the form of pensions.

The W/C^* ratio has clearly declined over time and varies across firms and industries. Empirical studies that fail to recognize the changing relation of W to C^* could contain serious biases.[36] The increasing impor-

Table 2.7 **Percentage of Employees Receiving Selected Benefits (all industries by firm size, May 1979)**

Firm Size	Hourly Wage	Group Health	Pension Plan	Disability Insurance
		All Employees		
1–24	4.90	34.1	21.1	8.2
25–99	5.61	64.8	48.1	20.2
100–499	6.26	76.4	70.8	33.0
500–999	6.36	80.0	80.1	40.8
1,000+	7.33	85.6	88.5	55.1
Total	6.23	67.4	61.9	33.7
		Male Workers		
1–24	5.63	41.9	23.2	11.0
25–99	6.58	73.4	52.8	26.0
100–499	7.43	83.7	73.3	41.1
500–999	7.73	87.9	82.0	50.5
1,000+	8.49	91.7	91.4	63.7
Total	7.34	75.7	66.4	41.8

Source: Current Population Survey, May 1979, unpublished data.

tance of deferred pay suggests that the labor input is becoming more firm specific. Long-term contracts and quasi-permanent jobs are evidently assuming greater importance in the labor market.

The organization of production and the composition of employment are obviously influenced by industrial affiliation. Within an industry, the Census of Manufactures data show that the ratio of production to nonproduction workers is inversely related to establishment size. Additional empirical regularities are revealed by data from the May 1979 *Current Population Survey* (CPS). Selected characteristics of all employed persons by firm size and sex are shown in table 2.8, and similar data for male employees in manufacturing appear in table 2.9. The attributes identified in these tables are familiar, and I shall remark on only some of them.

Education. Larger firms demand more highly educated persons who have already demonstrated their capacity to absorb training. The relation is stronger for males and for salaried workers. In fact, there is almost no relation between firm size and years of schooling for hourly male workers in manufacturing.[37] The patterns conform to the monitoring cost hypothesis in which more productive workers are matched with more able entrepreneurs.

Race, sex, and city size. Nonwhites accounted for 9.5 percent of employed persons. The percentage varies across industries, but there is no

systematic relation to firm size. Females, who make up 44 percent of employment, are more heavily represented in smaller firms because of their propensity to hold part-time jobs. The percentage of workers in cities with a population of a million or more is unrelated to firm size.

Part-time employment. Some part-time jobs are permanent, but many are staffed by temporary workers who require more supervision. In all industries, 34.0 percent of employees in small firms held part-time jobs, and this falls to 11.7 percent in the largest firms. Differences in monitoring costs could have been responsible for this inverse relation which is observed within an industry.

Age and job tenure. Larger firms have older workers, and the relation is stronger for males. Table 2.9 reveals, however, that the mean age of male production (hourly) workers varies little across size categories. The striking relation is that between years of job tenure and firm size. The mean duration rises from 4.04 years in small firms to 8.68 years in large firms. Longer job tenure generally corresponds to lower labor turnover rates which squares with the hypothesis that workers in large firms have more firm-specific human capital.[38] Males and salaried workers who receive more training have longer job tenures.

Hourly wage rates. Several economists have observed that wages are positively related to firm size.[39] As one moves from small to large firms (<25 vs. 1,000 + employees), the average hourly wage in 1979 climbed from $4.897 to $7.327 for an unadjusted differential of 49.6 percent.[40] Workers at small firms are more likely to be females, hold part-time jobs, and have less job tenure and education. When Mellow (1981) included personal characteristics in a log-linear wage equation, the estimated firm size differential fell to 24.7 percent. The inclusion of union membership further reduced the differential to 14.3 percent.[41]

Capacity utilization and shift work. When continuous production is dictated by technology, workers must be hired for around-the-clock operations. However, in the vast majority of industries, firms can choose the length of the workweek of fixed capital by varying the number and length of shifts. The equilibrium capital utilization rate will be higher the larger the share of costs attributable to capital, and the smaller the wage differential for shift work. According to Foss (1981), the workweek of fixed capital varied from a high of 140 hours in petroleum refining to a low of 42 hours in apparel. Instead of the usual division into durables and nondurables, I separated manufacturing into two sectors by using Foss's estimates of capital utilization rates.[42] Some 81.5 percent of manufacturing employees worked on the day shift, and the percentage was higher for females and salaried workers. Nearly a fourth of employees in high-use

Table 2.8 **Characteristics of Employees in All Industries by Firm Size: May 1979**

Characteristic	In Firms with an Employment of					
	Total	1–24	25–99	100–499	500–999	1,000+
	All Employees					
No. of workers	17,301	4,548	2,521	2,479	951	6,802
Age	36.44	35.33	36.57	36.59	36.29	37.09
Job tenure	6.51	4.04	5.18	6.30	6.80	8.68
Hourly wage	6.231	4.897	5.606	6.258	6.358	7.327
Education	12.66	11.94	12.30	12.91	13.25	13.10
Percentage of workers:						
Female	43.8	48.2	44.0	47.2	50.5	38.5
Nonwhite	9.5	9.5	8.7	9.2	9.2	10.1
Part-time	19.4	34.0	20.0	14.2	17.4	11.7
Union	22.7	5.9	17.3	25.5	28.3	34.2
Pension plan	61.9	21.1	48.2	70.8	80.1	88.5
Large cities	35.0	28.7	33.2	34.3	34.0	40.3
Hourly workers	55.4	57.7	58.3	58.3	53.8	53.7
	Male Workers					
No. of workers	9,731	2,357	1,411	1,310	471	4,182
Age	36.80	34.26	36.41	37.01	36.91	38.28
Job tenure	7.62	4.34	5.80	7.26	7.85	10.18
Hourly wage	7.340	5.628	6.583	7.426	7.729	8.490

Education	12.7	11.8	12.1	12.9	13.5	13.2
Percentage of workers:						
Nonwhite	8.3	8.7	8.2	8.3	7.9	8.2
Part-time	9.2	17.8	9.4	7.0	7.2	5.3
Union	28.2	9.1	21.8	30.2	33.8	39.9
Pension plan	66.4	23.3	52.8	73.3	81.9	91.4
Large cities	34.9	28.1	32.7	36.4	35.7	39.0
Hourly workers	53.0	57.1	56.3	50.0	45.4	51.4
			Female Workers			
No. of workers	7,570	2,191	1,110	1,169	480	2,620
Age	36.0	36.5	36.8	36.0	35.7	35.2
Job tenure	5.11	3.72	4.39	5.22	5.76	6.30
Hourly wage	4.805	4.111	4.364	4.949	5.013	5.471
Education	12.6	12.1	12.5	12.9	13.0	12.8
Percentage of workers:						
Nonwhite	11.1	10.4	9.2	10.1	10.4	13.1
Part-time	32.5	51.4	33.3	22.2	27.3	21.8
Union	15.6	2.4	11.5	20.3	39.6	25.1
Pension plan	56.1	18.8	42.3	67.9	78.3	83.9
Large cities	35.1	29.3	33.9	31.8	32.3	42.4
Hourly workers	58.4	58.3	60.8	57.1	62.1	57.3

Source: Current Population Survey, May 1979, unpublished data.

Table 2.9 **Male Employees in Manufacturing: May 1979**

Characteristic	Total	In Firms with an Employment of				
		1–24	25–99	100–499	500–999	1,000+
	Salaried Workers					
No. of workers	1,012	69	103	130	58	652
Age	40.21	36.06	40.15	41.12	42.76	40.26
	(11.87)	(13.98)	(13.10)	(12.80)	(11.74)	(11.16)
Job tenure	10.58	4.84	8.17	9.10	9.47	11.97
	(10.02)	(7.07)	(8.72)	(8.83)	(11.31)	(10.25)
Hourly wage	9.864	7.713	8.500	8.849	9.905	10.506
Education	13.99	12.41	12.90	13.62	13.97	14.41
Percentage of workers:						
Nonwhite	7.1	10.1	8.7	4.6	10.7	7.4
Part-time	3.0	20.2	5.8	3.0	5.1	0.6
Union	9.1	10.1	8.7	13.1	6.9	8.6
Pension plan	84.6	21.7	48.5	78.5	93.1	97.5
Large cities	43.3	46.3	37.9	42.3	32.8	45.1

	Hourly Workers					
No. of workers	1,781	154	233	251	93	1,050
Age	36.74	32.81	36.93	36.65	34.70	37.48
	(13.21)	(14.61)	(15.45)	(13.51)	(11.19)	(12.43)
Job tenure	8.75	3.92	5.89	7.54	8.13	10.44
	(9.32)	(6.09)	(8.11)	(8.76)	(7.89)	(9.75)
Hourly Wage	6.757	5.316	5.727	6.104	6.407	7.385
Education	11.27	10.69	10.79	10.79	11.31	11.57
Percentage of workers:						
Nonwhite	9.0	9.7	9.4	7.6	7.5	9.2
Part-time	2.8	9.7	7.7	2.0	2.2	1.0
Union	53.6	15.6	25.8	43.0	55.9	67.7
Pension plan	77.9	26.6	45.9	70.9	87.1	93.4
Large cities	33.5	46.1	33.0	38.2	27.9	31.1

Source: Current Population Survey, May 1979, unpublished data.

NOTE: Numbers in parentheses are standard deviations.

manufacturing worked on late/night shifts, and 14.2 percent in low-use manufacturing. There is a strong positive association between the percentage on late/night shifts (the complement of the percentage on day shifts) and firm size as evidenced by the data in table 2.10. It rises from 6.3 to 27.9 percent in high-use manufacturing and from 5.7 to 18.5 percent in low-use. The higher incidence of shift work in large firms is a consequence of the decision to engage in the volume production of standardized goods.

Firm size is systematically related to differences in wages, the organization of production, and the composition of employment. The received theory of labor markets acknowledges the presence of heterogeneous workers, but we still cling to Marshall's concept of a "representative firm." The latter convention must be abandoned to explain the empirical regularities exhibited in tables 2.8, 2.9, and 2.10.

2.3.5 Monitoring and Fixed Employment Costs in a World of Heterogeneous Firms

The uniqueness of a *firm* is essential for the existence of specific human capital. It is meaningless otherwise to speak about training that raises a worker's productivity in *one firm*, while leaving his productivity unchanged in other firms. If that is a true result, that *one firm* must have been different from all others. Fortunately, firms are heterogeneous, and profits can be increased by incurring those fixed employment costs that are needed to recruit, train, and retain a firm-specific labor force. Firms can differ in many ways, and the dimension which I have emphasized is that of entrepreneurial ability. Able entrepreneurs have the capacity to convert calendar time into larger supplies of managerial effort which allow them, through the usual law of variable proportions, to assemble large production teams. Following Lucas (1978), I assumed a background distribution of entrepreneurial abilities, $\psi(\lambda)$, which yields a critical ability level λ_0 such that if $\lambda < \lambda_0$, the individual does *not* become an entrepreneur. The relative frequencies of entrepreneurs of moderate abilities [λ just slightly greater than λ_0 in the truncated distribution of $\psi(\lambda)$] will be large, resulting in numerous small firms bearing a close resemblance to their neighbors. The rare, high-λ entrepreneurs from the extreme right tail of $\psi(\lambda)$ assemble very large corporations that are few in number. These giant firms are almost unique.

Entrepreneurs enter different industries, and in each industry competition produces an equilibrium size distribution of firms.[43] Employment is an imperfect but readily available measure of size. Small firms are defined here as those with less than twenty-five employees, while large firms have one thousand or more workers. Firm size distributions vary across industries. In all industries, 26.3 percent of total employment was located in

small firms and 39.3 percent in large. The corresponding figures in manufacturing were 8.3 and 57.8 percent.

Talented entrepreneurs can economize on the time losses to monitoring and supervision in several ways. Production can be organized around teams and units, and authority can be delegated by stacking these teams in a hierarchical structure. The standardization of products and procedures (exemplified by the assembly line) limits the opportunities for discretionary behavior thereby simplifying the monitoring task. Capital is intensively utilized by operating multiple shifts. This mode of operation is profitable because of the high ratios of specialized capital to labor which accompany volume production.[44] Further, the planning of production is extended to recruiting and personnel management. Job descriptions and task assignments are spelled out in great detail. Applicants are screened and tested to determine if they can meet the rigid job specifications. Recruiting costs are higher in large firms which will hold job vacancies open until a suitable candidate is found. More highly educated persons are recruited to staff the salaried, firm-specific positions, as evidenced by data in the top panel of table 2.9. Specific training is provided to new employees to adapt them to the firm's prescribed operating practices. These recruiting and training activities "produce" a labor force of reasonably homogeneous individuals, and such teams are obviously easier to supervise. In a sense, higher fixed employment costs can be substituted for lower subsequent monitoring costs. Those firms that incur high fixed costs have a strong incentive to design selection and compensation practices that reduce the turnover of specifically trained workers. If firms are successful in retaining specifically trained workers, data on job tenures tell us something about the relative importance of firm-specific human capital.[45] Job tenures are longer (suggesting larger investments in firm-specific human capital) for persons who are males, are more highly educated, work at large firms, hold salaried positions, and are employed in industries which exhibit higher capital utilization rates.

Less able entrepreneurs command smaller teams and produce customized goods.[46] They spend less on recruiting because jobs are flexible and work can be redesigned to fit the individual applicant. Small firms have shorter life expectancies and hence have less to gain from firm-specific training and specialized durable machinery. Fixed employment costs are small, and labor costs are almost entirely composed of wages that must be paid to attract general human capital from competing employments.

The behavioral differences among firms are confounded by the presence of unions. Some 22.7 percent of all employed persons were members of trade unions or employee associations, and the membership climbs to 34.2 percent of employees in large firms.[47] On the supply side, there are likely to be scale economies in organizing workers, collecting

Table 2.10 **Wages, Age, and Job Tenure in Manufacturing: May 1979 (by firm size)**

A. High-Use Manufacturing

Characteristic	In Firms with an Employment of					
	1–24	25–99	100–499	500–999	1,000+	Total
	All Employees					
No. of employees	126	163	237	89	1,223	1,838
Percent day shift	93.7	89.6	79.3	77.5	72.1	76.3
Hourly wage	5.256	6.069	6.482	6.997	8.012	7.405
Age	35.23	37.94	37.46	46.97	38.08	37.74
Job tenure	4.15	6.96	7.67	7.63	10.68	9.37
Percent female	40.5	34.4	30.0	25.8	20.9	24.8
	Salaried Male Workers					
No. of employees	18	29	60	25	338	470
Percent day shift	88.9	100.0	93.3	88.0	87.6	89.1
Hourly wage	7.924	9.774	8.998	10.361	10.855	10.413
Age	37.44	44.93	40.52	42.32	40.55	40.79
Job tenure	4.78	12.90	10.88	9.44	12.50	11.86
	Hourly Male Workers					
No. of employees	57	78	106	41	630	912
Percent day shift	89.5	83.3	72.6	64.4	61.6	66.8
Hourly wage	5.422	5.902	6.084	6.517	7.552	7.061
Age	32.23	35.49	35.31	35.10	37.84	46.87
Job tenure	3.11	6.19	7.54	7.32	11.00	9.52

B. Low-Use Manufacturing

Characteristic	In Firms with an Employment of					
	1–24	25–99	100–499	500–999	1,000+	Total
			All Employees			
No. of employees	212	358	387	145	1,132	2,234
Percent day shift	94.3	90.2	90.4	82.8	81.5	85.8
Hourly wage	5.423	5.449	5.709	6.075	7.185	6.412
Age	35.33	38.16	38.65	34.96	37.65	37.51
Job tenure	4.36	5.41	6.76	7.27	9.21	7.59
Percent females	30.2	35.8	44.4	41.4	35.2	36.8
			Salaried Male Workers			
No. of employees	51	74	70	33	314	542
Percent day shift	96.1	90.5	95.7	87.9	92.4	92.6
Hourly wage	7.639	8.001	8.722	9.561	10.131	9.389
Age	35.57	38.28	41.61	43.09	39.95	39.71
Job tenure	4.86	6.31	7.57	9.48	11.40	9.48
			Hourly Male Workers			
No. of employees	97	155	145	52	420	869
Percent day shift	90.7	85.8	83.4	75.0	69.8	77.6
Hourly wage	5.254	5.639	6.119	6.320	7.134	6.440
Age	33.15	37.66	37.63	34.38	36.94	36.61
Job tenure	4.40	5.74	7.54	8.77	9.61	7.94

Source: Current Population Survey, May 1979, unpublished data.

NOTE: The two-digit industries included under high use were industries 29, 33, 26, 28, 22, 30, 32, 21, 37, and 27. The remaining industries were put into low use. See note 42 of text.

dues, and enforcing compliance. The demand for unionism is also likely to be greater in larger firms for at least three reasons: First, a union may be the preferred institution to supply certain services of a public goods nature (e.g., grievance procedures, negotiating better working conditions, or resolving conflicts). Second, a union that can restrict membership may be able to raise wages above competitive levels if the host firm is earning economic rents. Third, big firms in small labor markets may have some monopsony power. The data of tables 2.8 and 2.10 indicate that the incidence of unionism is indeed positively correlated with firm size.

Some interesting interactions are observed in table 2.11 which presents data for male production workers in manufacturing classified by firm size, job tenure, and union membership. The incidence of unionism was 58.3 percent for the entire sample, and by firm size groups, the percentages in unions were 21.9 small, 46.5 medium, and 67.7 large. New employees with less than one year of job tenure accounted for 17.1 percent of employment; this measure of the annual accession rate varied from a high of 35.0 percent in small, nonunionized firms to a low of 8.7 percent in large unionized firms. Holding firm size constant, new employees are less likely to be assigned to jobs on the regular day shift. The relative frequency of shift work which reflects the firm's capital utilization rate is considerably higher in the union sector, but the reason for this is unclear. High capital utilization rates may be associated with larger union wage gains, or trade unions may be better able to supply the services and contractual arrangements demanded by employees on late/night shifts.

Wages of the blue-collar workers in table 2.11 are positively related to firm size and job tenure. The percentage wage gains due to size and tenure are larger in the nonunion sector, but in each size/tenure cell, unionized workers were uniformly better paid.[48] However, wage differences understate the differentials in total employee compensation because union workers receive proportionally more in fringe benefits. Although collinearity makes it difficult to disentangle firm size and union effects, the data of table 2.11 and the results reported by Mellow (1981) support the conclusion that other things equal, workers in large firms are paid higher wages and receive more fringes. These higher wages may contain elements of economic rents or compensating differences for working conditions, or they may simply represent the equilibrium payments to superior employees whose higher productivity cannot be linked to observable traits. The latter interpretation is in line with the monitoring cost hypothesis in which able entrepreneurs are matched with more productive workers.

Based on data from two longitudinal surveys, Mincer (1981) reported that the wage-experience profile was flatter for unionized workers. Union members claimed that they got less "training" at their jobs which could

account for the flatter profile. However, an employee's response to a survey question is not a reliable basis for estimating the amount of specific training that he has received. Further, the wage rate is an imperfect measure of pecuniary returns. A regression equation in which wages of employed persons are related to job tenures will yield a smaller slope coefficient relative to a second regression in which expected wages are related to job tenures.[49] Both regressions will understate the pecuniary returns to job experience. Unionized workers are covered by collective bargaining agreements and implicit long-term contracts whose compensation packages exhibit the property that the wage to pay ratio (W/C^*) falls with increasing job tenure; i.e., the value of a union member's claims to pensions and other fringe benefits rises with seniority. Expected pay per workhour is a better measure of pecuniary returns. We have to estimate regressions using an expected pay variable to determine if the experience profile of pecuniary returns is really flatter for union members.

Finally, it has been observed that unionized workers receive a larger fraction of *pay* in the form of pensions and fringes. Mincer (1981) attributed this magnification of the demand for fringes to the fact that union members are more highly paid and hence face higher marginal tax rates. If true, the ratio of fringe benefits to wages should be larger in states that have higher state income tax rates. Alternatively, it can be argued that fringe benefits and deferred pay are incorporated into compensation packages of those firms (union and nonunion) that are trying to retain their specifically trained workers. Reference to the left panel of table 2.11 reveals that in the nonunionized sector, the percentage of employees who are eligible for pension plans is positively related to firm size. The phenomenon is not unique to the unionized sector.

The data on job tenures are consistent with a model in which specific training assumes greater importance in larger and unionized firms. A tenured worker is defined here as one who has been with his current employer for five or more years. Table 2.11 reveals that 55.4 percent of male production workers in manufacturing were tenured, while 71.7 percent of employees in large unionized firms were tenured. For nonunion workers, the mean duration of job tenure was 2.7 years longer in large versus small firms, and the difference is due to higher retention rates during the first five years. Once workers pass beyond the five-year point, the conditional mean duration of job tenure is unrelated to firm size.[50] The job tenure differential between small and large unionized firms is 5.9 years, and marginal retention rates appear to be higher in large firms at each year of service point. If years of job tenure are subtracted from the mean age, we can derive the mean age of workers at entry. This exercise reveals that in large firms, tenured workers were recruited at younger ages.[51] Those individuals who obtain tenure at small firms evidently hold

Table 2.11 **Characteristics of Hourly Male Workers in All Manufacturing: May 1979 (by union status, job tenure, and firm size)**

Characteristic/ Firm Size	Nonunion with Job Tenure of				Union with Job Tenure of			
	<1	1–5	≧5	Total	<1	1–5	≧5	Total
No. of employees								
1–99	106	107	90	303	19	27	39	85
100–999	36	77	71	184	21	34	105	160
1,000+	65	102	172	339	62	139	510	711
Percentage on day shift								
1–99	80.2	86.9	95.6	87.1	73.7	85.2	94.9	87.1
100–999	66.7	81.8	88.7	81.5	52.4	55.9	81.0	71.9
1,000+	69.2	68.6	75.6	72.3	41.9	51.1	66.5	61.3
Hourly wage								
1–99	4.571	4.964	5.691	5.042	5.099	6.717	7.813	6.858
100–999	4.809	5.481	6.504	5.744	6.029	6.353	6.701	6.539
1,000+	5.106	6.226	7.077	6.443	6.908	6.982	7.795	7.559

Percentage on pension plans								
1–99	25.4	25.2	38.9	29.4	36.8	70.4	84.6	69.4
100–999	41.7	57.1	74.6	60.9	81.0	91.2	94.3	91.9
1,000+	80.0	84.3	96.5	89.7	90.3	92.1	96.7	95.2
Age in years (mean and standard deviation)								
1–99	29.06	30.07	46.14	34.49	25.84	32.00	48.05	37.99
	(13.42)	(12.58)	(14.02)	(15.29)	(9.17)	(12.19)	(11.79)	(14.77)
100–999	24.97	31.95	42.73	34.74	27.76	28.09	42.82	37.71
	(8.69)	(11.81)	(12.36)	(13.33)	(8.34)	(8.25)	(11.00)	(12.33)
1,000+	26.88	31.27	39.79	34.75	29.11	29.39	42.52	38.78
	(10.31)	(10.65)	(11.36)	(12.16)	(10.73)	(9.27)	(11.23)	(12.35)
Job tenure in years (mean and standard deviation)								
1–99	—	2.36	12.92	4.67	—	2.30	12.77	6.59
		(0.98)	(8.52)	(7.18)		(1.17)	(8.27)	(8.10)
100–999	—	2.31	12.62	5.84	—	2.71	14.12	9.84
		(1.05)	(7.63)	(7.24)		(1.06)	(8.90)	(9.37)
1,000+	—	2.25	13.22	7.38	—	2.35	15.95	11.90
		(1.01)	(8.78)	(8.67)		(1.06)	(8.83)	(9.91)

Source: Current Population Survey, May 1979, unpublished data.

several jobs before they find a suitable match with a viable firm that will remain in business for fifteen or more years. It would be interesting to discover whether the older tenured workers at viable small firms had received more or less firm-specific training in relation to tenured workers at large manufacturing firms.

2.4 Concluding Remarks

The labor market for the economy as a whole is populated by a wide diversity of workers and firms. The entrepreneurs who control small firms confront a lower shadow price of time which gives them a comparative advantage in monitoring worker performance, coping with disruptions and high labor turnover rates, providing maintenance for used equipment, and haggling over frequent changes in prices and wages. Their production teams include less experienced workers and more part-time employees. They apparently lack the organizational ability to operate multiple shifts. They choose to supply products where technology discourages standardization and volume production. Specialization and specific training might be profitable if the firm could be assured of its survival and its ability to retain specialized resources. But such assurances cannot be supplied, even by government regulation. Some workers obviously dislike the discipline and rigidity of employment at large firms and choose to work for small employers. They receive little or no specific training, are paid lower wages, and get relatively few fringe benefits. A relatively small number of individuals form permanent attachments with viable small employers.[52] However, most workers in this part of the labor market possess general human capital that can readily be shifted to numerous small firms that are only slightly differentiated from one another. The uncertain and possibly short lives of these firms reduce the returns to specific investments, but this uncertainty is apparently not the source of employment instability.[53] The important fact is that variable wage payments comprise almost all of the full costs of the labor input. A neoclassical model in which labor's marginal value product is equated in each period to the market wage rate describes the behavior of firms and workers in the portion of the labor market populated by "small firms."

The neoclassical model has been replaced by a loosely knit theory in which the labor market is characterized by implicit long-term contracts, rigid wages, formal layoff policies, lifetime tenured employment, deferred pay, and mandatory retirement. Fixed employment costs are, according to Hall (1980), the glue that binds workers and firms together. The authors of the new labor economics recognize that firm-specific investments in recruiting and training are endogenous, but they largely ignore this endogeneity. The existence of these fixed costs is simply

assumed to focus attention on the implications of a quasi-fixed labor input.

The thesis advanced in this paper is that recruiting and training only make sense in a world of heterogeneous, differentiated firms. The heterogeneity in my model is generated by a distribution of entrepreneurial abilities. The outliers succeed in building very large corporations that are few in number and are spread across industries. Each very large firm is nearly unique. Managerial efforts are directed to the development of standardized products and the organization of integrated but inflexible production lines. Companies make large investments in recruiting and training firm-specific labor forces. Workers in large firms are paid higher wages as well as compensation in the form of pensions and fringes that are designed in part to reduce the turnover of specifically trained employees. The full cost of the labor input is thus the sum of total employee compensation, the amortization of fixed employment costs, and the implicit costs of monitoring worker performance. Since wages represent only a part of full labor costs, they are unlikely to be frequently adjusted in response to short-run changes in demand. The prices quoted by large firms also tend to be rigid because every price change has to be closely supervised to prevent chiseling and cheating by numerous subordinates who staff a complex distribution network. Price and wage rigidities may have been responsible for more employment instability and may also have contributed to an increased demand for outside representation of workers by organized labor unions.[54] The specificity of the labor input, personnel management, and the organization of production along rigid assembly-line techniques have surely been influenced by the shadow price of the entrepreneurial input. The structure of each firm is rationally determined to maximize profits in a world where there are trade-offs between monitoring costs and the fixed costs of specialized resources. The portion of the labor market in which we find large firms is described by the perceptive picture painted by Hall (1982). However, it is a picture that applies to only a part of the economy, albeit an important part containing at least 40 percent of total employment.

Specific human capital has proven to be an important concept in the theory of labor economics. However, firm-specific capital can only be produced and employed in a segment of the economy occupied by very large firms. The predictions of a theory that embraces this concept can be borne out by the empirical evidence if these large firms account for a dominant share of the aggregate labor market. The empirical studies of the last two decades suggest that this is indeed the case. But there is another important sector of the labor market where there is little room for specialized labor. The caricatures of the large and small firms in these two sectors are like Marshall's "representative firm." They are simply

analytic prototypes picked from a continuum of firm sizes, production adaptability, and labor specificity. Firms and workers are indeed heterogeneous. A theory of labor economics that explicitly acknowledges this heterogeneity and incorporates it into its analytic models can, I believe, add considerably to our understanding of the behavior of labor markets.

Notes

1. Details of the theory can be found in Oi (1962) and Becker (1964). Reder (1955) offered an alternative theory in which the cyclical changes in occupation wage differentials were explained by countercyclical variations in hiring standards that result in the upgrading and downgrading of employees. The Reder model must be tied to a theory of factor demand shifts in order to explain employment responses.

2. Hutt (1977) identified six categories of idleness: (1) valueless resources, (2) pseudoidleness which is defined as a state in which the capital value of an asset exceeds its scrap value even though its net hire value is nil, (3) preferred idleness exemplified by the labor-leisure choice, (4) participatory idleness in arrangements to share monopoly rents, (5) enforced idleness due, for example, to legal limits on workhours, and (6) withheld capacity to obtain monopoly rents. Idleness of the first three types constitutes an efficient allocation of resources.

3. Additional evidence and citations to the literature can be found in Clark and Summers (1979), pp. 53–54.

4. This model helps to explain academic tenure; see Oi (1979). It also provides an explanation for mandatory retirement which is more fully analyzed by Lazear (1979).

5. The rationale for this argument was developed by Becker and Stigler (1974). Deferred pay can be viewed as an alternative to bonding, which puts the Becker-Stigler model in the spirit of the principal-agent literature.

6. I was introduced to the Koike article by Hall (1982). In addition to his discussion on job tenures, Koike points out an important difference in trade union behavior. The employment agreements negotiated by Japanese unions contain *no* seniority rules for layoffs and recalls.

7. Total compensation does not vary in direct proportion to man-hours because some components (e.g., disability and health insurance) are linked to the number of employees rather than man-hours. The nonlinearities in the relation of compensation to man-hours can be put into the fixed employment costs.

8. A specially designed machine or plant can be purchased on a "made-to-order" basis. However, when such inputs are demanded on a regular basis, the firm is likely to engage in vertical integration to control the source of supply. Klein, Crawford, and Alchian (1978) point to the concept of "postcontractual opportunistic behavior" to justify vertical integration in cases where the situation could result in a bilateral bargaining game. The principle is illustrated by the acquisition of Fisher Bodies by General Motors.

9. The durability of firm-specific human capital is jointly determined by the expected job tenure of a trained worker and the rate of technical obsolescence/depreciation applicable to such capital. Rapid changes in technology and in product demands increase the obsolescence rate, thereby reducing the durability of specific human capital.

10. Rosen (1972*b*) constructed a model in which inputs of labor L and firm-specific knowledge Z produced two joint products in fixed proportions: output Q and an increment to knowledge ΔZ. Knowledge is a permanent, nontransferable asset which produces a volume effect similar to the one examined by Alchian (1959). In a second model, Rosen

(1972*a*) assumed that each multiproduct firm supplied its workers with work plus training that added to the workers' general human capital. Employees recognized the composition of the tied package and were thus willing to accept lower wages which reflected the implicit market value of general human capital. Training and work (the production of goods) are presumably tied because of some unspecified economies of joint production. The economies argument is more plausible when the training is firm specific.

11. Suppose that the productivity of the ith individual, μ_i, depends on inputs of general and specific human capital, $\mu_i = \mu(G_i, S_i)$. If μ is homogeneous of the first degree, $\partial\mu/\partial S$ is a function of the ratio (G/S). The marginal product of S is equalized when $(G_i/S_i) = (G_j/S_j)$ for all i, j. If market wages are proportional to general human capital, an optimal allocation of specific human capital will result in more S allocated to individuals with more G. However, the degree of fixity will be a constant.

A positive relation between the wage rate and the degree of fixity can be derived by relaxing the assumption of first-degree homogeneity or by introducing additional arguments into the μ function.

12. Alfred Marshall and his followers were mainly interested in issues of allocative efficiency across commodities which were equated to industries. The concept of a "representative firm" was sufficient for this purpose, but it left little room for heterogeneity. The theory was mainly concerned with the determinants of an optimum (equilibrium) firm size that could be reconciled with two maintained assumptions: (1) first-degree homogeneity of the production function and (2) perfect competition in factor markets. Kaldor (1934) obtained a determinate firm size by appealing to the fixity in supply of at least one input. The proposition that firm size is ultimately limited by a scarcity of the entrepreneurial input was also advanced by Robinson (1958), Georgescu-Roegen (1967), and Friedman (1976). I shall also invoke this same assumption.

13. Capital is assumed to require *no* monitoring. In the adjustment cost model of Lucas (1967), the installation of new capital entailed an opportunity cost of forgone output. This is not the same as monitoring. In his distinction between man and machine, J. Clark (1923) pointed out the importance of monitoring and metering worker performance:

> Having learned one way of doing a thing, a worker tries variants on it, sometimes with a purpose, sometimes aimlessly, but always following the bent of "monkeying." . . . He is very imperfectly adapted to continuous toil and when he does work, he works now faster and now slower with an irregular rhythm. . . . Especially when working for a purely collective end, his ardors while often strong appear to be characteristically intermittent and unreliable. As a class, he needs personal incentives to work, rewards for good performance, and penalties for bad, more immediate and substantial than his share in the welfare of the whole industry or the whole community. (p. 8)

14. The shadow price of an efficiency unit of managerial effort is Pf_T, but a unit of calendar time yields λ efficiency units so that entrepreneurial time has a shadow price of $P\lambda f_T$. Since h hours are required to monitor each worker, we get the implicit monitoring cost, $\delta = P\lambda h f_T$.

15. Homogeneity implies that $Q = Kf_K + Nf_N + Tf_T$. Substitution into the expression for profits yields $\pi = P(M\lambda h f_T + Tf_T)$. Equation (5) is obtained by recalling that $T = \lambda(\bar{H} - hM)$.

16. If wages are proportional to productivity, $W(\mu)/\mu$ will be a constant. In this event, $[\mu W'(\mu) - W(\mu)] = 0$, and in the presence of positive monitoring costs, all firms will try to substitute higher quality for fewer numbers. The resulting increase in demand for more productive workers will raise their wages. Hence, in equilibrium, $W(\mu)$ must be convex so that $W''(\mu) > 0$.

17. Friedman (1976) explicitly acknowledged the presence of heterogeneous firms in his analysis of the relation of the firm to the industry. He examined a case in which firms with different cost curves had to adjust to an increase in the demand facing the industry. If the increase in product demand raises the price of a factor that is specific to the industry, the

quasi-marginal cost curve for an exceptional firm could bend backward; i.e., the exceptional firm contracts output in response to an increase in product price. Friedman did not try to explain the reasons for different cost curves. In my model, a dispersion of entrepreneurial abilities generated cost differences wherein more able entrepreneurs enjoyed lower marginal cost curves.

18. Tailors, punch press operators, and door-to-door salesmen are often paid by "piece rates." Such compensation methods can be implemented when output is easily observed and directly linked to particular employees. If each worker in Adam Smith's pin factory performed all tasks (sharpening the pin and placing the head on it), they could have been paid by the piece, thereby reducing monitoring costs. The gains from specialization and the division of labor evidently outweighed the monitoring cost savings. Payment by results may provide incentives for greater work effort which can explain its adoption in some firms. An analysis of this method of compensation can be found in Pencavel (1977).

19. Rees and Shultz (1970) found, for example, that a secretary's pay was positively related to typing speed, but the relation was nonlinear. Only a small part of the dispersion in pay could be explained by typing speed. A significant part of pay evidently represented compensation for productive activities other than typing. The relation of total productivity to length of job service cannot be determined by observing only one dimension of productivity.

20. Alchian argued that unit costs will decline as a function of the planned volume of output. Wright (1936) observed the same regularity in his study of progress functions in the production of air frames. Oi (1967) explained the progress function in terms of intertemporal factor substitutions and the economies of joint production. Planned volume (batch size) is obviously important in designing the production organization.

21. The critical volume at which an assembly line constitutes the least-cost production method obviously depends on the product. The requisite volumes are likely to be large for goods like pogo sticks, toasters, and bikes, but a volume of ten to twenty oceangoing oil tankers is sufficient to justify the construction of an assembly line.

22. The compensating wage difference will obviously be larger, the larger the size of the team that is asked to conform to the same common schedule and working conditions, and the greater the dispersion in worker preferences. These results are rigorously derived by Deardorff and Stafford (1976). Union workers are typically employed in large firms which provide them with inferior working conditions. According to Duncan and Stafford (1980), the union wage differential cannot be interpreted as simply a monopoly return. Part of it represents a compensating difference that must be paid to attract workers into accepting employment in less desirable work settings.

23. This implication provides yet another reason for the positive association between wages and firm size. Individuals who seek work at large firms will, on average, incur higher search costs because rigid job specifications will not be modified to meet individual worker differences. In equilibrium, larger firms must pay higher wages to compensate employees for the higher expected search costs. The fixed and flex-wage models of Pissarides (1976) incorporate the idea of search on the part of firms.

24. If input is more cheaply monitored, a master carpenter might teach his apprentices certain standardized ways of performing various tasks. When all apprentices use the same work methods, the quality and quantity of output may be more accurately and cheaply gauged by observing the input of apprentice time rather than measuring the flow of output. Investments in entry-level training and screening could thus reduce subsequent monitoring costs.

25. The argument is put as follows: "In slack weeks, hours of work are set at lower levels and the intensity of work may fall as well. The general flavor of the arrangement is that workers work harder when there is much or more work to do." However, in an earlier passage dealing with employment bargains for salaried workers, Hall writes, "Employers have the right to demand intense effort for a few weeks or months but not permanently.

Periods of extraordinary effort must be counterbalanced by restful periods. For salaried workers, arrangements of this kind develop by custom and are rarely spelled out in formal contracts." But why are implicit arrangements preferable to explicit incentives for extra effort? We could appeal to the difficulties of enforcing contingent contracts. There is another weakness in the argument. If "intense effort" can only be demanded for short bursts, we are still left with the puzzle of explaining Okun's law which sometimes applies to adjustment periods extending over several quarters.

26. Gordon and Walton appeal to Stigler's concept of "adaptability" to explain the rapid postwar recoveries in Europe. Wars do not destroy productive factors in balanced proportions. As a consequence, the surviving stocks of capital are not efficiently utilized. A small investment in a particular type of capital that restores the designed factor proportions can lead to a sharp increase in output. The installation of a conveyor belt might, for example, enable a company to make efficient use of its specialized but idle mining equipment. This investment will increase output and the output/capital ratio, but this does not contradict the law of variable proportions as a principle applicable to the ex ante planning of production.

27. With only one technology and two types of labor, the latter effect dominates, and the cyclical adjustments generate a positive correlation between output changes and changes in the output/labor ratio. The pattern is, however, attenuated by the presence of diverse technologies because the standby capacity usually has a lower output/labor ratio. The strong procyclical behavior of labor productivity suggests that the effect of fixed proportions outweighs the influence of diverse technologies.

28. Shinohara (1962) reported that 40 percent of the capital assets of small Japanese firms were purchased as used equipment, compared to only 6 percent for large firms. The flow is evidently one in which new machines tend to be purchased by large firms and, as they age, some are sold to small firms. A similar pattern is observed in international trade. The high-wage, industrialized nations regularly export used durable machinery to less-developed countries. Smith (1974) appeals to differences in relative factor prices to explain the trade in used assets.

29. The present value of the net quasi-rents that can be earned by a durable asset declines with age because of rising maintenance costs. A machine is scrapped when its present value falls below its scrap value. Parks (1979) analyzed the interactions of maintenance, scrapping, and the replacement demand. Grunfeld (1960) provided an early empirical study which emphasized the substitution of maintenance for investment.

30. The tacit assumption here is that capital can be more cheaply transported over time; i.e., increments to capital values depreciate more slowly than increments to inventories of final goods. Internal investments will be biased toward physical capital because the firm is contracting employment in a recession.

31. "Total product" is the sum of the value of final goods plus the implicit value of increments to capital assets. More maintenance during a recession reflects a firm's rational responses to changing relative factor and product prices. When the prices of goods fall in relation to the shadow prices of internally produced investments, the "output mix" understandably shifts away from the production of goods.

32. See, for example, *The Economic Report of the President, 1980*, table B-37, p. 246.

33. In addition to employee compensation, full labor costs must include the amortization of fixed employment costs. If complementary inputs of protective clothing, noise suppressors, and so forth are supplied by the firm, their costs are properly included in full labor costs.

34. I am unaware of any studies that analyze the factors which determine paid leisure time across firms and industries and over time. Paid leisure hours H_L are like the "income in kind" in the British truck system analyzed by Hilton (1957). Paid holidays and vacations place a lower bound on an individual's leisure time consumption. Additional leisure via absenteeism entails a loss of earnings. Rest and recuperation may increase productivity, but these gains redound to the individual. Why do firms impose a lower bound on rest and

recuperation time? I am forced to appeal to an argument like Hilton's in which workers are myopic, and merciful employers nudge them toward the "right" mix of work and leisure.

Banks allegedly insist on vacations so that an employee's books can be audited while he is away. The Rand Corporation gives their employees higher rates of pay when they are on vacation because they *need* more money for travel and lodging. Finally, we have to explain why employees in the public sector and in social service agencies get considerably more paid leisure than workers in the private business sector.

35. Postponing compensation may be an efficient means of controlling executives and public servants. It discourages them from engaging in theft, larceny, and dysfunctional acts. The deterrent effects of deferred pay must, however, be balanced against higher wages that can elicit greater work effort.

The growing demand for private pensions may partially be traced to lower after-tax prices. Discontinuities in the structure of social security benefits may also help to explain the growth. Failure to take early retirement at age 62 is accompanied by a loss in social security wealth. The defined social security benefits may not be enough to warrant full retirement as a utility maximizing choice. If private pensions or savings were available to supplement social security, a worker could avoid the implicit taxation of social security wealth by retiring at age 62. Legislation that raised the defined social security benefits may have prompted the growth in employer contributions to private pensions.

36. The supply of labor is not a function of the wage rate but is, instead, a function of pay per workhour C^*, working conditions, and anticipated future rates of pay. In a demand study, the appropriate "price" should be the full labor cost. The data of tables 2.4, 2.5, and 2.6 indicate that the wage rate is an imperfect measure of pay and probably an even poorer proxy for the full labor cost.

37. Although the CPS asked for the individual's occupation, I classified respondents according to whether they were or were not "paid by the hour." I shall refer to the hourly paid employees as production workers, even though this differs from the census definition of a production worker.

38. In a steady state, the annual turnover rate is equal to the proportion of workers with less than one year of job tenure, but the mean duration is determined by the entire frequency distribution. The correlation between turnover and mean job tenure is thus imperfect, except in the special case where the functional form of the frequency distribution is the same across firms and industries.

39. See, for example, Lester (1967), Masters (1969), and Mellow (1981). Several arguments have been proposed to explain the firm size profile. Lester (1967) and Duncan and Stafford (1980) argued that large firms must pay higher wages which contain a compensating difference for less desirable working conditions. Employees must accept the greater discipline and rigidities of working in large teams. A slightly different argument was proposed by Stigler (1962):

> It is well known that wage rates are less in small plants than in large, and the difference reflects at least in part (and perhaps in whole) the lower cost of the small-scale employer of judging quality. . . . Men should in general enter smaller companies, the greater their ability. (p. 102)

This argument is contrary to a model in which small employers are "small" because they lack the ability to judge and to organize large production teams.

40. The sample means shown in table 2.8 differ from those in Mellow (1981) because I excluded individuals who did not report the timing of work. The timing question was used to determine the frequency of shift work.

41. Mellow estimated two separate regressions. The coefficient of the largest firm size dummy variable fell to .056 in the union regression and to .119 in the nonunion regression. The treatment of unionism as exogenous (either as a dummy variable or as a classificatory variable) can be questioned.

42. The Foss estimates are based on the 1976 Census of Manufactures. The two-digit manufacturing industries wre divided into two groups as follows:

Workweek of Fixed Capital in Manufacturing Industries, 1976

High-Use Manufacturing		Low-Use Manufacturing	
Industry	Hours	Industry	Hours
29. Petroleum	140	35. Machinery	80
33. Primary metals	119	20. Food	75
26. Paper	115	34. Fabr. metals	75
28. Chemicals	112	36. Elec. machinery	73
22. Textiles	110	38. Instruments	69
30. Rubber	108	24. Lumber	58
32. Stone, glass	98	39. Misc. mfg.	57
21. Tobacco	91	25. Furniture	51
37. Transport. equip.	87	31. Leather	44
27. Printing	82	23. Apparel	42

Source: Foss 1981, table 2, p. 9.

Over the period 1929–76, the capital utilization rate in manufacturing rose by 24.5 percent.

43. Although λ was assumed to be an exogenous parameter, it can surely be influenced by economic forces. The process by which a chief executive officer is selected and retained varies across firms. In some cases, an owner-operator begins with a small firm and, through on-the-job experience, he gains the skills to expand the size of his team. In other instances, a candidate may be picked and groomed for the position, which is another way of saying that the firm is investing in specific human capital to raise the value of λ for this candidate. For my purposes, it is sufficient to assume that entrepreneurial ability λ corresponds to firm size.

44. The sparse use of part-time employees by large firms reflects their higher monitoring costs. Disruptions are also costlier when firms adopt inflexible production plans. Large firms enter into vertical integration to avoid unanticipated breaks in the supplies of raw materials. They purchase new as opposed to used equipment. They also invest more in safety to reduce the frequency of industrial accidents. The data on work injury rates examined by Oi (1974) clearly show that work injury risks are substantially lower in the largest establishments.

45. Job tenure is obviously an imperfect proxy for firm-specific human capital. Specifically trained workers ought to remain with their employers for longer periods, but a host of other factors affects the mean duration of job tenure. These include things like the worker's age, the cyclical volatility of demand, the survival probabilities of firms, wage levels, mobility costs, and so forth.

46. The customization may be evident in the product (the tailor-made suit), or it may be incorporated in other, not directly observable, attributes such as credit terms, delivery service, or implicit warranties.

47. Unions and employee associations were combined in the CPS. I shall refer to the combined group as "unions."

48. The impact of unionism on wages has been extensively studied by Lewis (1963), Mellow (1981), Mincer (1981), and in numerous studies cited by Freeman and Medoff (1981). The union wage differential has increased in the 1970s. Over the course of the last thirty years, the data reveal an upward trend in the ratio of nonproduction to production workers in manufacturing. If salaried nonproduction workers can be substituted for blue-collar workers, the rising union wage differential in the 1970s should have accelerated this trend.

49. The expected wage is equal to the wage times the proportion of the period that the individual is employed. Mincer (1981) found that the probability of a temporary layoff was

inversely related to seniority and that union members experienced higher layoff rates. Most unions establish seniority rules that place the burden of temporary layoffs on junior employees.

50. This conclusion is tentative. I have not examined the job tenure distributions by single years which might reveal the source of the difference in mean durations.

51. The age at entry for the ith worker, α_i, is the difference between his current age A_i and his years of job tenure T_i. The mean age at entry can thus be calculated from the tables as $\bar{\alpha} = \bar{A} - \bar{T}$.

Constructed Mean Age at Entry of Tenured Workers
(men with five or more years of job tenure)

Firm Size	Nonunion	Union
1–99	33.2	35.3
100–999	30.1	28.7
1,000+	26.6	26.6

Tenured workers at large firms joined their employers at a younger mean age (26.6 years) than the mean age of all new employees in 1979. However, the tenured workers at small firms were drawn from the right side of the age distribution of new employees.

52. Table 2.11 reveals that only 33.2 percent of male workers in small firms had five or more years of job tenure. Employment durations tend to be shorter for at least two reasons: First, these jobs may serve as stepping stones and training grounds for new entrants. The age distributions of employees in different industries show that young persons are more heavily represented in trade, personal services, and the low-wage manufacturing industries. This allocation can be explained by a reciprocal search process in which individuals seek high-paying jobs and firms look for suitable candidates. New entrants who lack job experience and work histories may take jobs at small firms that provide little specific training, but these jobs enable them to establish track records documenting their reliability, honesty, and capacity to work with others. Second, the turnover of small firms due to bankruptcies and takeovers are responsible for some job terminations.

53. The data of table 2.2 collected by King (1923) indicate that in the recession of 1920–22 the variability of employment was less in small firms. A perusal of data in *County Business Patterns* also suggests that employment is less volatile in small establishments. However, both data sets describe the behavior of employment for the aggregate of firms in each size category and could thus conceal considerable churning among firms within each size group. It is unclear how a prior job affects the behavior of a released employee. Individuals who worked at small firms possess only general human capital that can readily be transferred to many jobs. Workers who are separated from large companies may try to find an employer who can utilize his specific human capital. But if training is truly *specific*, it has little or no value to others. This reasoning suggests that specifically trained, unemployed workers (ignorant of the nature of their human capital) are more likely to experience the long spells of unemployment that were reported by K. Clark and Summers (1979).

54. The unionized sector in manufacturing is mainly located in large firms and in industries with high capital utilization rates. The unionized firms that have adopted rigid production techniques tend to experience greater employment variability, and the burden of unemployment is mainly placed on junior employees. The evidence examined by Mincer (1981) and the studies cited by Freeman and Medoff (1981) indicate that the incidence of temporary layoffs is higher in the unionized sector. This finding in combination with the fact that premiums for unemployment insurance are *not* based on actuarially fair experience ratings, led me to the tentative conclusion that the present unemployment insurance program redistributes income from nonunionized workers to union members. This conclusion must await further empirical study.

References

Alchian, A. A. 1959. Costs and outputs. In *The allocation of economic resources*, ed. M. Abramovitz. Stanford: Stanford University Press. (Reprinted in 1977. *Economic forces*. Indianapolis: Liberty Press.)

———. 1969. Information costs, pricing, and resource unemployment. *Journal of Economic Inquiry* 7: 109–28.

Alchian, A. A., and H. Demsetz. 1972. Production, information cost, and economic organization. *American Economic Review* 62: 777–95.

Baily, M. N. 1974. Wages and employment under uncertain demand. *Review of Economic Studies* 41: 37–50.

Becker, G. S. 1964. *Human capital: A theoretical and empirical analysis*. New York: Columbia University Press for the NBER.

Becker, G. S., and G. J. Stigler. 1974. Law enforcement, malfeasance, and compensation of enforcers. *Journal of Legal Studies* 3: 1–18.

Ben-Porath, Y. 1967. The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75: 352–65.

Clark, J. M. 1923. *Studies in the economics of overhead costs*. Chicago: University of Chicago Press.

Clark, K. B., and L. H. Summers. 1979. Labor market dynamics and unemployment: A reconsideration. *Brookings Papers on Economic Activity*, part 1, 13–60.

Coase, R. H. 1937. The nature of the firm. *Economica* 4: 386–405.

Deardorff, A. V., and F. P. Stafford, 1976. Compensation of cooperating factors. *Econometrica* 44: 671–84.

Director, S. M., and S. I. Doctors. 1976. Racial differences in blue-collar turnover rates. *Industrial Relations* 15: 338–42.

Duncan, G. J., and F. P. Stafford. 1980. Do union members receive compensating wage differentials? *American Economic Review* 70: 355–71.

Foss, Murray R. 1981. *Changes in the workweek of fixed capital*. Washington, D.C.: American Enterprise Institute.

Freeman, R. B., and J. L. Medoff. 1981. The impact of collective bargaining: Illusion or reality? NBER Working Paper no. 258. Cambridge, Mass.: National Bureau of Economic Research.

Friedman, Milton. 1976. *Price theory*. Chicago: Aldine Publishing Co.

Georgescu-Roegen, N. 1967. Chamberlin's new economics and the unit of production. In *Monopolistic competition theory: Studies in impact*, ed. R. Kuenne. New York: Wiley.

Gordon, Donald F. 1974. A neoclassical theory of Keynesian unemployment. *Journal of Economic Inquiry* 12: 431–59.

Gordon, D. F., and G. M. Walton. 1982. A theory of regenerative growth and the experience of post–World War II West Germany. In *Explorations in the new economic history*, ed. R. L. Ransom, R. Sutche, and G. M. Walton. New York: Academic Press.

Grunfeld, Yehuda. 1960. The determinants of corporate investment. In *The demand for durable goods*, ed. A. C. Harberger, 211–66. Chicago: University of Chicago Press.

Hall, R. E. 1972. Turnover in the labor force. *Brookings Papers on Economic Activity*, part 3, 709–64.

———. 1980. Employment fluctuations and wage rigidity. *Brookings Papers on Economic Activity*, part 1, 91–123.

———. 1982. The importance of lifetime jobs in the U.S. economy. *American Economic Review* 72: 716–24.

Hilton, G. W. 1957. The British truck system in the nineteenth century. *Journal of Political Economy* 65: 237–56.

Hutt, W. H. [1939] 1977. *The theory of idle resources*. Reprint. Indianapolis: Liberty Press.

Jensen, M., and W. Meckling. 1976. Theory of the firm, managerial behavior, agency costs, and the ownership structure. *Journal of Financial Economics* 4: 305–60.

Kaldor, N. 1934. The equilibrium of the firm. *Economic Journal* 44: 60–76.

King, W. I. 1923. *Employment, hours, and earnings in prosperity and depression, United States 1920–1922*. New York: National Bureau of Economic Research.

Klein, B., R. J. Crawford, and A. A. Alchian. 1978. Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics* 21: 297–326.

Koike, Kazuo. 1978. Japan's industrial relations: Characteristics and problems. *Japanese Economic Studies* 7: 42–90.

Lazear, E. P. 1979. Why is there mandatory retirement? *Journal of Political Economy* 87: 1261–84.

Lester, R. 1967. Pay differentials by size of establishment. *Industrial and Labor Relations Review* 7: 57–67.

Lewis, H. Gregg. 1963. *Unionism and relative wages in the United States*. Chicago: University of Chicago Press.

Lucas, Robert E., Jr. 1967. Adjustment costs and the theory of supply. *Journal of Political Economy* 75: 321–34.

———. 1978. On the size distribution of business firms. *Bell Journal of Economics* 9: 508–23.

Masters, S. L. 1969. An interindustry analysis of wages and plant size. *Review of Economic Statistics* 51: 341–45.

Medoff, J. M., and K. G. Abraham. 1981. The role of seniority in U.S. workplaces: A report on some new evidence. NBER Working Paper no. 618. Cambridge, Mass.: National Bureau of Economic Research.

Mellow, Wesley. 1981. Employer size, unionism, and wages. In *New approaches to labor unions, Research in labor economics*, supplement 1982, ed. R. G. Ehrenberg. Greenwich, Conn.: JAI Press.

Mincer, Jacob. 1981. Union effects: Wages, turnover, and job training. In *New approaches to labor unions*, *Research in labor economics*, supplement 1982, ed. R. G. Ehrenberg. Greenwich, Conn.: JAI Press.

Oi, W. Y. 1962. Labor as a quasi-fixed factor. *Journal of Political Economy* 70: 538–55.

———. 1967. The neoclassical foundations of progress functions. *Economic Journal* 77: 579–94.

———. 1974. On the economics of industrial safety. *Law and Contemporary Problems* 38: 669–99.

———. 1979. Academic tenure and mandatory retirement under the new law. *Science* 206: 1373–78.

Parks, Richard W. 1979. Durability, maintenance, and the price of used assets. *Economic Inquiry* 27: 197–217.

Pencavel, J. H. 1977. Work effort, on the job screening, and alternative methods of remuneration. *Research in Labor Economics* 1: 225–58.

Pissarides, C. A. 1976. *Labor market adjustment*. Cambridge: Cambridge University Press.

President's Council of Economic Advisers. 1980. *Economic report of the president*. Washington, D.C.: Government Printing Office.

Reder, M. W. 1955. Theory of occupational wage differentials. *American Economic Review* 45: 833–52.

Rees, Albert, and George P. Shultz. 1970. *Workers and wages in an urban labor market*. Chicago: University of Chicago Press.

Robinson, E. A. G. 1958. *The structure of competitive industry*. Chicago: University of Chicago Press.

Rosen, S. 1972*a*. Learning and experience in the labor market. *Journal of Human Resources* 7: 326–42.

———. 1972*b*. Learning by experience as joint production. *Quarterly Journal of Economics* 86: 366–82.

———. 1974. Hedonic prices and implicit markets. *Journal of Political Economy* 82: 34–55.

Rosenfeld, Carl. 1977. Job search of the unemployed, May 1976. *Monthly Labor Review* 100 (November): 39–43.

Shinohara, M. 1962. *Growth and cycles in the Japanese economy*. Tokyo: Kinokuniya Bookstore Co.

Smith, Freeman. 1977. Wage trends as performance displays productive potential: A model and application for academic early retirement. *Bell Journal of Economics* 8: 419–43.

Smith, M. A. M. 1974. International trade in second-hand machines. *Journal of Development Economics*, pp. 261–78.

Steiner, P. O. 1957. Peak loads and efficient pricing. *Quarterly Journal of Economics* 71: 585–610.

Stigler, George J. 1962. Information in the labor market. *Journal of Political Economy* 70: 94–105.

Turvey, Ralph. 1968. *Optimal pricing and investment in electricity supply*. Cambridge, Mass.: MIT Press.
Wenders, J. T. 1976. Peak load pricing in the electric utility industry. *Bell Journal of Economics* 7: 232–41.
Wright, T. P. 1936. Factors affecting the costs of air frames. *Journal of Aeronautical Sciences* 3: 122–28.

Comment Ernst R. Berndt

For quite some time now, Walter Oi has been working on ideas concerning fixed and variable labor inputs into production processes; the idleness, slack capacity, and utilization of these fixed inputs; and implications for cyclical variations in the employment and wage rates of labor. Professor Oi is very much aware of the complexity of the labor market. This paper represents, I think, Professor Oi's converging ideas on how the labor market really works, and how its complexities can be unraveled and understood most usefully and succinctly. It is most appropriate that we open the NBER conference with Professor Oi's paper, for it deals in an original way with classic research issues that have a long and distinguished tradition within the NBER.

Essentially, the paper consists of three distinct essays, each dealing in a different way with the single theme that it is fixed employment costs which provide the glue that binds together workers and firms. The first essay consists of a review and assessment of the literature dealing with the notion of labor as a quasi-fixed factor of production. In the second essay, Oi summarizes search theory and the implicit contract literature and then analyzes implications for wage flexibility and turnover. In brief, Oi argues that search and implicit contract theories are in fact appropriate only for one portion of the labor market, albeit an important one, namely, large firms employing specialized labor. The third essay builds on the first two and sets out novel insights and hypotheses. Specifically, in this essay Oi puts forth a somewhat different notion of "dual labor markets," based here on the heterogeneity of firms.

I begin with a brief review of the first essay. Total labor cost to a firm consists of variable wages paid to workers in return for a flow of productive services plus the periodic rent on the firm's investment outlay incurred while hiring and training its workers. Hiring costs include the direct costs of recruiting and payroll processing, plus such indirect costs as those incurred in terminating, laying off and recalling workers, and

Ernst R. Berndt is professor of applied economics at the Sloan School of Management, Massachusetts Institute of Technology, Cambridge, and a research associate with the National Bureau of Economic Research.

incremental costs of unemployment insurance resulting from higher labor turnover rates.

Following Gary Becker, Oi notes that training can either be general (when benefits could possibly be realized by several different employers) or specific (when benefits in the form of enhanced marginal productivity of the worker can be extracted by only one employer). In competitive markets, firms will provide general training only if the costs are borne by the recipient. When training is specific, however, competitive firms may willingly undertake investment costs, which will result in a wedge between current marginal revenue product and the current wage, the wedge consisting of the periodic rent earned by the firm on its investment in specific training. The value to the firm of any specific training investment can be increased by extending the expected period of employment through, for example, offering different age-income profiles and more attractive pension plans (as has been argued by Donaldson and Eaton 1976).

Professor Oi notes that, in a production cycle, the timing of demand changes for different labor types depends on the relative size of the periodic rent in total labor costs, so that employment variations—both up and down—are less volatile and less frequent for workers with higher degrees of fixity. Oi neglects to mention here the fact that to some extent alternative hypotheses can produce the same cyclical behavior. For example, more than a decade ago Griliches (1969) put forward an hypothesis concerning technology, namely, capital-skill complementarity. Recently Morrison and Berndt (1981) have shown that when physical capital is the only quasi-fixed input, the elasticity of demand for skilled labor with respect to output will be less than one if and only if skilled labor is a Hicks-Allen complement with physical capital. Given such capital-skilled labor complementarity, short-run increasing returns to aggregate labor can easily occur, even when skilled labor is a fully variable factor. Hence, somewhat different frameworks can "explain" the same procyclical phenomena. However, as Professor Oi notes, while the technological substitutability-complementarity story can explain relative shifts in factor demands, by itself it is unable to explain the exact *timing* of employment turning points, the latter seemingly requiring at least some story on costs of adjustment for physical capital and skilled labor.[1]

The above discussion points out, I believe, that the notion of quasi-fixity of certain inputs is related quite closely to the notions of Lucas (1967*a*, *b*) and Treadway (1971) concerning internal and external increasing marginal costs of adjustment. I would have preferred to have seen Professor Oi provide a more detailed and rigorous comparison of these two conceptual frameworks. I conjecture that the Oi quasi-fixed factor, static equilibrium framework is more likely to yield corner solutions, since in the Lucas-Treadway dynamic framework, adjustment costs are

increasing at the margin in a continuous way. Also, in the empirical review, I would have liked to have seen some discussion of the contributions of Brechling (1975) and Nadiri and Rosen (1973) in which the cyclical behavior of employment is broken down into number of employees and average hours at work per employee.

In his second essay, Professor Oi begins by noting that quasi-fixity, a demand notion, cannot by itself explain the sluggish response of wages to changes in aggregate demand or the persistence of involuntary unemployment. For some time now, a number of economists have viewed a great deal of unemployment as frictional, voluntary, and in some sense "optimal," since in this view persons are envisaged as investing time and resources while unemployed into searching for, finding, and securing better paying and more satisfying jobs. As noted by Professor Oi, such search models tend to imply large flows into and out of unemployment, with only a thin tail of individuals experiencing long unemployment spells. Recent empirical studies cited by Oi cast considerable doubt on the quantitative significance of such search behavior. For example, according to Clark and Summers, turnover and search accounted for only about one-fourth of 1 percent of unemployment, and in fact 64 percent of job changes were made with no intervening spell of unemployment at all. Moreover, Clark and Summers observed that most workers take the first job offer received, with this job lasting typically less than two years. In such a world, the assumption that search can be conducted more efficiently when unemployed must be rejected—it is irrational for a person to remain unemployed in order to allocate time to job search. Oi concludes, therefore, that such search theories which rationalize voluntary unemployment are not very useful empirically.

While search theories examine the behavior of an individual seeking a permanent job, contract theories based on different attitudes toward risk by firms and workers attempt to explain the behavior of firms in designing compensation packages and employment policies that attract and retain "permanent" workers. If the benefits of risk sharing and the costs of mobility are important to workers, and if fixed employment costs comprise a substantial portion of a firm's total labor cost, then one would expect to observe job durations with long mean tenure. Arguing both analytically and with the benefit of empirical research, Oi contends that such behavior is to be found only in sectors of the economy consisting of large firms. For example, Oi cites empirical research results recently reported by Hall indicating that a representative worker could be expected to hold ten jobs over a lifetime, and that by age thirty, 40 percent of workers will be at a job they will hold on average for twenty years. Hall also reports that the time profile of employment turnover is most important: job tenure turbulence is high during the first five to ten years in the labor force, when young people experience high turnover rates in search

of "permanent" lifetime jobs. As their ages increase, more workers settle into permanent jobs lasting for twenty or more years.

Oi concludes this essay by noting that firms and individuals are heterogeneous and therefore exhibit different patterns of fixed employment costs and job tenure. Specifically, a marginal firm whose survival probabilities are slim is unlikely to make substantial investments in recruiting and specific training; hence, argues Oi, small firms should be expected to be less interested in contracting. Moreover, an individual with few assets and only general training will tend to be less interested in incurring the high fixed costs of job search; hence younger and generally trained workers should be expected to be less interested in searching. By contrast, large stable firms are more interested in contracting, and experienced workers with some specific training have more incentive to search.

This then brings us to the third essay in which Professor Oi displays his considerable skill as an insightful and strategic craftsman. The notion that both employers and employees are heterogeneous has been around for quite some time and has been formalized in a number of stochastic search and contract models. Oi, however, strategically simplifies by suggesting a particular structure to the forms of the underlying distributions. Let there be two groups of firms—call them giants and small firms (the firm is viewed as a team in the sense of Alchian and Demsetz—more on this later). Let there also be two types of workers—those with general human capital training and those with specific training. A firm that "produces" quasi-fixed labor inputs can be thought of as a multiproduct firm whose outputs are the regular products plus laborers embodying additional specific training.[2] More specific capital will be invested in those individuals who have longer expected employment tenures and who can manifest larger productivity increases through training. Oi then suggests that specific training would be concentrated on more highly skilled and malleable/educable workers; this implies a positive relation between the current wage rate and the degree of fixity.

Turning then to the heterogeneity of firms, Oi identifies and highlights several systematic differences in the way heterogeneous firms organize their production. First, firms differ in their ability to exploit internal gains of specialization. The central agent performing this task for the team is the entrepreneur (coach) who both supplies managerial input and monitors worker performance. Hence the full labor costs of a worker to the firm include at least the wage rate and the cost of monitoring performance. Differences in workers and entrepreneurs are described in terms of two parameters: μ, the implicit monitoring costs each worker generates by being employed, and λ, the ability of entrepreneurs to transform hours into effective managerial input. High-λ entrepreneurs are found in large firms, for they supply more effective managerial input, thereby increasing the scale of output produced by capital and labor inputs. However, since

these high-λ entrepreneurs have a comparative disadvantage in monitoring performance, they attempt to compensate by adopting capital-intensive production methods and by hiring more productive workers who command higher wages. But how do giant firms with assembly-line and batch production processes reduce monitoring costs of workers? Here Oi evokes the old image of IBM employees. New employees at giant firms receive more firm-specific training which "adapts" them to a particular team production process, encourages compliance with the prescribed job description, and teaches them preferred ways of performing tasks. When all of the team's members are more or less homogeneous and work in the same way, monitoring costs can be sharply reduced. In a sense, then, large firms substitute additional fixed employment costs now for lower monitoring costs later on. For small firms producing more customized outputs, the fixed costs of assembling and training a homogeneous team may not be warranted. Hence, in Oi's view, since it is the entrepreneurial input that is most scarce, differing firms organize production teams and supply those kinds of products in various ways, each so as to yield the largest return to the scarce entrepreneurial input.

Turning to somewhat related issues, Oi notes that capital typically requires maintenance which in turn, he suggests, necessitates use of specific-trained workers.[3] Maintenance, however, is a labor-intensive activity requiring extensive monitoring. Firms facing high wages and high monitoring costs will tend to substitute new machines for lower maintenance. Such substitution between maintenance and investment can occur across time in response to cyclical fluctuations in marginal productivities—more maintenance in recession when the marginal revenue product of production work is lower, and correspondingly less maintenance and more production in the upturn. Assignments of specific workers to maintenance and training from the production of goods is therefore countercyclical and can help explain Okun's observed, short-run, increasing return to labor. I might add here that I have not seen much evidence yet that in the current recession workers are devoting more time to rebuilding human and nonhuman capital, nor that this rebuilding is greater, as Oi would suggest, in large, capital-intensive firms. One industry worth examining in this regard is the electric utility industry, for its variations over time in excess capacity are well known, good data are available, and it does not contain the additional problem of using output inventories as a buffer stock, for electricity is not easily stored.

Turning now to the compensation of fixed factors, Oi notes that perhaps more generous but less portable fringe benefits reduce labor turnover and thereby increase the capitalized value of specific training. Moreover, deferred payment in the form of pensions discourages malfeasance and shirking of tasks by employees, thereby reducing monitoring costs. Oi points to evidence that deferred compensation is becoming

increasingly important, and from this he concludes that in the United States labor is becoming more firm specific. In my judgment, such an important conclusion is not yet warranted on the basis of the evidence Oi cites. Some of the recent increasing importance of deferred compensation and pensions may be due simply to the changing age distribution of the labor force, and some may also be due to the U.S. experience of wage and price controls in the 1970s, compliance with which created incentives for firms to increase the fringe and deferred payment items of the compensation package. Moreover, if Oi's hypothesis were true, labor productivity should be more procyclical today than before. I recall recent *Economic Reports of the President* in which it was mourned that in the last decade, during upswings, labor productivity had risen much less than previously. Also, is labor productivity more procyclical in the capital-intensive manufacturing sector today than in the service sector? I know of no careful study on this issue.

In the closing pages of his paper, Professor Oi examines empirical relationships among firm size, wage rates, education, race and sex, age and job tenure, capacity utilization, shift work, and unionization. The remarkable and, I think, most significant feature of Oi's paper is that fixed employment costs and the particular highly structured heterogeneity he envisages among firms and individuals has clear implications for the signs of correlations among these variables. My only criticism is that, by and large, the way in which Oi examines these relationships empirically is just two at a time, using bivariate regressions or simple correlations. The rich set of testable hypotheses generated by Professor Oi deserves a much more careful and detailed examination within a multivariate regression and partial correlation framework.

This paper suggests numerous directions for future research, in addition to those noted by Oi and suggested by me earlier. Specifically, I would hope that, in the future, attempts be made to obtain direct estimates of fixed employment costs over time and space, that dynamic optimization be incorporated more explicitly, that implications of specific training for market structure be examined more carefully in a multi-industry framework, and that the empirical notion of the firm be considered in greater detail. Regarding this last point, I am uncertain what best corresponds in the real world to the firm or team envisaged by Oi, particularly given numerous recent mergers, growth of conglomerates, and heterogeneity of firms across countries such as the United States and Japan.

These further research issues offer each of us great opportunities. If only there were more applied theorists, like Walter Oi, whose insightful analyses generate such well-structured opportunities for important additional empirical research.

Notes

1. See, however, section 4 in Morrison and Berndt (1981) on the issue of whether costs of adjustment are either necessary or sufficient for the existence of short-run increasing returns to aggregate labor.

2. This proposition was developed in the internal costs of adjustment literature about a decade ago by Brechling and Mortenson (1971, p. 5) who stated that:

> The assumption that internal costs of adjusting input levels exist is equivalent to the proposition that the inputs used by the firm at one point in time are at least partially "produced" by the firm at some earlier date. For example, the existence of hiring and training costs imply that the raw material, a newly-employed worker, must be processed and modified in certain ways by the firm before his services are appropriate for use in the production process. . . . In other words, the production rates and the time rates of change in input levels are measures of jointly produced output and inputs, respectively. Hence, more rapid changes in input levels can be obtained only either at the expense of output, if the firm's resources are given, or by increasing resource levels, if output is maintained at some predetermined level.

3. Such a relationship could generate the capital-skill complementarity observed by Griliches (1969).

References

Brechling, Frank P. R. 1975. *Investment and employment decisions*. Manchester: Manchester University Press.

Brechling, Frank P. R., and Dale T. Mortenson. 1971. Interrelated investment and employment decisions. Paper presented at the winter meetings of the Econometric Society in New Orleans.

Donaldson, David, and B. Curtis Eaton. 1976. Firm-specific human capital: A shared investment or optimal entrapment? *Canadian Journal of Economics* 9: 462–72.

Griliches, Zvi. 1969. Capital-skill complementarity. *Review of Economics and Statistics* 51: 465–68.

Lucas, Robert E., Jr. 1967*a*. Optimal investment policy and the flexible accelerator. *International Economic Reveiw* 8: 78–85.

———. 1967*b*. Adjustment costs and the theory of supply. *Journal of Political Economy* 75: 331–44.

Morrison, Catherine J., and Ernst R. Berndt. 1981. Short-run labor productivity in a dynamic model. *Journal of Econometrics* 16: 339–65.

Nadiri, M. Ishaq, and Sherwin Rosen. 1973. *A disequilibrium model of demand for factors of production*. NBER General Series no. 99. New York: Columbia University Press for National Bureau of Economic Research.

Treadway, Arthur B. 1971. On the multivariate flexible accelerator. *Econometrica* 39: 845–55.