NBER WORKING PAPER SERIES

THEORIES OF WAGE RIGIDITY

Joseph E. Stiglitz

Working Paper No. 1442

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1984

The research reported here is part of the NBER's research program in Economic Fluctuations and project in Macroeconomics. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #1442
September 1984

Theories of Wage Rigidity

ABSTRACT

This paper considers two sets of theories attempting to explain wage rigidities and unemployment: implicit contract theory and the efficiency wage theory. The basic thesis of the paper is that the former set of theories do not provide a convincing explanation of the kind of wage rigidity which is associated with cyclical unemployment, while the latter theories do. Several of the more recent versions of implicit contract theory are considered: implicit contracts with asymmetric information may give rise to over employment rather than underemployment, and the forms of contracts to be expected, were asymmetric information considerations paramount, are not observed. Other versions of the asymmetric information implicit contract model, explicitly long term in nature, may give rise to full employment. One version of implicit contract theory which does give rise to lay-offs arises when search is costly and cannot be monitored. But even this extension does not explain certain important features of observed patterns of unemployment.

In contrast, the efficiency wage models not only provide an explanation of the existence of unemployment equilibrium in competitive economies, but they also provide part of the explanation of the observed patterns of unemployment. They also explain why different firms may pay similar workers different wages, why wages may be sticky, why firms may not loose much if they fail to adjust their wages, and why, when they adjust their wages optimally, they adjust them slowly.

The policy implications of the efficiency wage model are markedly different from those of models in which wages are absolutely rigid as well as from those in which unemployment arises from asymmetric information.

Joseph E. Stiglitz
Department of Economics
Dickinson Hall
Princeton University
Princeton, NJ 08544

THEORIES OF WAGE RIGIDITY

by

Joseph E. Stiglitz*

It is widely recognized that the assumption that wages are rigid is central to Keynes' explanation of the persistence of unemployment.[1] Indeed, in the fixed price (temporary equilibrium) models, which are currently so much in fashion in Europe, it is the rigidity of wages and prices which provides the sole explanation for unemployment in the economy.[2]

These theories have, however, simply assumed that wages and prices are rigid; they have not attempted to explain the rigidity. This has left

---

*Paper prepared for presentation at a conference on Keynes' Economic Legacy University of Delaware, January 12-13, 1984. Financial support from the National Science Foundation is gratefully acknowledged. This paper was completed while I was a Visiting Fellow at the Hoover Institution, Stanford. I am indebted to R. Arnott, C. Azariadis, G. Akerlof, J. Butkiewicz, Russell Cooper, O. Hart, S. Grossman, B. Greenwald, A. Hosios, John Pencavel, A. Weiss, J. Yellen, and D. Newbery for helpful discussions.

---

[1] We shall follow the well honored tradition in Keynesian' analysis of obfuscating whether it is the rigidity in real or nominal wages which is crucial for the analysis. Most of the theories that we discuss in this paper are concerned with rigidities in real wages. In section II we discuss alternative explanations of the rigidity of nominal wages.

[2] Among recent expositions include two books by Malinvaud (1977, 1980), and one by Grandmont (1982). Earlier studies included those by Solow and Stiglitz (1967) and Barro and Grossman (1971). Although most studies in this tradition place primary emphasis on the rigidity of wages and prices, other rigities may give rise to unemployment. For instance, Neary and Stiglitz (1983) analyze the consequences of rigidities in the rate of interest.

these models open to charges of adhocery and inconsistency: why should firms which are assumed to act in such a rational, profit maximizing, competitive manner with respect to production and employment decisions, not act in a similarly rational, profit maximizing, competitive manner when it comes to decisions concerning wage and price determination? Worse still, the models may not provide an adequate basis for policy analysis: even if wage and prices are fixed in the very short run, might they not be affected by policies, such as the level of unemployment compensation?

In this paper, I wish to discuss two sets of theories which attempt to explain the observed rigidities in wages.[3] The first, which I shall refer to generically as "implicit contract theories," explain wage rigidity as a consequence of implicit insurance provided to risk averse employees by their employers. The second I shall refer to generically as "efficiency wage" theories; these hold that workers' productivity depends on the wage that they are paid; firms may not lower wages, even in the presence of unemployment, because to do so would lower the productivity of their labor force. The basic thesis of this paper is that the former set of theories do not provide a convincing explanation of the kind of wage rigidity which is associated with cyclical unemployment, while the latter theories do.[4]

Our discussion is not intended to be a survey of what have developed into large literatures, but rather to provide a critical assessment of

---

[3]Elsewhere, I have addressed the question of price rigidity. (Stiglitz (1984)).

[4]As I note below, the two theories are not mutually inconsistent: a full theory will need to incorporate elements of both approaches.

the central issues associated with each of the theories.

1. What it is that is to be Explained.

A theory of unemployment must explain not only the level of unemployment, but its form and composition. For instance, workers may be put on short weeks as well as laid off. If individuals have quasi-concave indifference curves, they prefer work sharing to having some probability of working a full week, and some probability of not working at all.

There are two easy explanations of the prevalence of lay offs in the US (in Europe, work sharing seems to be more common). The first has to do with unemployment compensation: an individual who is on half time cannot collect the public subsidy associated with unemployment compensation. There are two basic objections to this "explanation": First, given the limitation on the number of weeks which individuals can collect unemployment compensation (and the fact that there is only a one week waiting period for collecting it) firms should rotate individuals who are laid off. The argument for job rotation is even more compelling if the marginal utility of leisure diminishes significantly, the longer the individual is unemployed.[5] Moreover, the phenomenon of unemployment pre-dated the extensive provision of unemployment insurance factors.

Another explanation for lay-offs is that technologies require individuals to work 8 hour days, can similarly be dismissed: in most

---

[5]There is, moreover, considerable evidence that being without work for an extended period of time is extremely disruptive to the individual and his family; the loss of income is only partially responsible; the loss of sense of worth from not being gainfully employed, the loss of respect from others, and the fear of this loss, are probably even more important. Beyond some point, leisure appears to have negative value.

production jobs, the economies of time-scale seldom extend beyond the day, and certainly not beyond the week or month.

It is not, however, enough to explain why firms might lay off individuals: one must explain why other firms do not rehire them. Unemployment arises when layoffs exceed hires. Most of the implicit contract theories have called themselves theories of unemployment simply by assuming that there is no labor mobility; in that case, obviously, any worker who is laid off is unemployed. But even in the midst of the Great Depression, there was a high job accession rate -- not high enough to compensate for the even higher lay off rate. Unemployment cannot be explained on the basis of the behavior of a single firm, but only by the analysis of the market as a whole. Any implicit contract theory which purports to explain unemployment must, thus, be a part of a theory of the market.

Finally, any theory should be able to explain not only why certain shocks give rise to unemployment (an increase in lay offs relative to hires), but also why the unemployment should be concentrated within certain parts of the labor force.

Furthermore, any theory which we construct should be consistent with certain other characteristics of economic fluctuations: among these are (a) those who are unemployed often look very much like those who are employed; individuals with similar qualifications seem to be treated very differently; and (b) those who are unemployed are, for the most part, unhappier than those who remain employed. For the most part, individuals prefer not to be layed off, rather than to be layed off.[6]

---

[6]As we note below, there are a few instances where this is not true.

Of course, even were we able to explain why there are lay-offs, and why those layed off are not rehired, we would not have a complete explanation of economic fluctuations. We would still need to explain what are the sources of the disturbances to the economy. Our task here is the more limited one of explaining why certain shocks might lead to unemployment, why the adjustment mechanisms, which in traditional competitive analysis lead the economy to full employment[7] seems to fail so frequently.

I

Implicit Contract Theories of Wage Rigidity

1. The Basic Theory[8]

The basic hypothesis of implicit contract theory is that workers are risk averse, and have limited access to capital markets. They would like to obtain insurance against fluctuations in their income. They cannot obtain such insurance from conventional insurance companies;[9]

---

[7]Variations in economic circumstances may lead to variations in hours worked, or to some individuals deciding not to seek work, even in a Walrasian economy. We wish to explain why it appears to be the case that variations in employment are greater than they would in a Walrasian equilibrium.

[8]The standard references for the Implicit Contract Theory are Azariadis (1975), Gordon (1974), and Bailey (1974). For two recent surveys, see C. Azariadis (1979) and C. Azariadis and Stiglitz (1983).

[9]Most of the literature has not addressed the question of why individuals cannot or do not purchase wage and/or employment insurance companies. Arnott, Hosios, and Stiglitz (1983) provide a brief discussion; noting the central problems of moral hazard (to be discussed below), they observe that the employer has an informational advantage over other potential suppliers of insurance.

Similarly the standard implicit contract models do not provide a

(Footnote continued)

but their employers are less risk averse and have greater access to the capital market than they do. As a result, employers provide some form of wage and employment insurance as part of the employment package. To put it another way, a firm which offered such insurance as part of its employment package would be able to attract workers at a lower (average) wage than a firm which did not provide such insurance.

1.1 Simple Implicit Contract Theory May Explain Wage Rigidities But Not Unemployment

If all states of nature are observable (and verifiable), then the implicit contract would specify the amount of labor and the wage to be paid in each state. In such circumstances, though there may be relatively little fluctuations in wage (incomes), the implicit contract would not give rise to unemployment: the marginal rate of substitution of each individual between income and leisure would be equal to the marginal rate of transformation, and there would be no lay-offs.[10] Recall our original objective was to find an explanation of wage rigidity which could help explain unemployment; the implicit contract theory may provide us some insights into the movement of wages, but (at

---

[9](continued)
good explanation of why access to the credit market is limited, or why it is more limited for individuals than for firms. One explanation of credit rationing in competitive markets with imperfect information has recently been provided by Stiglitz and Weiss (1981, 1983). Because the firm is more informed about its employees' future income potential than, say, a bank, it has a natural advantage in serving as an intermediary in the provision of (possibly implicit) credit.

[10]Except if some individual would have remained voluntarily unemployed in the corresponding Walrasian equilibrium, because his marginal role of substitution exceeded his marginal rate of transformation at zero work.

least in the simple version) does not explain movements in employment.[11]

This can be seen most easily if we think of the firm as having two departments, an insurance department and a production department. The production department pays workers the ordinary "Walrasian" (spot) wages, which may fluctuate considerably. The worker then takes this wage to the insurance department; the insurance department agrees to supplement his income in bad states, in return for a premium (payable only in good states.) This results in the individual's take home pay (his wage plus or minus payments to the insurance company) varying much less than his Walrasian wage. What we observe, of course, is the take home pay; but what enters into the analysis of the production decision is the Walrasian (shadow or spot) wage; and since the Walrasian wage is perfectly flexible, there is no unemployment.

(If the marginal rate of substitution was less than the marginal rate of transformation in some states of nature, so that the amount individuals required to be compensated for an increase in work by an hour was less than the value of the extra amount of output they produced, then it would pay the firm to ask the worker to work the additional hour, and increase his pay accordingly; if the individual is risk averse, and the utility function is separable in consumption and leisure, he will increase his take-home pay equally in all states;[12]

---

[11] Even its implications for the movements in real wages may not be totally convincing. If workers' utility functions are separable in consumption and leisure (a case we shall discuss at greater length below), and if workers' have no access to the capital market (so consumption and income at each date are identical), then consumption of all those retained throughout the business cycle should be constant, i.e. real wages should move inversely to the number of hours worked. If leisure and consumption are complements, real wages should increase even more in a recession.

more generally, the increase in pay will be spread among states to equate his marginal utility of income.)[13]

The result that implicit contracts do not give rise to unemployment fluctuations[14] should not come as a surprise: it has long been recognized that insurance contracts (and recall, implicit contracts are really nothing more than insurance contracts provided by the employer) improve the functioning of competitive economies.

Indeed, one of the basic results of Arrow and Debreu was to show that a competitive economy would be pareto efficient if there were a complete set of risk markets. Subsequently, Borch showed that the absence of insurance markets may make all individuals worse off.[15]

1.2 Market Equilibrium Is Not Constrained Pareto Efficient

More relevant for our purpose is the question of the consequences of providing employment related insurance in economies in which there are not a complete set of insurance contracts. As we have aruged, under the conditions given above, such implicit contracts will be associated

---

[12] That is, the optimal contracts entails equal pay in all states; without separability the level of pay depends on the level of work in each state.

[13] Assuming that the firm is risk neutral.

[14] This view of the implicit contract, and the corresponding implication that implicit contract theory did not give rise to unemployment, was set forth in Stiglitz (1978). See also Azariadis and Stiglitz (1983) and Akerlof and Miyazaki (1980).

[15] For general results on the constrained inefficiency of competitive economies with an incomplete set of markets, see Greenwald and Stiglitz (1983), Hart (1975), Newbery and Stiglitz (1981,1982) and Stiglitz (1973,1982); for a proof that competitive economies with implicit wage contracts are inefficient, see Newbery and Stiglitz (1983).

with full employment (the marginal rate of substitution of each individual is equal to the marginal rate of transformation); they are "locally" efficient, in the sense that, given the distribution of prices, employer's expected profits (or expected utility of profits) is maximized, subject to the constraint that workers obtain a particular level of expected utility. But the economy is not a constrained pareto optimum; that is, given the limited set of insurance markets, there are feasible interventions on the part of government (say taxes and subsidies, on employment or output) which will make some individuals better off without making others worse off. (Newbery and Stiglitz (1983)).

Our interest here, however, is more in the descriptive implication of implicit contracts than its welfare consequences. But again, we would have thought that, particularly if individuals have limited access to the capital market, and their consumption is constrained by this, that the provision of employment related insurance would have smoothed out variations in consumption, and hence would serve to dampen, rather than to exacerbate, the cyclical fluctuations of the ecomony.

There have been several attempts to modify and extend Implicit Contract Theory, to derive conditions under which such contracts might give rise to unemployment.

1.3 Unemployment Compensation and Voluntary Unemployment

The first, and most obvious, is a consequence of our system of unemployment compensation: given that this is provided not on an actuarially sound basis, it pays firms to devise employment strategies to take advantage of the public subsidies. In the limiting case where

the payments of the firm are not related at all to the experience of its employees, the firm would lay off workers all states of the world in which worker's productivity is sufficiently low. The precise condition depends on whether firms supplement the unemployment compensation of individuals. Then the layed-off workers are, in a sense, all voluntarily unemployed provided that leisure is normal: the level of utility that they get from leisure together with the unemployment compensation exceeds the level of utility they would have obtained if they worked. (See Appendix A). Most (or at least many) of those who are classified as unemployed do not fall within this category; and the problems of unemployment with which we are concerned here antedated the spread of unemployment compensation and welfare systems with more than subsistence level of payments.

Further evidence that there is more at issue is provided by the fact that the pattern of unemployment does not conform to that which would arise if lay-offs were simply a result of firms attempting to take advantage of unemployment compensation. In particular, one would expect to see more extensive use of job rotation, ensuring that no individual exhausts his unemployment benefits.

1.4 The Theory of Insurance and Implicit Contracts

A more fundamental approach is to ask, what are the generic problems associated with insurance contracts, and implicit insurance contracts in particular, and do any of these problems manifest themselves as what might look like involuntary unemployment.

There are at least six problems which are relevant for our analysis:

(a) The observability problem: the insurance contract can only cover events which are observable to the insurance company.

(b) The verifiability problem: even if the event can be observed, to enforce a contract through legal means requires that a third, outside party must be able to verify its existence.[16]

This is an example of a wider class of problems which we refer to generically as:

(c) The enforcement problem. Whenever the "trade" between two individuals does take place simultaneously, there must be some method of ensuring that the party which has agreed to make the later delivery lives up to his promise. This problem may be particularly acute, as we shall see, with implicit contracts.

(d) The complexity problem. Any insurance contract usually covers not a single event (state of nature) but a range of events; writing contracts so that each contigency which should be distinguished from other contingencies is so distinguished is a difficult if not impossible task. Usually, several different contingencies are covered by the same set of clauses, though in the absence of "transactions costs" each would be treated differently.

(e) The moral hazard problem. The provision of insurance often affects the likelihood that the insured against event occurs. To avoid the resulting inefficiencies, (individuals failing to

---

[16] A closely related problem is referred to as the Valuation problem: even when it is possible to ascertain that some loss has occured, it may be difficult to assess its magnitude.

take due care against the occurrence of the accident) insurance firms frequently insist on co-insurance clauses.

(f) The adverse selection problem. When a group of individuals are offered a given insurance policy, those among the group with the highest (subjective) probability of the occurrence of the accident are most likely to purchase the policy. The insurance firm may attempt to design contracts which separate low risk individuals from high risk individuals, e.g. by offering policies with high deductibilities, limited coverage, etc.

Each of these problems has its manifestation in the provision of employment related insurance (implicit contracts.) Surprisingly, it is only the first of these problems which has received much attention in the literature; and it does not, in my view, provide a convincing explanation of unemployment.

3. The Observability Problem.[17]

The contracts we described above had hours and wages (income) depending on the state of nature, which we assumed was observable to both the firm and the workers. What happens if the state of nature is observable to the firm, but not to the worker?

The firm will not, in general, have the incentive to tell the

---

[17] Implicit contracts with asymmetric and incomplete information have been analyzed by (among others) Arnott, Hosios, and Stiglitz (1980), Azariadis (1983), Chari (1983), Cooper (1981), Green and Kahn (1983), Grossman and Hart (1981, 1983), Grossman, Hart, and Maskin (1983), and Hart (1983). For a partial survey, see Azariadis and Stiglitz, (1983).

truth. Consider the case discussed earlier, where individuals have utility functions which are separable in consumption and leisure. The optimal contract entailed equating consumption in all states, but varying hours, with the individual working more hours in the bad state than in the good. Clearly, the firm has an incentive to announce that it is the good state, regardless of what the true state is: what it pays the workers is the same, but the amount of work it extracts from them is greater in the good state.

It is easy, however, to design a contract which induces the firm to tell the truth. We restrict the choices of the firm: if it announces that it is a good state, we compel it to make the workers work more than they would in the Walrasian perfect information equilibrium and to pay their workers more. The extra output produced by a firm in the good state is much larger than that produced by one in the bad; for firms in the good state, the extra output is enough to compensate for the increased wages which they pay; for firms in the bad state, the extra output is not enough to compensate for the increased wages they would have to pay if they falsely announced it was a good state. Such a contrat introduces two inefficiencies relative to the first best contract (with perfect information.) Workers have to bear risks, and in the good state there is over-employment.

Assume, in contrast, that firms are very risk averse, and individuals are risk neutral. Then the optimal contract would entail workers providing insurance to the firm; in the bad state, they would accept a wage below their marginal product, while in the good state, they would be paid more than their marginal product. In that case, firms might have an incentive to announce that it is a bad state, to get

workers to accept a low wage, when in fact it is a good state. Again, firms can be induced to tell the truth by restricting their choices. Now, we curtail the amount of employment which they can undertake in the bad state, relative to the Walrasian equilibrium. The "cost" of this reduction is greater in the good state than in the bad; indeed, if we curtail the maximum level of employment which we allow enough, the loss in profits in the good state exceeds the gains that it obtains from having lower wages, and thus it will pay the firm to tell the truth.

Thus, the "theory" may explain either under or over employment. Which occurs depends on the specific assumptions concerning the utility functions of workers and risk aversion of firms. The assumptions under which over employment occurs seem more plausible to me than those under which under employment occurs; but we do not have to resolve the matter here: the theory can be dismissed on other grounds, as we shall see later.

First, however, we turn to a more detailed analysis of the conditions under which over or underemployment results. Which occurs depends whether, in the first best equilibrium with complete observability, the firm would, in the good state, prefer the hours and wage associated with the low state, or vice versa. To ascertain which occurs, we need to examine in greater detail the structure of implicit contracts.

3.1 The Structure of Implicit Contracts.

To examine the structure of implicit contracts, we assume that there are only two states, each occurring with a probability of 1/2; we shall denote the good and bad states by superscripts and subscript 1 and

2 respectively. We represent the individuals utility function by U = U(y,h) where y is income and h is hours worked. The worker's productivity in state i is denoted by θ_i so firm profits per worker, i, are:

(1) $\rho^i(h_i,y_i) = \theta_i h_i - y_i$.

We begin our analysis by assuming firms are risk neutral, and individuals have a reservation expected utility level of U.

Formally, we can characterize the implicit contract with complete observability as the solution to the following problem:

(2) max $\rho^1 + \rho^2$

(2a) subject to $U^1 + U^2 = 2U$

the solution to which satisfies the first order conditions:

(3) $\theta_1 U_y{}^1 + U^1{}_h = \theta_2 U^2{}_y + U^2{}_h = 0$

(4) $U^1{}_y = U^2{}_y$

The first two conditions simply say that the marginal rate of substitution equal the marginal rate of transformation (the full employment condition while the third equation is the provision of complete insurance (in the sense that the marginal utility of income of the worker is equated in the two st tes.) This follows from the risk neutrality of the producer; it would not be true, as we shall see, if the producer is risk averse.

We have represented the equilibrium in figure 2. ρ^1 represents a constant profit line in state 1, (i.e. θh-y=constant) and ρ^2 in state 2. As we have drawn it, the firm makes a positive profit in state 1, but a loss in state 2. For any given level of profits, the hours-income package is efficient, i.e. $E_1 = \{h_1,y_1\}$ is the point where the worker's indifference curve between hours worked and is tangent to the iso-profit

line, and similarly for E_2. (Equation 3 is the mathematical representation of this tangency condition). (We have not yet explained how the levels of profits or losses in each state are determined. These are given by the solutions to (2a), (3) and (4). The diagram is useful only in portraying certain conditions which the implicit contract must satisfy.)

There are three possible patterns which can emerge. Figure 1 illustrates the case where profits in state 1 are higher with E_1 than E_2, i.e.

(5) $\rho^1(h_1, y_1) > \rho^1(h_2, y_2)$

and profits in state 2 are higher with E_2 than E_1, i.e.

(6) $\rho^2(h_2, y_2) > \rho^2(h_1, y_1)$

(5) and (6) are referred to as the <u>self-selection constraints,</u> and in figure 2a, both are satisfied.

On the other hand, in figure 2b, in the good state the firm would prefer E_2 to E_1, while in figure 2c, in the bad state the firm would prefer E_1 to E_2. In these cases, if the state of nature is not observable, firms cannot be relied upon to tell the truth. The contract must be modified to recognize this.

3.2 <u>The Implicit Contract with Unobservability.</u>

The implicit contract with unobservability is the solution to the following problem

(7) max $\rho^1 - \rho^2$

(8) subject to $U^1 + U^2 = U$, utility constraint

and

subject to

(9) $\rho^1(h_1, y_1) > \rho^1(h_2, y_2)$ self-selection constraints

(10) $\rho^2(h_2, y_2)$ $\rho^2(h_1, y_1)$.

This problem is identical to the previous one, except that we have appended the self-selection constraints. Under fairly standard conditions, one, and only one, of these constraints are binding. Which depends on whether, in the absence of the constraint, firms would have lied claiming it was a bad state when it was good, or a good state when it is bad. Before turning to this, let us note diagramatically what the self-selection constraints imply. When in the absence of the self section constraint, the firm would have said it was bad when it was good, then the firm can be induced to tell the truth by limiting the amount of labor that he can hire in the bad state. Thus, in figure 3a, the contract $\{E_1, E_2'\}$ satsifies the self-selection constraint, but entails underemployment in the bad state. (In the good state, the firm is just indifferent between E_1 and E_2'). Similarly, in figure 3b, we show the equilibrium when the firm would have announced that it was a good state when it was bad. As we argued earlier, truth telling is induced by requiring the firm to hire additional hours; the contract $\{E_1', E_2\}$ induces truthtelling. Note that this contract entails overemployment: the marginal rate of transformation is less than the marginal rate of substitution.

3.3 Over Versus Underemployment.

We now ask, which do we expect to occur, over or underemployment? To help fix our ideas, consider first the case where utility functions are separable in hours and income:

(11) $U = u(y) - v(h)$

Thus, equality of marginal utility implies that $y_1 = y_2$; this, together with (3) implies that $h_1 > h_2$ and firms always want to declare that it is a good state (to elucidate more effort). The equilibrium is always characterized by overemployment. (See figure 3).

The result holds, a fortiori, if $U_{hy} < 0$: since increasing the number of hours worked decreases the marginal utility of income, individuals in the good state receive less than in the bad state. Thus, a necessary condition for implicit contracts with asymmetric information to give rise to unemployment equilibria is that leisure and consumptions good are strong Edgeworth substitutes.

In Appendix B, we establish a somewhat stronger and more general result:[18] <u>if firms are risk neutral and the two states of nature are near each other, then there is under or over-employment as leisure is an inferior or normal good. A sufficient condition for there to be over or full employment is that leisure is a normal good. (Thus, a necessary condition for under-employment is that leisure is inferior.)</u>

On the other hand, the over-employment equilibria have the unattractive feature that individuals are always better off in the bad state than in the good state. (See figure 1c).[19]

---

[18]These results extend results independently derived by Cooper (1983).

[19]This result does not depend upon any property of the individual's utility function, other than quasi-concavity. Though there are some instances in which workers complain about being forced to work more than they would like, the result that workers are better off in the bad state, suggests to me the inappropriateness of the model for explaining over-employment. One should not conclude that therefore, the relevant equilibria are those entailing underemployment, but rather than the model itself is at best of only limited relevance.

3.4 Risk Averse Firms

We argued in the preceding section that, with risk neutral firms, there is a strong presumption that implicit contracts either entail full employment or overemployment, but not underemployment. One attempt to obtain underemployment from the implicit contract theory entails assuming risk averse firms.

Consider the limiting case where firms are infinitely risk averse, and individuals have a finite degree of risk aversion. Then, in equilibrium, their profits in the two states must be the same, as depicted in figure 4. It is immediately obvious that with infinitely risk averse firms, equilibrium will always be characterized by unemployment.

In appendix B, we show that there is still a presumption that there will be over-employment rather than underemployment. We show that a necessary condition for over-employment to result is that the firm's degree of risk aversion be large; in the case of separable utility functions, the critical value increases with the worker's risk aversion and decreases with the worker's elasticity of marginal disutility of labor (see figure 5).

Curiously enough, while implicit contract theory began as a explanation of wage rigidity in which risk neutral firms provide insurance for risk averse workers, in this new version, underemployment only results when firms are very risk averse.

Though the hypothesis of risk averse firms seems inconsistent with a well-working capital market, in the New Theory of the Firm, in which shareholders are imperfectly informed concerning the actions of the mangers, firms do behave in a risk averse manner. The question is do

they behave in a sufficiently risk averse manner.

3.5 Some Special Examples

In this section, we present some simple utility functions, for which the first best equilibrium can be easily calculated; for these utility functions, it can easily be ascertained whether there will be under or over employment.

(a) Separable utility functions.

We have already noted that if the utility function is of the form

$$U = u(y) - v(h)$$

then, if firms are risk neutral, there will always be over-employment. If firms are risk averse, the condition for underemployment (if the two states are near each other) is that

$$A \alpha \nu > (1-\alpha) R$$

where

$A = A(\theta h - y)$, the firm's measure of relative risk aversion

$R = -\frac{U_{yy} Y}{U_y}$, the individual's measure of relative risk aversion

$\nu = \frac{U_{nnh}}{U_h}$, the elasticity of marginal disutility of labor

α = share of labor

Since h_1 increases with θ_1, if the two states are far enough apart, we will obtain over or full employment. (See Appendix B).

(b) Infinite elasticity of subsitution between leisure and consumption. In this case, the utility function is of the form

(12) $U = u(y - ah)$, $h \leq 1$, $\theta_2 < a < \theta_1$,

income and leisure are perfect substitutes. This peculiar utility function gives rise to corner solutions, where either individuals do no work or work to capacity. With this utility function, the self-selection constraints are <u>never</u> binding if firms are risk neutral, as illustrated in figure 6. (Note that equating the marginal utility of income means that:

$$y_1 - ah_1 = y_2 - ah_2,$$

individual's utility in the two states are equalized.) This result does not depend on assuming a linear technology, as figure 6b illustrates.[20] The unconstrained implicit contract is described by the tangencies of the iso-profit curves to the worker's indifference curve. The iso-profit curve for the good state is assumed to be always steeper than the corresponding iso-profit curve through the same point for the bad state. (Because of diminishing returns, the extra income that the firm can pay to a worker who works more diminish the more the individual works). It is clear that the self-selection constraints are always satisfied if $h_1 > h_2$.

When firms are risk averse, the marginal utility of income of workers in the two state will not be equalized. The implications of this can be seen in figure 6c for the utility function (12). Assume first that the firm has a linear technology. Then again, it is easy to show (figure 6) that the self-selection constraint is always satisfied.

If, however, we assume diminishing returns <u>and</u> risk averse firms, then, as figure 6d illustrates, it is possible that in the good state,

---

[20] Though the results that $h_1 = 1$ and $h_2 = 0$ clearly do.

firms will announce that it is in fact the bad (E_2 lies on a higher iso-profit function for state 1 than does E_1; E_2 entails fewer hours and less income than does E_1, but the decline in pay is greater than the loss in output.

(c) A Slight Generalization

If the individual has a utility function of the form

$$U = U(y - v(h)),$$

the indifference curves between income and leisure are no longer straight lines; however, the results obtained in the previous section still hold. This utility function has the property that leisure is neither normal or interior. In the first best equilibrium, with risk neutral firms,

$$U'_1 = U'_2$$

(where $U'_i = U^{i'}(y_i - v(h_i))$

so the individual is on the same indifference curve in both states. Hence, so long as the iso-profit curves satisfy the single crossing property, the self-selection constraints are satisfied in the first-best equilibrium. (See figure 6e).

(d) Risk Neutral Individuals

If individuals are risk neutral, then the objective of a long term contract is not to provide insurance to the worker. (Nor is it plausible to think of the function of the insurance contract as a mechanism for insuring the firm.) For a variety of reasons, long term contracts may save on transactions costs. Though the reason that workers and employers engage in long term contracts may have little to do with the provision of insurance, attitudes towards risk play an important role in the design of efficient contracts, in the presence of

asymmetric information.

With risk neutral individuals and firms, contracts which satisfy the self-selection constraint are easy to construct. All we require is that

$$\theta(h_1-h_2) > y_1 - y_2 > \theta_2(h_1-h_2)$$

Clearly, if the difference in productivity in the two states are low, differences in income will be low. Note that in the case where hours worked in the bad state exceed that in the good state, income paid in the bad state must exceed that in the good state.

<u>With risk neutral individuals, either there is full employment or underemployment.</u>[21]

Identical results obtain if there is a third party providing insurance to the firm, and workers are risk neutral.[22]

The hypothesis of low effective risk aversion on the part of

---

[21] For the case where the two states are near each other, the result follows directly from the calculations of Appendix B. More generally, with risk neutral individuals, the optimal employment contract is described by the solution to the problem

$$\max \rho^1(\theta_1 h_1 - y_1) + \rho^2(\theta_2 h_2 - y_2)$$

s.t.

$$U^1 + U^2 \leq U$$

and s.t.

$$\theta_1 h_1 - y_1 > \theta_1 h_2 - y_2$$

$$\theta_2 h_2 - y_2 > \theta h_1 - y_1$$

Letting λ_1 and λ_2 be the Lagrange multipliers associated with the self-selection constraints, we obtain

$$\rho^1 \cdot \theta_1 + \lambda_1 \theta_1 - \lambda_2 \theta_2 + \mu U^1 h = 0$$

(Footnote continued)

workers is not as implausible as it might seem at first glance. Our analysis has assumed that an individual's consumption is equal to his wage payment; that he cannot save or borrow (or does not receive implicit or explicit unemployment compensation.) In fact, most fluctuations are sufficiently small that it is easy for individuals to tide themselves over without any significant reduction in their consumption. (There is some recent econometric evidence suggesting that there is very limited variability in consumption, as opposed to the purchase of consumer durables.) If that is correct, then the model we have just presented seems more appropriate than one in which the worker is risk averse. As we emphasized here, workers are not entering long term contracts to have firms provide insurance, but to save on transactions costs. If firms are also risk neutral, the efficient contract give rise to neither unemployment or overemployment.

If the motivation of long term contracts is to save on transaction

---

[21] (continued)

$$\rho 2' \theta_2 - (\lambda_1 \theta_1 - \lambda_2 \theta_2) + \mu U^2 h = 0$$

$$\rho^{2\prime} - \rho^{1\prime} - 2(\lambda_1 + \lambda_2) = 0$$

$$\rho^{1\prime} - (\lambda_1 - \lambda_2) + \mu U^1{}_y = 0$$

$$\rho^{2\prime} + (\lambda_1 - \lambda_2) + \mu U^2{}_y = 0$$

Subtracting the last equation form the next to the last equation, we obtain

$$\rho^{1\prime} - \rho^{2\prime} = 2\lambda_1 + \mu[U^1{}_y - U^2{}_y] = 2\lambda,$$

under the hypothesis that $U_y{}^1 = U_y{}^2$. Hence, if

$$\lambda_1 = 0, \ \rho^{1\prime} = \rho^{2\prime}; \text{ and thus } \lambda_2 = 0$$

[22] The problem of the firm is identical to that described above, except
(Footnote continued)

costs, them simple contracts have a great advantage over complex contracts; the simplest contract is one which pre-specifices wages and hours.[23] The welfare loss (relative to the Walrasian equilibrium) can be shown to be of the order of magnitude $(\theta_1 - \theta_2)^3$

The welfare loss relative to the self-selection equilibrium is clearly smaller, and hence provided the fluctuations in productivity are not too large, it is plausible that a rigid wage-employment contract dominates the self-selection contract.

Though in the absence of enforcement problems (as we have assumed thus far) such fixed wage-employment contracts cannot explain unemployment, once we introduce enforcement problems, they may easily give rise to unemployment.[24]

3.6 Interpretation

The conditions which we have derived show, at the very least, that the asymmetric information theory of implicit contracts does not provide a robust explanation of unemployment: whether there exists

---

[22](continued)
now we write the self-selection constraints as

$$\theta_1 h_1 - y_1 - I > \theta_1 h_2 - y_2 + I$$

$$\theta_2 h_2 - y_2 + I > \theta_2 h_1 - y_1 - I$$

where I is the payment from the insurance company to the firm in the bad state (from the firm to the insurance company in the good state).

[23]In the general theory of self-selection, it is shown that with only two states of nature, a self-selection equilibrium dominates a pooling equilibrium, in the absence of transactions costs. See Stiglitz (forthcoming).

[24]See below, section 3.4. Note, however, that with enforcement problems all the calculations concerning the relative merits of self-selection equilibria versus rigid-wage employment equilibria need to be redone.

underemployment is sensitive to special assumptions concerning the nature of the utility function and the degree of risk aversion of firms.

The theory has several implications: first, it is reasonable to assume different individuals and firms differ in their utility functions and attitudes towards risk. Thus, if the theory provided the explanation of unemployment, one should expect to find different firms attracting different classes of workers. Risk neutral firms which experienced limited risk would always be characterized as having over-employment in good states, full employment in bad; risk neutral firms which had underemployment equlibria would have workers with special attitudes towards risk. Firms whose behavior (in other respects) seemed to indicate high degrees of risk aversion should experience under-employment in bad states.

Notice too that when there is a dispute about how bad circumstances are, unions try to persuade the firm to hire more workers at the given wage. Unions do not say to management, "we will be willing to accept a further wage cut if you prove that the state of nature is so bad, by throwing more of us out of work," but rather, "we will be willing to accept a further wage cut if you reduce your planned level of lay-offs." If these statements are to be taken at anything like their face value, they suggest that what is at issue is not a question of inducing truthful revelation of the state of nature.

Notice that the theory of asymmetric information implicit contract has turned the original theory of implicit contracts on its head. While the original theory was based on risk averse workers, and risk neutral firms, an essential part of the new theory is that firms are risk averse. While the original theory was used to explain wage rigidities

(and was successful in this, but not in explaining unemployment), the new theory entails greater fluctuations in employee's income than in the Walrasian equilibrium.

Though these observations lead me to the view that this extension (or reversal) of implicit contract theory does not provide an adequate basis for understanding macro-economic fluctuations, there are other, perhaps less contentious grounds on which to object to this version of the Implicit Contract Theory as an explanation of unemployment.

3.4 Further Objections to the Assymetric Information Implicit Contract Theory.

There are several further objections to the theory. The first concerns what is observable. The theory assumes both too much and too little. The theory assumes that the firm's hours and wage decisions are observable to the worker, while the state of nature (or other surrogates for the state of nature) are not. Though there is some presumption that each worker knows the amount of his own work and wages (though even here, there may be ambiguities arising from changes in jobs and the complexity of compensation packages), the worker is unlikely to be informed concerning total employment. The way that the underemployment contracts "force" the firm to tell the truth is to restrict the amount of labor the firm can hire. If profits depend on aggregate employment, then the firm can evade the force of these restrictions by hiring outside workers. The analysis, as presented, only applies if the firm cannot hire workers (or if there are significant fixed costs per worker, so that the restriction in hours worked per old worker is costly.)

Not only are workers likely to be uninformed about the total number

of employees, but even if they were informed about the total number of employees, it would not convey all the requisite information; first, the firm is concerned with the effective labor supply, and it can increase that keeping the number of employees unchanged by increasing the quality of the labor force; secondly, the boundaries of the firm are often ill-defined. The firm could - and would have an incentive to - sell off some of its underutilized assets to a subsidiary. Distinguishing such transfers from "legitimate" transfers would seem a difficult task for the average blue collar worker.

Non-observability of hours (employment) and state of nature. If hours (employment levels) as well as the state of nature are not observable, (and there is nothing else relevant upon which to make wages contingent) then the implicit contract will entail a fixed (wage, employment (hours)) contract. Clearly, if the worker cannot be bound to stay with the firm, and his productivity at different firms is perfectly correlated, he will leave whenever his wage is less than his marginal product; but in those states when the worker works for the firm his average marginal product (plus any insurance premium paid initially) must equal his wage. Thus, if there is no initial insurance premium, the only viable markets are spot markets. (Insurance premiums may take the form of the firm paying the individual less than his marginal product during the initial period in which the worker is hired.) Conversely, if firms cannot be bound not to fire workers, they will do so whenever the workers marginal product is less than the wage, and again, the only viable labor market is a spot market. When opportunity costs are not perfectly correlated with the workers productivity in his current firm, then there will exist a set of states of nature in which

the worker will not quit and the firm will not fire him.[25] With bilateral asymmetries of information, with hours worked and state of nature not observable, with imperfectly correlated shocks, and with contracts in which employment levels are not enforceable implicit contracts can give rise to unemployment. Critical to this result is the hypothesis that implicit contracts which specify wages are enforceable, while those which specify employment levels are not; the reasonableness of this hypothesis is addressed in section 3.6.[26]

Contingent Contracts. In the previous paragraphs, we have argued that the standard asymmetric information implicit contract theory assumes that variables which are probably not observable are. It also (implicitly) assumes that variables which are observable are not.

Though the state of nature is not perfectly observable, there are many good surrogates: Consider the recent recession in the car industry; sales of automobiles (of each firm), inventories, and foreign imports are readily observable. Moreover, aggregate variables, like unemployment, money supply, and the rate of inflation etc., which may be relevant to the demand facing an industry or firm are also readily observable. Direct indications of profits are observable. Though

---

[25] See Hall and Lazear.

[26] There are, of course, other versions of asymmetric information models, yielding somewhat different results. Cooper, for instance, has explored a model in which firms are uniformed concerning workers marginal rates of substitution. This model may easily give rise to under-employment equilibria. But the underlying assumption behind the analysis is unconvincing. While firms may plausibly be assumed to be imperfectly informed concerning an individual's opportunity costs, it is not plausible to assume, given that a worker remains on a job, that there are significant variations in his marginal rates of substitution between leisure and income. (An exception would be provided by individuals who hold second jobs, such as taxi-cab driving.)

audited profits and true economic profits may well differ, and making wages contingent upon audited profits would result in incentives to distort audited profits there are ways around this problem. First, wages may be made contingent upon profits of other firms (thus eliminating this "moral hazard problem", except if the firms collude together.). Secondly, wages may be made contingent upon the market value of the firm; in a well functioning capital market, these will reflect true profits, and not audited profits, and again, there will be no incentive for a small investor to alter his market valuation because of the wage consequences; thirdly, there may exist compensation plans, in which workers are paid partly with the shares of the firm, where the incentives for distorting behavior on the part of firm managers may be mitigated. The fact of the matter is that relatively few firms have (explicit) contracts employing such contingencies as part of wage determination. (The difficulty with implicit contract is that, since the terms are never explicit, one might claim that such contingencies are implicitly there.) Of course, even if such information were employed, there would be still a residual of imperfect information concerning firm specific shocks (states of nature). Thus, it is conceivable that the under or overemployment contracts we have described above would still be employed. The question remains, of course, of the extent of the unemployment which can be so explained.

3.6 Enforcement Problem.

A second, and equally damaging criticism of the theory concerns the Problem of Enforceability, which we noted earlier. The contracts described are one period contracts; yet the essence of the earlier

(symmetric information) implicit contract is the long term relationship between the employer and the employee. Although the older theory slid quietly over the issue of how such contracts were to be enforced, we can no longer do so. The contracts we have described above make sense only as one period explicit contracts, not as many period implicit contracts.

The distinction between explicit and implicit contracts has to do with the enforcement mechanism. Explicit contracts are primarily enforced through the legal system (or the threat of resort to the legal process.) Implicit contracts are enforced through two mechanisms: firms that violate their implicit contract loose their reputation, and find it difficult to recruit additional workers; and workers within the firm who feel that they have been cheated against may reduce their effort (or act in other ways which reduce the firm's profitability.) It should be emphasized that one enforcement mechanism does not dominate the other: to enforce a contract through the legal system requires that the alleged breach of contract can be verified; both sides may know, in their heart of hearts, that the contract has been breached, but the side that has breached it may also know that the other side cannot establish that fact to the court. In that circumstance, the firm may still loose its reputation; reputation may be an effective enforcement mechanism.

On the other hand, when interest rates are positive, firms must balance the gains from cheating (violating the implicit contract) with the losses (say from the loss of reputation); in sufficiently bad states, it may pay them to cheat, to violate the implicit contract. More accurately, the implicit contract must take into account the fact that certain contract provisions which could (under appropriate conditions of verifiability and observability) be enforced through the

legal mechanism cannot be enforced by a reputation mechanism. The view of implicit contracts we are putting forward is that they represent perfect equilibrium (wage, employment) strategies in a repeated game between employers and employees.[27]

What do equilibrium contracts (implicit or explicit) between workers and employers engaged in a long term relationship look like? The answer depends on what is observable, what the rate of interest is, and the structure of the shocks to the economy. Consider, for instance, an infinite period contract with no discounting, in a world in which the probability, each period, that the state of nature is good, is 1/2. Then by standard arguments it can be shown that the employer can be forced to tell the truth (almost) all of the time: in the long run, unless he announces that it is good half the time and bad half the time, he will be severely punished. And if he must announce good half the time and bad half the time, he would announce good when it is good, and bad when it is bad; moreover, the optimal contract under such circumstances entails full employment. Once again, implicit contracts (even under asymmetric information) have explained wage rigidity, but failed to explain unemployment.[28]

---

[27] We have drawn the line between implicit and explicit contracts somewhat more finely that we should. Even when there is an explicit contract, reliance may be placed on reputation as an enforcement mechanism, both because of the costs of resorting to the legal mechanism and because of the inevitable ambiguities associated with the terms of any contract. In addition, there is a long legal tradition stipulating that not all the terms of the contract which can be enforced through legal processes need to be made explicit, and not all the terms which are made explicit can be enforced through the legal process.

[28] Obviously, if the shocks at date t and $t+1$ are perfectly correlated, there is, in effect, only one shock; but then the assumption that the worker remains uninformed about what the state of nature is

(Footnote continued)

Though the structure of the optimal multi-period contracts with positive interest rates has not received detailed analysis, what is known about multi-period self selection problems suggests that their structure (entailing elaborate interperiod contingency provisions) will be even more unlike what is observed than the one period structure.[29]

Implicit contracts as perfect equilibria. One difficulty with analyzing implicit contracts as perfect equilibria is that it appears that there may be a superabundance of such equilibria. We are particularly interested, however, in ascertaining whether there are equilibrium which entail wage rigidity and unemployment. If contracts which ensure that wages do not vary can be enforced, why cannot contracts that ensure full employment be enforced? Let me suggest a tentative answer: recall earlier that we distinguished two reputational enforcement mechanisms, one based on the reaction of potential employees, the other of current employees. If the firm believes that it will be some time before it wishes to hire workers again, the present discounted value of the loss of the "outside" reputation may be relatively low; yet the firm may still be very concerned with the good will of its present employees, a central part of the efficiency wage

---

[28](continued)
becomes implausible. Eventually, it should learn that the firm faces bad prospects. So long as the state of nature eventually becomes known, the asymmetric information restrictions are irrelevant.

[29]From the general theory of self-selection, contracts may entail randomness (See Stiglitz (1982)); how such random contracts are to be enforced appears even more problematic than how the non-stochastic contracts discussed so far are to be enforced.

For a discussion of optimal multi-period contracts, see Stiglitz and Weiss (1983). Note that since some of the shocks to the economy are common environmental shocks, the optimal contract should employ information from other firms, as in Nalebuff-Stiglitz (1983a, 1983b).

hypothesis presented in Part II. Thus, an equilibrium contract may entail the firm paying all retained employees a fixed wage which entails a subsidy in all states; if the state is bad enough, the firm will reduce its loses by firing (laying off) a fraction of its workers.[30]

3.7 Lay-offs and Contract Complexity

The version of the implicit contract theory with asymmetric information we have presented explains at best work reductions (work sharing) but not layoffs, and not unemployment. Those versions of the theory which attempt to explain layoffs usually simply assume it: the worker is assumed either to work or not to work, so that reductions in work can only show up as layoffs. As we argued earlier, whether with asymmetric or symmetric information efficient contracts should entail work sharing. Moreover, since they do not provide an explanation of job hiring, knowing the determinants of job lay-offs provides an explanation of unemployment only when it is assumed that labor is immobile.

A final unattractive feature of most versions of asymmetric information implicit contracts is there complexity. If there are many states of nature, the efficient contract will entail specifying, for each set of observable variables, an hours, wage schedule; each schedule

---

[30] We have focused here on one side of the enforcement problem: ensuring that the firm does not cheat on the worker. There is another side which has received attention: the worker who has been subsidized in the bad state leaving the firm in the good state. If the worker always left the firm whenever his wage was less than the value of his marginal product, the firm would never be able to pay the worker more than his marginal product in the good states. There is an obvious solution: make the individual pay his insurance premium before he starts to work; equivalently, since workers are normally hired in good states, pay the worker less than his marginal product the first period of employment. See Holmstrom.

will have many points on it; the schedule is likely to be highly non-linear. (Specifying the contract as an explicit contract seems hard enough; knowing whether such an implicit contract had been broken would seem nigh impossible.)

In the next two sections, we present two models which remedy these two deficiencies.

4. Complexity of Contracts.

So far, we have not been able to elicit out of the implicit contract theory a convincing basis of a theory of unemployment. The explanations provided in this and the next section are, I think, more convincing.

The first is based on the problem noted earlier on the complexity which can be encompassed within an (explicit or, a fortiori, implicit) contract. The contracts described earlier resulting in full employment required wage payments and hours to vary from state to state. If we restrict this even a little, we obtain the possibility of unemployment equilibria. For instance, Newbery and I have considered a simple macro-economic model in which the "shock" to the economy is the variability in export prices. All contracts are made contingent upon the export price. We show that with any linear indexing rule <u>even when the level of indexing is optimally</u> chosen there may be unemployment. Similar results obtain with other simple (log linear, quadratic, etc.) rules.

5. Theory of Moral Hazard and Labor Turnover.

Layoffs represent a non-price response to a market disturbance:

rather than lowering wages in the face of a decrease in demand, employment is rationed. It is natural, thus, to look for an explanation for this in other instances where markets are characterized by rationing; it is by now well known insurance markets in which adverse selection and moral hazard problems arise are, in general, characterized by quantity rationing (See Rothschild and Stiglitz (1976), Wilson (1977), Stiglitz and Weiss (1981), Arnott and Stiglitz (1983)).

Recall earlier we argued that a theory (of implicit contracts) which purports to explain unemployment must not only explain layoffs, but also why those who are layed off are not rehired. In general, the shocks facing different firms are not perfectly correlated. If information were costless, individuals who are at firms where the value of the marginal productivity of those workers has decreased would move to other firms; labor would be reallocated, until the value of the marginal productivity of all workers at all firms was the same. Note in this case the optimal contract would not require the firm provide any unemployment insurance; individuals would never be unemployed; the only insurance required would be wage insurance, and it need not be provided by the firm. Now assume that information (search) is costly and the process of gathering information is stochastic, so that some individuals who search are unsuccessful.

The risk that is to be insured is not just that the firm has a bad state; it is the risk that the firm has a bad state, that the individual searches, and is unsuccessful in finding a job.[31]

---

[31] Even if search were costly in time, but not in goods, and if it were always successful, if individuals had a separable utility function, an individual who happened to be in a low productivity firm and was layed

(Footnote continued)

If the firm provided complete wage and employment insurance, the individual at the low productivity firm would have no incentive to search. The more complete the insurance provided, the less incentives for search, and the less efficient the resulting resource allocation. The firm, of course, is not interested in the efficiency of the societal resource allocation; but by inducing individuals to move elsewhere in bad states, the firm saves the subsidy which is implicit in the wage contract in those states. If search were observable, the optimal contract would provide for unemployment compensation which was conditional upon the level of search. But search is not observable, and this is what gives rise to the moral hazard problem.

There are two instruments which are available to the firm for inducing search; one of them is to lower the wage of individuals who are retained, and the other is to lay off workers. In general, Arnott, Hosios, and Stiglitz (1983) show that both instruments will be employed. The nature of the search process (the individual either does or does not obtain employment) introduces a natural non-convexity into the problem, which implies that even if all individuals are identical, it pays to lay off some workers; it is not optimal to rely simply on work sharing. But in addition, some workers are better than others, and for a number of reasons (both information, equity, and institutional) it may not be possible to differentiate wages among them. Lowering wages results in a differentially higher quit rate among the high quality workers; layoffs

---

[31] (continued)
off would not be compensated by any severance pay. Indeed, if leisure and goods are complements, the optimal contract would entail negative severance pay, were that feasible.

do not have a corresponding adverse effect on qualtity.

On the other hand, when individuals differ in their search costs, reduced wages have an advantage over layoffs. When wages are reduced, those with lower search costs will be induced to search for a better job (where they are more productive); layoffs force both high and low search cost individuals to search.[32]

This model thus has accomplished what the other theories of implicit contracts could not do: it has provided an explanation of the simultaneous occurrence of wage reductions and layoffs (of on the job and off the job search); and it has provided a model of the labor market in which there is an equilibrium level of unemployment. (The theory is, of course, not inconsistent with the theories of implicit contracts with asymmetric information; the structure of the implicit contracts discussed by Arnott, Hosios, and Stiglitz, the specification of wages, hours, and layoff rates in good and bad states, depends on whether the state of nature is or is not observable by workers.)

Patterns of Unemployment

We noted in section 2 that a "good" theory of unemployment should explain not only the presence of unemployment, but also its pattern. Thus, it should explain layoffs as well as work sharing. Our model does this. It should also explain which workers get layed off. To some

---

[32]Several other papers have also focused on the issue of interfirm mobility. Holmstrom (1983) has considered its implications for the enforceability of contracts (with workers leaving in good states.) The paper closest in spirit to the Arnott-Hosios-Stiglitz paper is that by Geanakoplos and Ito (1981), but they do not focus on the moral hazard issues which are central to Arnott-Hosios-Stiglitz. For a partial survey, see Ito (1982).

extent, our model does this too. It is consistent with the use of the seniority system of layoffs; younger individuals may have both lower costs and greater benefits from search. But it does not seem to explain why certain groups, like woman and minorities, should experience more cyclical unemployment that do white males. The model presented in Part II does provide an explanation both for the use of layoffs and for the concentration of layoffs among certain groups.

The Paradox of the Preference for Being Layed Off

The model has one other failing: it does not provide a satisfactory resolution of the seeming paradox which arises in many implicit contract theories, in which the workers who are layed off are better off than those who are not. Consider, for instance, the case of a separable utility function; with full insurance, and no opportunities for being rehired, the pay of those layed off must be the same as that of those retained; but since those layed off enjoy more leisure, their total utility is higher. All workers would be petitioning to be layed off.[33][34]

If (i) there is some probability that a worker[35][36] will be

---

[33]Obviously if U_{yh} is positive enough, the retained workers will be better off than layed off workers. In the terminology of Arnott and Stiglitz, being layed off is a marginal utility increasing or decreasing accident depending on whether $U_{yh} \gtrless 0$; with full insurance, marginal utility increasing accidents always lead the individual experiencing the accident to be better off.

[34]The result is even stronger if leisure and consumption, are complements, in the sense that $U_{yh} < 0$. Those who are layed off have more time to enjoy their goods, and thus receive more goods than those who are left to work. In Appendix A we show that if leisure is normal, layed off workers will be better off.

rehired, (ii) we ignore the effect of unemployment benefits on search; and (iii) the firm cannot monitor whether or not the individual is rehired, and so must set its unemployment benefit as a fixed payment, than the payment will be set to equate the expected marginal utilities; that is, if we assume a separable utility function, and let y_r be the individual's income if he is retained (in the bad state), b be his lay-off benefit, and y be his income if he is rehired (a random variable) then b is set so that

$$Eu'(y + b) = u'(y_r)$$

Whether the expected utility of income is greater or less for the laid off worker than the retained worker depends on the shape of the utility function, on whether utility is a concave or convex function of marginal utility. While with a quadratic utility function the retained worker is better off, a necessary and sufficient condition for the worker to be worse off is that there is decreasing absolute risk aversion. (It is straightforward to derive more general conditions with non-separable utility functions under which retained workers are better off.)

Now we must take into account the fact that, in the bad state, the pay of the layed off worker affects their quit propensity. By

---

[35] Though there are a few instances of "inverse seniority" -- where unemployment benefits plus layoff pay is sufficiently generous that workers with greater seniority prefer to be layed off -- this is more the exception than the rule. See, for instance, Bloom and Northrup (1977).

[36] In models with work sharing, the corresponding paradox is that workers' utility is higher in the bad state than in the good.

assumption, the firm is subsidizing workers in the bad state, so would like to encourage them to leave. Thus, the firm lowers the pay of the retained workers. This strenghtens the presumption that retained workers are worse off than laid off workers.

We have established the fact that workers who are layed off may be rehired means that it is possible that retained workers may be better off than layed off workers, under the plausible condition that leisure and consumption are complements or independent.[37] But whether retained workers are or are not better off remains dependent on specific properties of the utility function, and plausible utility functions still lead to the seemingly paradoxical results.

Our argument in favor of lowering wages for retained workers, to encourage them to search, ignored the effect of lowering wages on their productivity. This is not the first instance in which we have noted indirect effects of current wages on the costs facing firms: we earlier referred to the fact that to the extent that firms had to rely on reputation mechanisms to enforce contracts, firms had two motivations for paying high wages: to ensure the productivity of their current workers and to lower future recruitment costs.

These are but examples of a more general phenomena, which we refer to as the Efficiency Wage Hypothesis, which provides an alternative, and we believe more plausible, basis of the theory of unemployment, than the implicit contract theory.

---

[37] There are other ways of resolving this seeming paradox. For instance, as we noted earlier, if search is time intensive (but does not use goods), but always successful, then layed off workers will not be compensated for the loss of time associated with search. They will be worse off than retained workers.

II

Efficiency Wage Theories of Unemployment

6. The Basic Argument.

The basic hypothesis of the efficiency wage theories of unemployment is that the (net) productivity of workers is a function of the wage paid. If that is the case, then firms may be reluctant to lower wages, even in the face of an excess supply of labor; to do so might lower the productivity more than proportionately, so that total labor costs are actually increased. Competitive equilibrium is thus consistent with a situation in which there is an excess supply of laborers. The Law of Supply and Demand has been Repealed.

Moreover, since the relationship between productivity and wage may differ from industry to idustry, wages (for similar laborers) may differ across industries. The Law of the Single Price has also been repealed. (Thus, there may exist in equilibrium some sector in which these efficiency wage considerations are not relevant, in which a conventional competitive wage is paid, while in some other sector(s), higher wages are paid; these higher paid jobs are obviously rationed, and queues for these jobs may serve as equilibrating mechanism,[38] a substitute for the adjustments in wages which do not take place because of the effect of wages on productivity.)

---

[38] It seems to be merely a semantic quibble to claim that so long as there is some industry in which rationing does not exist, there cannot be involuntary unemployment. The models we construct have the property that individuals who are identical (or nearly identical) are treated differently, and have markedly different levels of expected utility.

It is important to note that the economy is not always characterized by unemployment,[39] only that it may be. And changes in the economy (the destruction of some capital, the change in technology, etc.) may move the economy from a full employment regime to an unemployment regime, or may change the equilibrium level of unemployment in the economy.

7. Alternative Explanations of the Wage-Productive Relationship

There are at least five different explanations of this phenomenona which have been discussed in the literature.

The earliest, noted in the development literature, was based on the hypothesis that, at least at low levels of nutrition, individual's productivity depended on their nutrition, which depended in turn on their pay. A productivity-wage curve of the form depicted in figure 6 was hypothesized by Leibenstein (1957) and subsequently analyzed in greater detail by Mirrless (1975) and Stiglitz (1976). If $\lambda(w)$ is the efficiency of a worker receiving a wage w, the firm chooses a wage which minimizes the wage costs per efficiency unit,

(13) $\min \quad w/\lambda(w)$

the solution to which entails

(14) $\lambda'(w^*) = \lambda/w^*$.

w^* is referred to as the efficiency wage and is depicted in figure 7 as the tangency between a line through the origin and the wage productivity curve.

Assume the aggregate production function of the industrial sector

---

[39] The Repeal of the Law of Supply and Demand is thus a selective repeal.

is of the form

$$Q = F(\lambda L)$$

where L is the number of workers, and λ is their efficiency. Let output of the sector be our numeraire; firms in the sector will pay a wage w^*, and hire workers up to the point where:

(15) $\lambda(w^*)F'(L\lambda(w^*))=w^*$

the (value of the) marginal product equals the real wage. Let $L^D(w^*)$ be the solution to (15). If $L^D(w^*) < L^S$, the supply of workers to the industrial sector, there will be unemployment: no firm will have any incentive to lower wages.

This version of the theory is useful in helping to bring out the basic structure of the argument, but nutritional considerations are of limited relevance for wage determination in more developed countries.

The second theory is based on labor turnover (Phelps (1970), Stiglitz (1974, 1982)): the lower the wage, the higher the rate of labor turnover; so long as the firm must bear some part of the turnover costs[40] this lowers the net productivity of a worker.

This can again easily give rise to unemployment. To see this most simply, assume that individuals leave firms for two reasons: they die, at an exponential rate, μ, or they quit, to obtain a higher paying job. Assume for simplicity, that they make s searches per unit time. Assume all firms paid the same wage w^*. Then any firm which paid a wage greater or equal to w^* would have a quit rate of μ, while any firm which paid a wage less than w^* would have a quit rate of $\mu + s$ (since

---

[40] So long as workers are risk averse, more risk averse than firms, firms will bear some of these costs. See Arnott and Stiglitz (1983). See also Hashimoto (1981).

every firm that the searcher sampled would be paying a wage in excess of the lower wage). The quit rate appears as in figure 8. The labor costs of the firm (per unit time) are

$$w + (r + q)T$$

where r is the rate of interest and q is the quit rate, and T is the training costs. ((r+q)T represents the amortization of the training costs.) Clearly, these are minimized at the wage w*. Thus, if the economy experiences a shock (a war which decreases the capital stock), and there is not a coordinated wage reduction, unemployment will develop. It does not pay any firm to lower its wage, even in the presence of unemployment (because the unemployed workers, though thankful now to get a lower paid job, will continue to search for a still better job.)

This model can easily be enriched to include individuals differing in their attitudes towards non-pecuniary characteristics of the firm, and searching for a good match between themselves and the firm. While in the simple version presented in the preceding paragraph, there is alway some full employment equilibrium (although many unemployment equilibria), in a more general theory, the only equilibria may entail unemployment.[41]

---

[41] Note that this problem could be resolved if individuals could be forced to pay all of their training costs but this would give rise to a moral hazard problem on the part of the firm: it would have an incentive to charge individuals for allegedly training them, and then fail to provide any training. This problem may be partly resolved by use of contests (Bhattacharya (1983)) or promotion ladders (Carmichael (1983)). Moreover, individuals may not have the capital to pay for all of their training costs. In any case, as we have noted, if individuals are risk averse, it is not optimal to force workers to bear the entire brunt of the risk that he may be ill suited to the firm. As a factual matter, firms do bear some of the turnover costs.

The third theory (Stiglitz (1976), Weiss (1980) Nalebuff and Stiglitz (forth-coming)) is based on imperfect information concerning the characteristics of workers (and/or an inability for legal or sociological reasons to differentiate wages among individuals whose characteristics differ.) The quality mix of applicants depends on the wages offered (and the quality mix of those who quit a firm depend on the wages paid to its current employees). In general, by paying higher wages, one obtains a higher quality labor force.[42]

The fourth theory is based on imperfect information concerning the actions of workers: the costs of monitoring them perfectly in the presence of such costs, firms must have some method of inducing "good behavior", some threat against workers that are caught shirking. If there were full employment, any worker who was fired would immediately simply find another job (at the same wage). To induce workers not to shirk, firms thus attempt to raise their wages relative to that paid by

---

[42]To obtain an efficiency wage, one must show more than a dependence of quality on wages, but that quality increases sufficiently fast with wages that it does not pay firms to lower their wages in the face of unemployment. Conditions under which this occurs have been derived in Stiglitz (1976, 1982).

Some have objected to this theory on the grounds that firms eventually learn individual's abilities; hence workers could be required to post a bond, which they would forfeit if it turned out later that they were not as good as they claimed. The objections to this are similar to those discussed earlier concerning the requirement to make individuals pay all of their training costs. Individuals may not have the capital to post the bond, and there is a moral hazard problem on the part of the firm. Some effective bond posting does occur, when individuals intially accept a lower wage, until they have proven themselves; but this imposes a further cost on individuals, in an inefficient intertemporal pattern of consumption.

Even when individuals are not well informed about their abilities, it may pay a firm to offer an above market wage, to elicit a larger applicant pool, among which it can select those which are best matched with the firm.

What is at issue here is in part an empirical question: do firms

(Footnote continued)

others. This has two consequences. If all firms were identical, then they would not raise their wages relative to each other, but as they raised wages, their demand for labor would decrease, and unemployment would increase. So long as the level of unemployment compensation is less than the market wage, the period of unemployment serves as a discipline device. The equilibrium level of unemployment may be depicted as in figure 9. Assume there are only two levels of effort (0,e) and that the unemployment compensation level is fixed. It seems plausible that the lower the level of unemployment, the higher the wage the firm must pay to make it worthwhile for an individual not to shirk. This no shirking wage is depicted as an upward sloping curve, increasing with the level of employment. The demand for labor (assuming that workers do not shirk) is the usual downward sloping function of the wage. The equilibrium is the intersection of the two.[43]

The major objection to this theory is that there are other methods of providing discipline to the labor force, e.g. forcing them to post bond. The limitations on these mechanisms are well known (see footnote 24 above), but it is worth noting that firms do employ indirect forms of bonding to some extent. Hall has noted that unemployment rates are highest among groups who, because of limited resources, are least able to post bonds.

The difference between the last two models and the standard

---

[42](continued)
believe that by lowering wages they will lower the quality of their labor force, with a loss in productivity exceeding the savings in wages?

[43]This version of the theory is developed by Shapiro and Stiglitz (1984). In their more general model, monitoring costs (and hence monitoring intensity) is endogenous. See also Calvo (1979).

competitive paradigm should be noted: the latter assumes that all individuals are paid strictly on a piece rate, that their actions can be perfectly monitored. If they perform the contracted for action, they get paid; if the don't, they don't. In fact, most individuals have at most a small fraction of their compensation depend directly on performance; in many cases, individual performance cannot be directly observed; at best, group performance can be observed, and then only with a lag. Firms do care about the quality of their labor force; they are worried about providing incentives to their workers (in the competitive paradigm, the firm could care less whether a worker decides to work or not; there is a competitive supply of workers readily available to perform any service; a worker who does not do what he has contracted to do is replaed by one who will).

The fifth set of theories are the sociological theories recently propounded by George Akerlof (1984). Though many of the patterns of behavior which he describes might equally well be described by one of the theories which we have provided here, there are some which are not. In particular, he noted that individuals' performance depends critically on whether they believe they are being "fairly" treated. We shall return to this observation later.

Though the five theories differ in a number of important respects, they have a common mathematical structure: the net productivity of a worker at the ith firm is a function of the wage paid by the firm, w_i, wage paid by other firms, w_i and the unemployment rate (or, more generally, the expected duration in the unemployment pool)

$$\lambda_i = \lambda_i (w_i, w_i, u)$$

The equilibrium level of unemployment, and wage structure may be

derived, and the consequences of various policy changes investigated, for alternative specifications of λ .

8. Patterns of Unemployment: A Further Implication of the Efficiency Wage Model.

The Efficiency wage models are not only consistent with there being unemployment in a competitive market equilbrium; they provide some insights into the patterns of observed unemployment. If groups differ in their relationships between wage and productivity, as illustrated in figure 10, equilibrium will be characterized by some groups being fully employed, other groups being partial employed, and still other groups being rationed out of the market. Changes in the aggregate demand for labor will have very large differential effects on the employment of different groups. (This is in contrast, for instance, to the standard theories, were the wages of different groups might be affected differentially, but there is not reason to expect, once wages have adjusted, differential unemployment rates among different groups.)

9. The Consequences of Policy changes.

The policy consequences may differ markedly depending on the explanation of the dependence of productivity on wages. Consider, for instance, an increase in the unemployment compensation. In the "shirking" version of the efficiency wage model, this results in firms having to raise their wages, to induce workers not to shirk (the penalty for being caught is smaller at any fixed wage and unemployed level); this in turn results in a higher equilibrium unemployment rate and a higher real wage. On the other hand, in the quality-efficiency wage

model, an increase in unemployment compensation may have a differential effect on the search intensities of individuals of different abilities, and thus shift the wage-productivity curve facing different firms. If low productivity workers' search is reduced relative to the high productivity workers, then the applicant productivity-wage curve may shift up, as in figure 11, with a consequent change in the wage (it may either increase or decrease), and increase in the demand for labor. In figure 12 we have depicted a case where there are low and high productivity individuals but the number of individuals of each productivity type who search for a job is affected by the level of unemployment compensation. In figure 12a the mean ability of those applying at high wages is increased; the wage is unchanged, but the demand for labor increased; while in figure 11b the mean ability of those applying at high wages is decreased sufficiently that the optimal wage is lowered (and unemployment is consequently eliminated.)

All of these models should be contrasted with the policy implication of the naive fixed price model, in which an increase in unemployment compensation would be unambiguously desirable, since wages and prices are (by assumption) unchanged, while the unemployment compensation increases aggregate demand.

Thus, though the structure of the equilibrium with the efficiency wage model and the fixed price model may look very similar (real wages do not respond to the presence of unemployment), and a careful general equilibrium analysis of an economy with efficiency wages would entail the same kind of detailed analysis of spill-overs and constraints that have characterized the fixed price literature, the comparative statics analysis and the analysis of the consequences of policy changes is

completely different. For in our models, real wages, though not falling to the market clearing levels, do respond to changes in policies.

10. The Efficiency Wage Hypothesis and Cyclical Fluctuations

We have seen in the previous section how the efficiency wage model can give rise to an equilibrium level of unemployment. Changes in the economy, in say the level of productivity of workers, in the capital stock, in the level of unemployment compensation, give rise to different levels of equilibrium unemployment. Thus, one approach to seeing whether (or the extent to which) the efficiency wage model provides insights into cycical fluctuations is to acertain the extent to which we can identify parameter changes which would result in a change in the equilibrium level of unemployment.

For instance, it is easy to show that a lowering of the demand curve for labor in figure 8 (for instance, as a result of the destruction of some of the country's capital stock) will result both in lower wages and a higher unemployment rate.

But the insights obtained from the efficiency wage hypothesis extend beyond those directly obtained from this kind of equilibrium analysis. Even if equilibrium were characterized by full employment, the economy's adjustments to disturbances may be such as to result (frequently) in unemployment.[44]

The argument may be seen most simply by considering the labor turnover model present above. We noted there that there exists a full

---

[44] The remarks in this and the next section are based on research in progress, and therefore are of a more speculative nature than results reported earlier in this paper.

employment equilibrium, in which all firms pay a wage, w^{**} at which the demand for labor equals the supply. Assume that some of the capital stock has been destroyed, so that there is a new full employment equilibrium with a lower real wage.

Consider now what happens if there is any friction in the wage setting process. Assume, in particular, that all contracts last for two periods (seconds, days, years?) and that some contracts come up for renewal in even periods, others in odd; assume further that firms can hire within a period. Then given even this slight amount of friction, the only equilibrium entails the real wage remaining unchanged, with the resulting increase in unemployment. (If the firms which have the option of lowering their wage did so, they would find that they experienced a higher quit rate, and lower profits.)[45]

The argument is, of course, more general than this simple example: since the optimal wage at one firm depends on the wage at the other firms (and the unemployment rate), given that (some firms) do not adjust, it does not pay others to adjust to the full equilibrium levels. Adjustments may occur, but they occur slowly.[46]

Furthermore, the private losses from not reducing wages in the efficiency wage model are of second order, while in the standard competitive model, they are first order; that is, if we write the

---

[45] See, for instance, J. Taylor, for a discussion of staggered contracts. The theory, as presented, does not explain why contracts should be staggered. Hosios has developed a theory, based on information costs, for why in equilibrium, contracts would not be signed simultaneously by all firms.

[46] For a more formal development of these ideas, see Stiglitz (forthcoming).

profits of the firm as a function of its wage, the wage paid by other firms, the amount of labor hired and a vector of other parameters,

$$\rho = \rho (w_i , \underline{w} , L , \eta)$$

in the efficiency wage model, the wage is set so that

$$\rho^i_w = 0$$

while in the standard competitive model, the wage is always the lowest wage at which the firm can obtain workers. Assume now that some paramenter η_j has changed and that the wage at which a firm can obtain workers is lowered. In the efficiency wage model, the wage will be lowered

$$\frac{dw_i}{d\eta_j} = - \rho^i_{w\eta} / \rho^i_{ww}$$

but the change in profits from this change in wages is zero (since $\rho_w = 0$). But in the standard competitive paradigm, if firms can obtain workers at lower wages, their profits are strictly increased.

As a result, one might expect some "almost rational" firms not to fully adjust their wages in response to disturbances which they face. (See Akerlof and Yellen (1983)). Such distortion in the behavior of one agent in the economy has, of course, its general equilibrium consequences, e.g. on prices elsewhere in the economy. But these are in the nature of pecuniary externalities, and were the economy initiallly

at a pareto efficient allocation (where rationing did not occur, e.g. because of efficiency wage considerations), the economy would still be pareto efficient; but under the circumstances described here, there are real welfare consequences of these failures on the part of some firms to adjust their wages.[47]

11. Nominal Versus Real Rigidities

There is a widespread belief among macro-economists that it is nominal rigidities, not real rigidities, which are to be explained. The evidence on this matter is not completely convincing. The few experiments on fully neutral monetary changes -- the change from old francs to new francs -- suggest that such changes may have relatively few real effects. Other forms of monetary injection are never uniformily distributed among the population, and it is easy to construct (not necessarily plausible) models in which non-uniform monetary injections will have real effects. Moreover, economies which have practiced extensive indexing (so money wages are not rigid) have experienced episodes of unemployment just as economies in which indexing is not so widely practiced.

Still, it is worth noting that two versions of the efficiency wage model may give rise to rigidities in money wages. In the labor turnover model, we noted an indeterminancy of equilibrium. If all firms were paying a wage w, it would pay them to continue to pay that wage, in spite of a change in the demand or supply of labor. The critical wage

---

[47] This result is an application of a more general result on market mediated extenalities in economies with incomplete markets and imperfect information (Greenwald and Stiglitz, 1983).

could be set in money or real terms. If each firm believes that all others are going to leave money wages unchanged, it would not pay any firm to change his (money) wage. In an economy which has not experienced rapid rates of inflation in the past, it may be natural to "fix" on the money rate; while in other economies, where indexing is more prevalent, it may be the real wage which is rigid.

The second explanation is based on Akerlof's sociological theories of the efficiency wage. If individuals come to believe that money wage reductions are unfair, then it does not pay firms to reduce their money wages. At one level, such an explanation seems unsatisfactory: why should individual come to view nominal wage changes of any moment? But if workers exhibit such irrationalities, it pays for firms to reflect those irrationalities in their wage setting policies. Such a theory may have some degree of plausibiltiy for the short run, in an economy which has had little experience with inflation; but it is unconvincing in the long run: but then do we have much evidence of nominal wage rigidities in such economies?

Conclusions

Those of us brought up in traditional Keynesian macro-economics were taught the importance of the assumption of wage rigidity. This was described as a Fact of Life, explained by vague reference to certain institutional factors. If wage rigidity is as central to the explanantion of unemployment as many modern renditions of Keynes seem to suggest, surely we need to explain this wage rigidity. The objective of this paper has been to investigate two important classes of explanations.

The first approach, the implicit contract theory, has had a long, but sad, history.[48]

The simpler models put forward a decade and a half ago provided an explanation of the lack of variability of real wages, but not an explanation of unemployment. The more complicated asymmetric information models were found unconvincing: they more easily gave rise to over-employment than underemployment, and the forms of contracts to be expected, were asymmetric information considerations paramount, are not observed. Other versions of the asymmetric information implicit contract model, explicitly long term in nature, may give rise to full employment. Two versions of the implicit contract model did give rise to unemployment: those with limitations on the complexity of the contracts which could be designed and those in which search was costly and could not be monitored. Though these may provide part of the explanantion of the observed patterns of wages and unemployed, of who becomes unemployed, and why unemployment takes the form of layoffs rather than work-sharing, additional insights may be obtained from the

---

[48] There is the suggestion in the implicit contract approach that, since the terms of the contract, which lead upon occasion to unemployment, are signed voluntarily, the unemployment generated is not really involuntary unemployment; and that since the contract maximized expected profits, given the level of expected utility of the worker, since the contract is, in this sense efficient, the market equilibrium is efficient. While the first question, whether the unemployment is voluntary or involuntary, is mainly a matter of semantics, the second conclusion, that the market equilibrium is pareto efficient, is as we have noted incorrect. There exist, in general, governmental interventions in the market which can both increase profits and increase the expected utility of workers. (See Newbery and Stiglitz, 1983, Arnott, Hosios, and Stiglitz, 1984.) (The result that -- even when contractual arrangments between parties are "locally efficient" -- the general equilibrium is not pareto efficient is more general, and holds whenever there are problems of moral hazard or adverse selection. See Greenwald and Stiglitz (1983).)

second approach, the efficiency wage models. These not only provide an explanation of the existence of unemployment equilibrium in competitive economies, they also provide part of the explanation of the observed patterns of unemployment, of who becomes unemployed, and why unemployment takes the form of layoffs rather than work-sharing. They provide an explanation for why different firms may pay similar labor different wages. They explain wage stickiness, both why firms may not loose much if they fail to adjust their wages, and why, when they adjust their wages optimally, they adjust them slowly.

The two theories are, of course, not mutally exclusive: employer-employee relationships are frequently long term relationships; what affects the quality of the labor force attracted to a firm, or the effort which a worker exerts, is not just the wage at the moment, but his life-time prospects.

The issues we have raised, concerning the nature of insurance, the presence of asymmetric information, the limitations of enforcement mechanisms, and restrictions on the degree of complexity of feasible contracts, are all relevant in the design of the contract. They must all be taken into account in explaining cyclical movements in wages, hours, and employments. Our argument is that while simple efficiency wage models can provide a plausible explanation of unemployment, simple versions of the implicit contract model (with or without asymmetric information) can do so only under highly restrictive and implausible conditions, and have some important counterfactual implications.

We have just begun the exploration of the full implications of these efficiency wage models. In the end, they may prove as unsatisfactory as the earlier versions of the implicit contract theory; empirical

predictions of the theory may be shown to be inconsistent with the observed facts. But for now, they seem to provide the most fruitful direction of research, in extending our understanding of wage rigidities, this central element in the Keynesian legacy.

Appendix A

Proof that layed-off workers are better off than retained workers, under hypothesis of normality of leisure.

We assume that firms provide supplementary unemployment benefits.

The optimal contract maximizes profits, subject to individual's obtaining an expected utility of U, i.e. it maximizes

$$\int_{A}[\theta\, h - y]f(s) + s\int_{A'} f(s)ds$$

subject to

$$\int_{A}[U(y_1 h)f(S)dS + \int_{A'}U(S + b_1 0)f(S)\, dS \gtrless U$$

as a function of his income, y, and hours worked, h; and $\theta(s)$ is the productivity of an individual in state S. It is straightforward to show that the solution to this entails

$$U_y = \text{constant};$$

individuals obtain perfect insurance, in the sense that their marginal utility of income in all states is the same. Differentiating (A.1), we obtain

$$\frac{dy}{dh} = -U_{hy}/U_{yy}$$

thus

$$\frac{dU}{dh} = U_h - \frac{U_y\, U_{hy}}{U_{yy}} \lessgtr 0 \text{ as leisure is } \begin{cases}\text{normal}\\ \text{inferior}\end{cases}$$

Appendix B

The optimal contract (without self-selection constraints) must satisfy

(B.1) $\theta_2 U_y(h_2,y_2)+U_h(h_2,y_2) = 0$

(B.2) $\theta_1 U_y(h_1,y_1)+U_h(h_1,y_1) = 0$

(B.3) $\dfrac{\rho^1,}{\rho^2,} = \dfrac{U_y(h_1,y_1)}{U_y(h_2,y_2)}$.

Without loss of generality, we let $y_2 = y_2^*$.

Then from (B.1), $h_2 = h_2^*$, where

$$\theta U_y(h_2^*, y_2^*)+U_h(h_2',y_2') = 0 \quad .$$

Let $U_y(h_2^*,y_2^*) = U_y^*$. Totally differentialy (B,2) and (B3) we obtain (letting $\rho'/\rho = A$)

$$\frac{U_{hh}}{U_h} \qquad \frac{U_{hy}}{U_h} \qquad dh_1 = 1 \qquad \frac{d\theta_1}{\theta_1}$$

$$-\theta_1 \ (A - U_{yh}) \qquad \frac{A-Uyy}{U_y} \qquad dy_1 \qquad Ah_1\theta_1$$

Let $c = \dfrac{-U_{yy}}{U_y}$, $b = \dfrac{U_{hy}}{U_h}$, $a = \dfrac{U_{hh}}{U_n}$.

$$\theta_1 D \quad \frac{dh_1}{d\theta_1} = (A+c) - bAh_1\theta_1 = A(1-bh_1\theta_1) + c$$

$$\theta_1 D \quad \frac{dy_1}{d\theta_1} = \theta\, aAh_1, + \theta_1 (A-b) = \theta_1 A\,[ah_1+1] - b\theta_1$$

where $D = (A+c)\, a + b\theta_1(A-b)$

$= A(a+b\theta_1) + ac - b^2\theta$

Define

$$S_i = \theta_i(h_1-h_2) - (y_1-y_2),\ i = 1,2$$

$$\frac{dS_1}{d\theta_1} = h_1 - h_2 + \frac{1}{D}\left[c+b - A\left(bh_1\theta_1 + ah_1\right)\right]$$

$$D\,\frac{dS_2}{d\theta_1} = \frac{(\theta_2 - 1)}{\theta_1}\left(A+c-Abh\,\theta_1\right) + \left(c+b - A(bh_1\theta_1+ah_1)\right]$$

Direct calculations verify that

$$\frac{dh}{dI} = \frac{(b+c)}{\theta(c+2b) + a} \gtrless 0 \text{ as } b+c \gtrless 0$$

$$\frac{dC}{dI} = \frac{a+b\theta}{\theta(c+2b) + a} \gtrless 0 \text{ as } a+b\theta \gtrless 0$$

where C = consumption.

Hence, at $\theta_1 = \theta_2$

$$\frac{dS_1}{d\theta_1} \gtrless 0 \text{ as } \frac{\alpha A}{1 - \alpha + A\alpha} \lessgtr \eta$$

where α = share of labor = $\frac{y}{\theta h}$

and $A = A(\theta h - y)$ = firm's measure of relative risk aversion

$\eta = -y\,\frac{d\ln h}{dI}$ = income elasticity of labor supply

It immediately follows that if A=0, the individual works more hours in the good state than in the bad ($h_1 > h_2$). It also immediately follows that for states near each other, whether there is over or underemployment simply depends on whether leisure is normal or inferior. Moreover, if leisure is normal, and firms are risk neutral, one always obtains either full employment or over employment, since $h_1 > h_2$.

Special cases: (a) separability

Separability ensures normality, and $\frac{dh_1}{d\theta_1} > 0$.

The condition for $\frac{dS_1}{d\theta_1} > 0$ can now be written (at $\theta_1 = \theta_2$

$$(1-\alpha)\,\nu - A\,\alpha\,\nu < 0$$

where

$C = U_{yy}Y$ = worker's relative risk aversion

$\nu = U_{hh}h/U_h$ = worker's elasticity of marginal disability of effort

(b) $U = U(y-v(h))$

This utility function has the property that $b+c = 0$.

$$a + b\,\theta = a - c\theta = \frac{v''}{v'} > 0$$

but since

$b = u''/u' < 0$, $dh_1/d\theta_1 > 0$, and if the two states are far enough apart, there will be full employment (the constraint $S_2 \leq 0$ is never violated, since $\frac{dS_2}{d\theta_1} < 0$).

REFERENCES

Akelof, G.A. and J.E. Stiglitz. June 1969. Capital, Wages and Structural Unemployment. Economic Journal. Vol LXXIX. pp. 269-81.

Akerlof, G. and H. Miyazaki. 1980. The Implicit Contract Theory of Unemployment Meets the Wage Bill Argument. Review of Economic Studies. XLVII. pp. 321-38.

Akerlof, George A. 1970. The Market for "Lemons": Qualitative Uncertainty and the Market Mechanism. Quarterly Journal of Economics, 84. pp. 488-500.

Akerlcf, George and Yellen, Janet. 1983. The Macroeconomic Consequences of Near Rational Rule of Thumb Behavior, mimeo. U.C. Berkeley.

Akerlof, George. 1984. Gift Exchange and Efficiency Wage Theory: Four Views. American Economic Review Proceedings.

Arnott, R. and A. Hosios, and J. Stiglitz. 1983. Implicit Contracts, Labor Mobility and Unemployment. mimeo. Princeton University. (revised version of a paper presented at NBER-NSF conference, Dec. 1980).

Arnott, R. and J. Stiglitz. 1981. Labor Turnover, Wage Structures and Moral Hazard. mimeo. Princeton University.

Arrow, K.J. 1964. The Role of Securities in the Optimal Allocation of Riskbearing. Review of Economic Studies. 31. pp. 91.

Azariadis, C. 1975. Implicit Contracts and Underemployment Equilibria. Journal of Political Economy. LXXXIII. pp. 1183-202.

Azariadis, C. 1979. Implicit Contracts and Related Topics: A Survey. Z. Hornstein et al, eds., The Economics of the Labour Market. London: HMSO.

Azariadis, C. 1983. Employment with Asymmetric Information. Quarterly Journal of Economics. XCVIII. pp. 157-72.

Azardias, C. and J.E. Stiglitz (1983), "Implicit Contracts and Fixed Price Equilibria," Quarterly Journal of Economics, 98, pp. 1-22.

Baily, M.N. 1974. Wages and Employment Under Certain Demands. Review of

Economic Studies. XLI. pp. 37-50.

Bhattacharya, S. 1983. Tournaments and Incentives: Heterogeneity and Essentiality. mimeo. University of California, Berkeley.

Bloom, Gordon F. and Northrup, Herbert R., Economics of Labor Relations, Richard D. Irwin, 1977

Borch, K. 1968. The Economics of Uncertainty. Princeton: Princeton University Press.

Burdett, Kenneth, and Dale T. Mortensen. 1980. "Search, Layoffs, and Labor Market Equilibrium," Journal of Political Economy, Vol. 88, No. 4, pp. 652-672.

Calvo, Guillermo and Wellisz, Stanislaw. 1979. Hierarchy, Ability and Income Distribution. Journal of Political Economy. 87. pp. 991-1010.

Calvo, Guillermo, and E.S. Phelps. 1977. Appendix: employment contingent wage contracts. Stabilization of the Domestic and International Economy. Vol. 5 ed. Karl Brunner and Allan H. Meltzer, Carnegie-Rochester Conference Series on Public Policy. Journal of Monetary Economics Supplement. pp. 160-68.

Calvo, Guillermo. May 1979. Quasi-walrasian theories of unemployment. American Economic Review. Vol. 69. No. 2. pp. 102-107.

Carmichael, Lorne. Spring, 1983. Firm-Specific Human Capital and Promotion Ladders. Bell Journal.

Chari, V.V. 1983. Involuntary Unemployment and Implicit Contracts. Quarterly Journal of Economics. XCVIII. pp. 107-22.

Cooper, R. 1981. Risk-Sharing and Productive Efficiency in Labor Contracts Under Bilateral Asymmetric Information. mimeo. University of Pennsylvania.

Cooper, R. 1983. A Note on Overemployment/Underemployment in Labor Contracts Under Asymmetric Information. Economic Letters, pp. 81-87.

Cooper, R. Worker Asymmetric Information and Involuntary Unemployment, Cowles Foundation Discussion Paper, 672R, April, 1984.

Debreu, G. 1959. Theory of Value. New York: Wiley.

Eaton, B. Curtis and White, William. 1982. Agent Compensation and the Limits of Bonding. Economic Inquiry. 20. pp. 330-43.

Gordon, D. F. 1974. A Neoclassical Theory of Keynesian Unemployment. Economic Inquiry. XII. pp. 431-59.

Grandmont, J.M. 1982. Money and Value. Cambridge University Press.

Green, J. and C. Kahn. 1983. Wage-Employment Contracts. Quarterly Journal of Economics. XCVIII. pp. 173-87.

Greenwald, B. and Stiglitz, J.E. 1983. Market Mediated Externalities: The Inefficiency of Competitive Economies with Imperfect Information and Incomplete Markets. mimeo. Princeton.

Grossman, S. and O. Hart. 1981. Implicit Contracts, Moral Hazard, and Unemployment. American Economic Review. (papers and proceedings). LXXI. pp. 301-07.

Grossman, S. and O. Hart. 1983. Implicit Contracts Under Asymmetric Information. Quarterly Journal of Economics. XCVIII pp. 1223-56.

Hall, R. 1980. Employment Fluctuations and Wage Rigidity. Brookings Papers on Economic Activity. pp. 91-124.

Hall, R. and D. Lilien. 1979. Efficient Wage Bargains Under Uncertain Supply and Demand. American Economic Review. LXIX. pp. 868-79.

Hall, R. and Lazear, E. 1982. The Excess Sensibility of Layoffs and Quits to Demand. NBER Working Paper. no. 864.

Hall, Robert. 1975. The Rigidity of Wages and the Persistence of Unemployment. Brookings Papers on Economic Activity. 2. pp. 301-35.

Harris, J.E. and M. Todaro. 1970. Migration, Unemployment and Development: a Two-Sector Analysis. American Economic Review. Vol. 60. No. 1. pp. 126-42.

Harris, M. and B. Holmstrom. 1982. A Theory of Wage Dynamics. Review of Economic Studies. XLIX. pp. 315-33.

Hart, O. 1975. On the Optimality of Equilibrium when the Market Structure is Incomplete. Journal of Economic Theory. XI. pp. 416-43.

Hart, O. 1983. Optimal Labor Contracts Under Asymmetric Information: An Introduction. Review of Economic Studies. 50. pp. 3-36.

Hashimoto, M. June 1981. Firm-Specific Human Capital as a Shared Investment. American Economic Review. pp. 475-82.

Holmstrom, B. 1983. Equilibrium Long-Term Labor Contracts. Quarterly Journal of Economics. XCVIII. pp. 23-54.

Hosios, A. 1981. Externalities, Efficiency and the Implicit Contract Theory of Employment. mimeo. Princeton University.

Imai, H., J. Geanakoplos, and T. Ito. 1981. Incomplete Insurance and Absolute Risk Aversion. Economic Letters, Vol. 8, 107-112.

Ito, T. 1982. "Implicit Contract Theory: A Critical Sur ey," Center for Economic Research, University of Minnesota.

Leibenstein, J. 1957. Economic Backwardness and Economic Growth. New York: Wiley.

Malcolmson, James. 1981. Unemployment and the Effciency Wage Hypothesis. Economic Journal. 91. pp. 848-66.

Malinvaud, E. 1977. The Theory of Unemployment Reconsidered. Oxford: Blackwell.

Malinvaud, E. 1980. Profitability and Unemployment. Cambridge University Press.

Mirrlees, J. 1975. A Pure Theory of Underdeveloped Economies. Agriculture in Development Theory. ed. L.A. Reynolds, New Haven: Yale University Press.

Moore, J. 1982. Optimal Labour Contracts When Workers have a Variety of Privately Observed Reservation Wages. mimeo. Birkbeek College.

Mortenson, D. 1978. On the Theory of Layoffs. mimeo. Northwestern University.

Nalebuff, B. and R. Zeckhauser. 1981. Involuntary Unemployment Reconsidered: Second-Best Contracting with Heterogeneous Firms and Workers. University of Wisconsin-Madison, Institute for Research on Poverty, Discussion Paper, #675-81.

Nalebuff, B. and Stiglitz, J.E. 1983. Prizes and Incentives: Towards a General Theory of Compensation and Competition. Bell Journal.

Nalebuff, B.J. and Stiglitz, J.E. 1983. Information, Competition and Markets. American Economic Review. pp. 278-283.

Nalebuff, B.J. and Stiglitz, J.E. (forthcoming). Quality and Prices. Quarterly Journal of Economics.

Neary, J.P. and J. Stiglitz. 1983. Toward a Reconstruction of Keynesian Economics: Expectations and Constrained Equilibria. Quarterly Journal of Economics. XCVII. pp. 199-28.

Negishi, T. 1979. Microeconomic Foundations of Keynesian Macroeconomics, Amsterdam, North-Holland.

Newbery, D. and J. Stiglitz. 1982. Wage Rigidity, Implicit Contracts and Economic Efficiency: Are Market Wages Too Flexible?, mimeo. Princeton University.

Newbery, D.M.G. and Stiglitz, J.E. 1981. The Theory of Commodity Price Stabilization. Oxford University Press.

Phelps, E.S. 1970. Micro-economic Foundations of Employment and Inflation Theory. New York.

Polemarchakis, H.M. January 1979. Implicit Contracts and Employment Theory. Review of Economic Studies. Vol. XLVI. pp. 97-108.

Polemarchakis, H. and L. Weiss. 1978. Fixed Wages, Layoffs, Unemployment

Compensation, and Welfare. American Economic Review. Vol. 68. pp. 909-917.

Rothschild, Michael A. and Stiglitz, Joseph E. 1976. Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Imperfect Information. Quarterly Journal of Economics, 90 no. 4. pp. 629-49.

Salop, S.C. August 1973. Wage Differentials in a Dynamic Theory of the Firm. Journal of Economic Theory. Vol. 6. No. 3. pp. 321-44.

Salop, Steven. 1979. A Model of the Natural Rate of Unemployment. American Economic Review. 69. pp. 117-25.

Sargent, R.J. 1979. Macroeconomic Theory. New York: Academic Press.

Schlicht, Ekkehart. 1978. Labour Turnover, Wage Structure and Natural Unemployment. Zeitschrift fur die Gesamte Staatswissenschaft. 134. pp. 337-46.

Shapiro, Carl and Stiglitz, Joseph. Equilibrium Unemployment as a Worker Discipline Device. American Economic Review. forthcoming.

Solow, R. and J. Stiglitz. 1968. Output, Employment and Wages in the Short run. Quarterly Journal of Economics. LXXXII. pp. 537-60.

Solow, Robert. 1979. Another Possible Source of Wage Stickiness. Journal of Macroeconomics. pp. 79-82.

Solow, Robert. 1981. On Theories of Unemployment. American Economic Review. 91. pp. 848-66.

Stiglitz, J. and Weiss, A. June 1981. Credit Rationing in Markets With Imperfect Information. American Economic Review. 71. pp. 393-410.

Stiglitz, J.E. 1973. Approaches to the Economics of Discrimination. American Economic Review. vol. 62 no. 2. pp. 287-295.

Stiglitz, J.E. 1974a. Theories of Discrimination and Economic Policy. Patterns of Racial Discrimination. ed. von Furstenberg. Lexington, Massachusetts: Lexington Publishing Company. pp. 5-26.

Stiglitz, J.E. 1974b. Equilibrium Wage Distributions. IMMSS Technical Report. No. 154. Stanford University. Forthcoming in Economic Journal.

Stiglitz, J.E. 1976b. The Efficiency Wage Hypothesis, Surplus Labor and The Distribution of Income in L.D.C.'s. Oxford Economic Papers. vol. 28. no. 2. pp. 185-207.

Stiglitz, J.E. 1978. Lectures in Macro-economics. mimeo. Oxford University.

Stiglitz, J.E. 1982a. The Wage-Productivity Hypothesis: Its Economic Consequences and Policy Implications. Paper presented at the New York Meetings of the American Economic Association.

Stiglitz, J.E. 1982b. The Inefficiency of the Stock Market Equilibrium. Review of Economic Studies. 49. pp. 241-261.

Stiglitz, J.E. 1983. Alternative Theories of Wage Determination and Unemployment in L.D.C.'s: The Efficiency Wage Model. M. Gersovitz, et al, editor.

Stiglitz, J.E. 1984. Price Rigidities and Market Structure. American Economic Review.

Stiglitz, J.E. and Weiss, A. 1983. Incentive Effects of Terminations: Applications to the Credit and Labor Markets. American Economic Review. 73. pp. 912-927.

Stiglitz, J.E., August 1976a. Prices and Queues as Screening Devices in Competitive Markets. IMMSS Technical Report no. 212. Stanford University.

Stoft, Steven. 1982. Cheat Threat Theory: An Explanation of Involuntary Unemployment. mimeo. Boston University.

Strand, J. 1983. The Structure of Implicit Contracts with Word of Mouth Reputational Enforcement. mimeo. University of Oslo.

Strand, J. 1983. The Structure and Efficiency of Reputational Labor Contracts. mimeo. Stanford University.

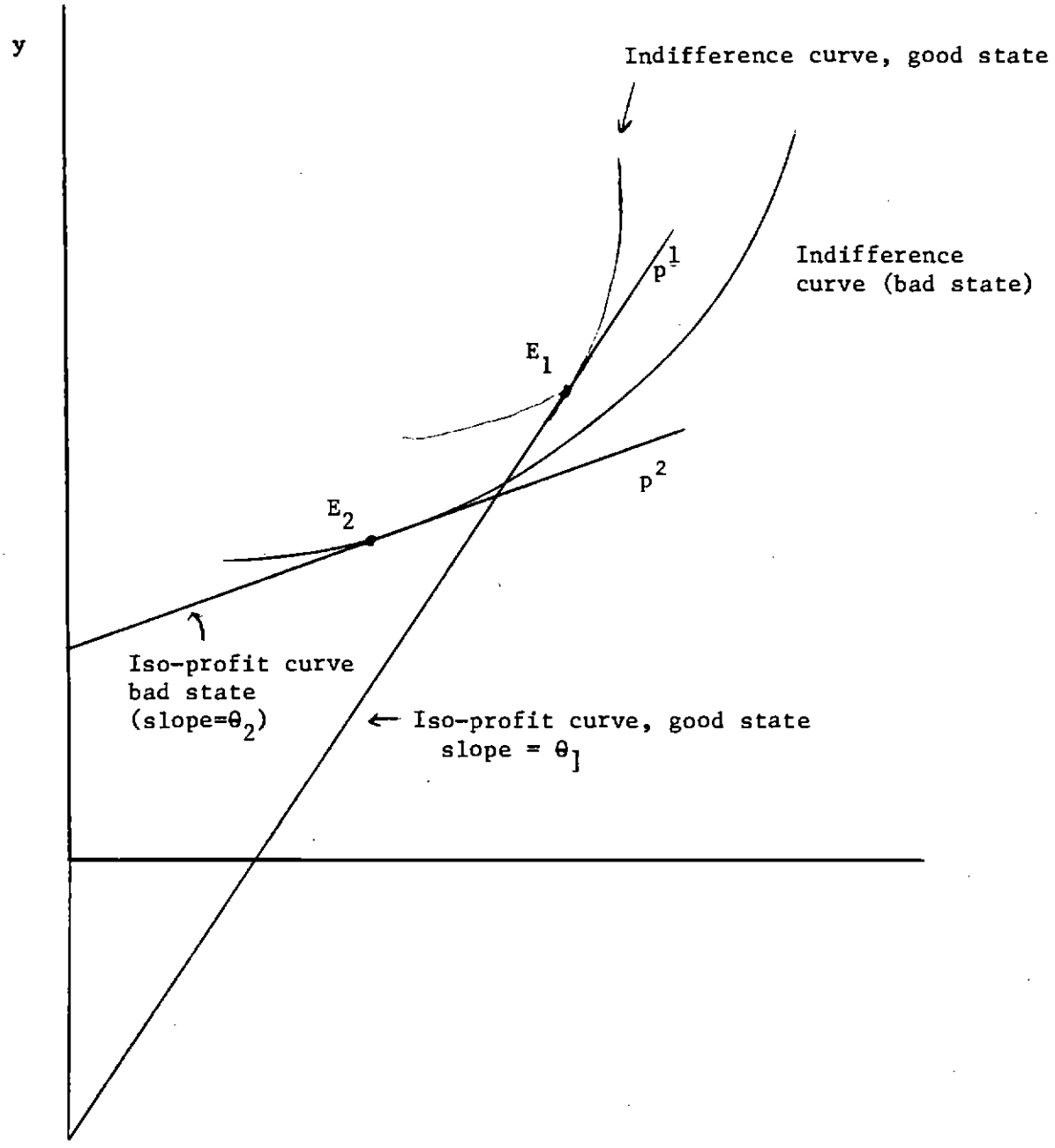
Weiss, Andrew. 1980. Job Queues and Layoffs in Labor Markets with

Flexible Wages. Journal of Political Economy. 88. pp. 526-538.

Wilson, Charles. 1977. A Model of Insurance Markets with Incomplete Information. Journal of Economic Theory, 16 no. 2. pp. 167-207.

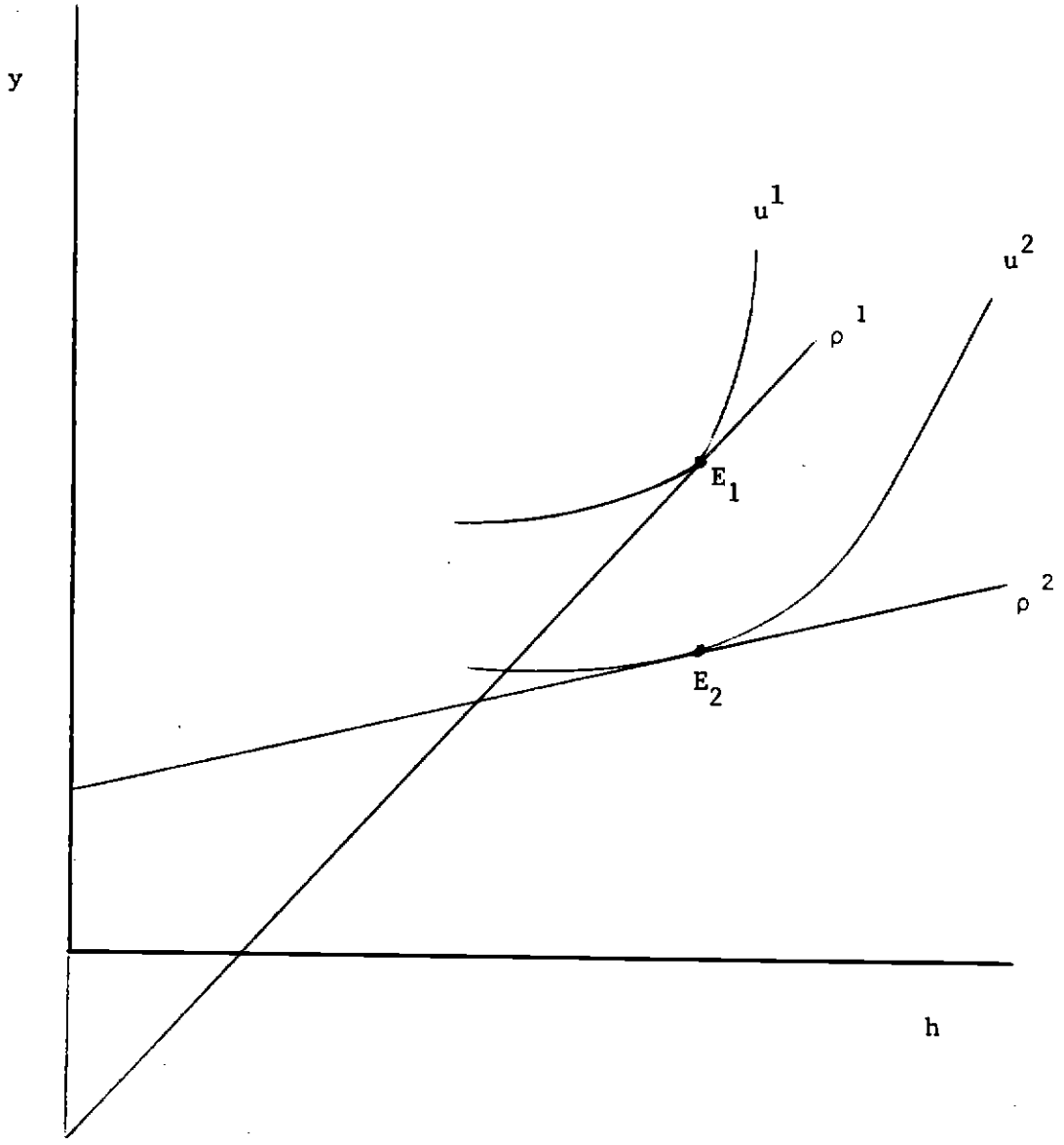
Yellen, J. 1984. Efficiency Wage Models of Unemployment. American Economic Review.

FIGURE 1a



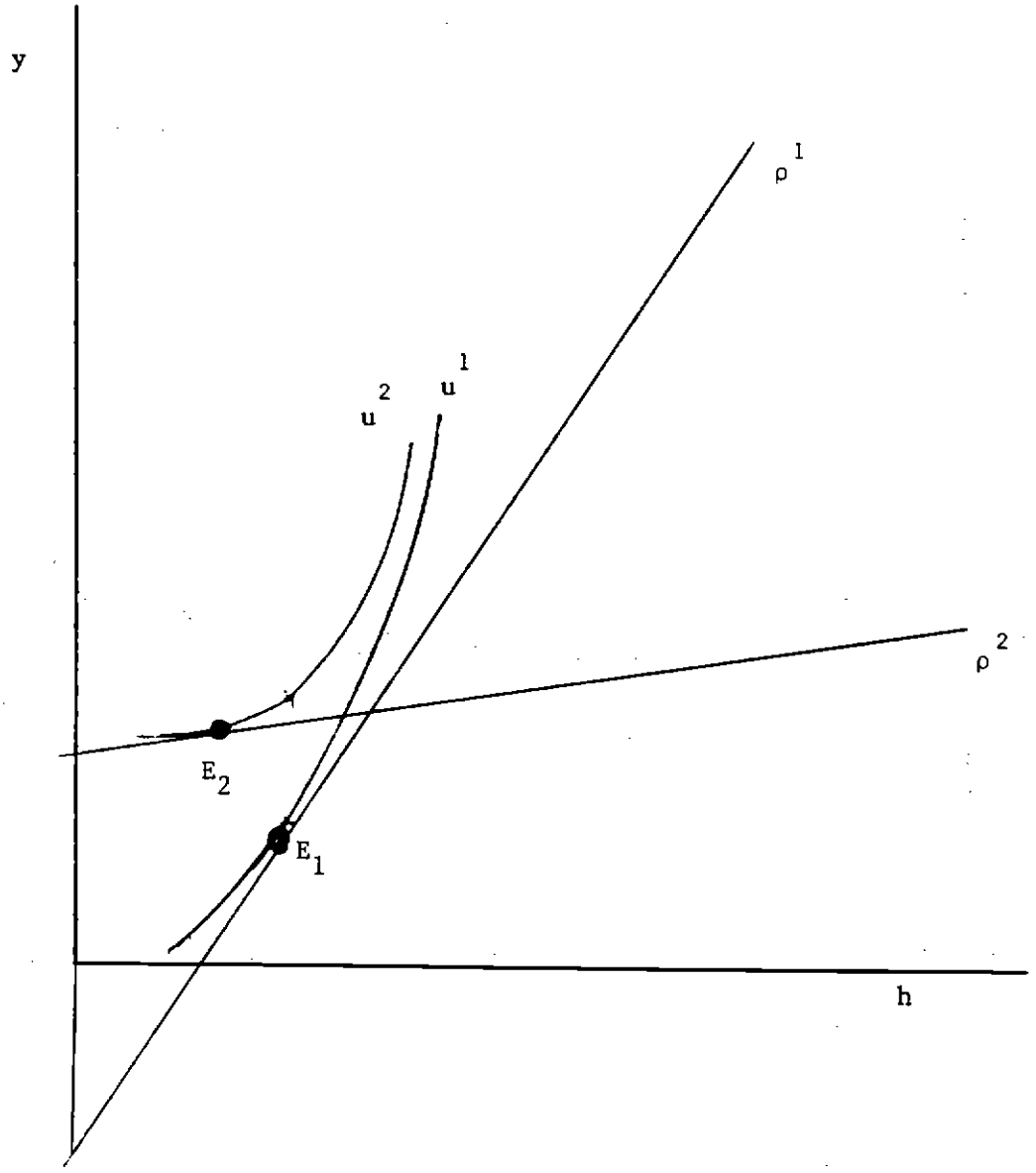
Self-selection constraints satisfied in first best equilibrium.

FIGURE 1b



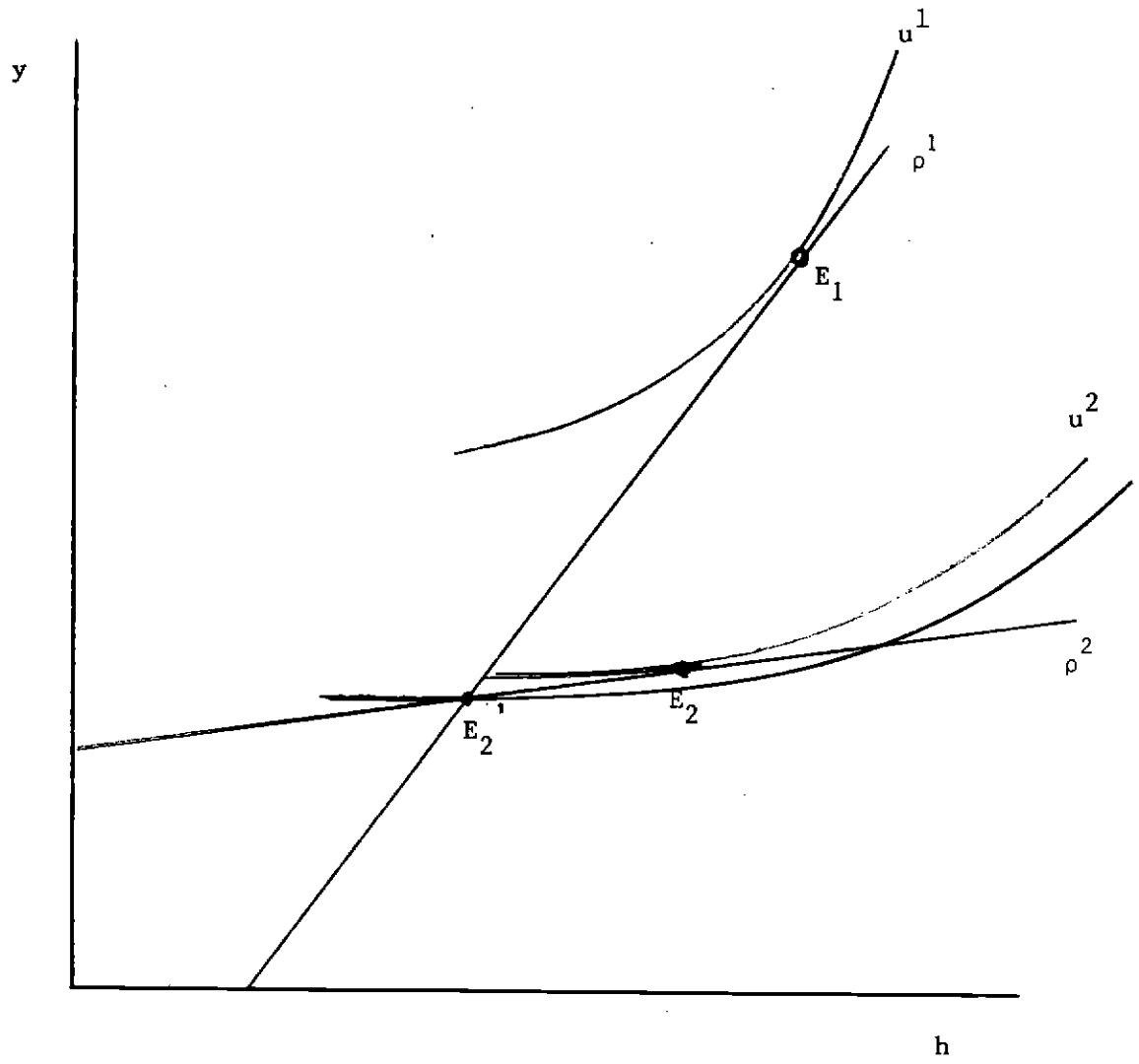
Self-selection constraint not satisfied in first best equilibrium. In the good state, the firm would prefer E_2 to E_1.

Figure 1c



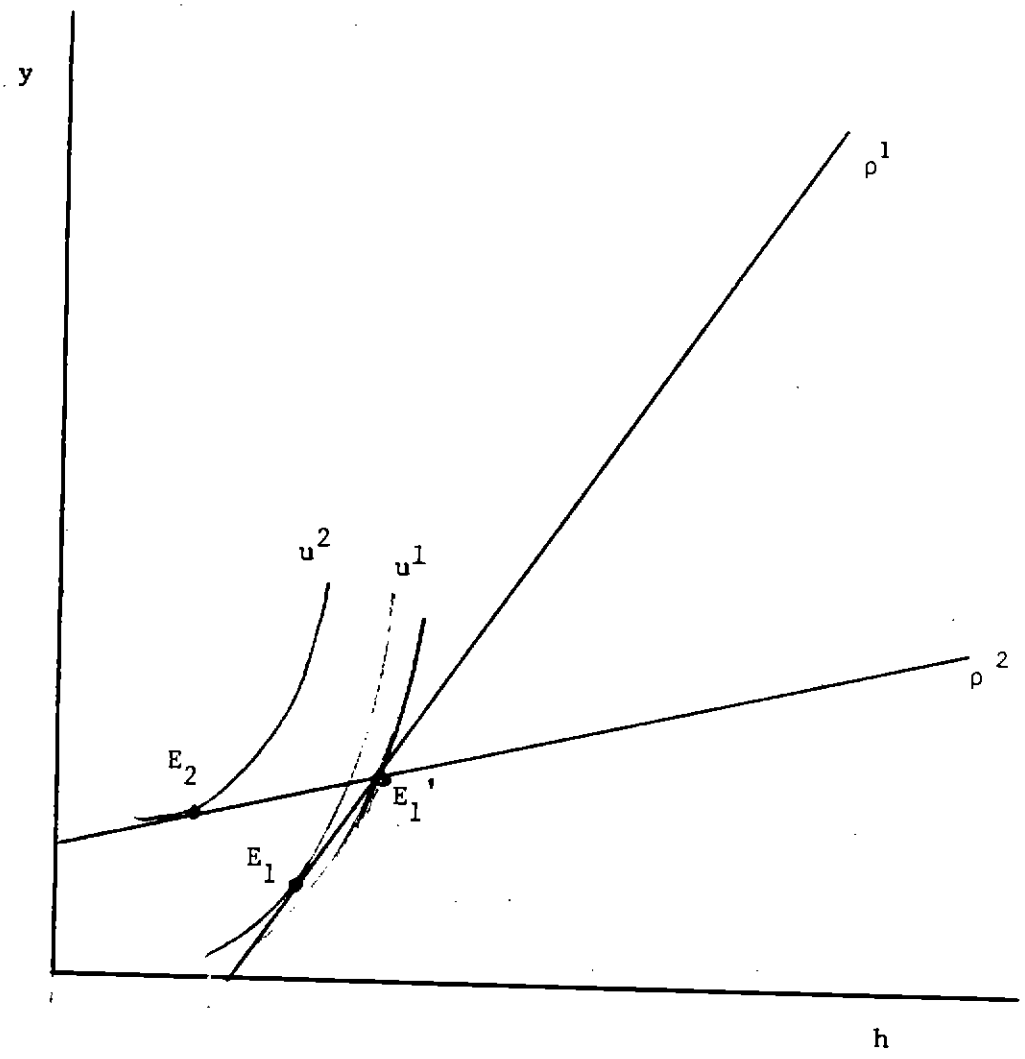
Self-selection constraint not satisfied in first best equilibrium: in the bad state, the firm would prefer E_1 to E_2. Note that individuals are better off in a bad state.

FIGURE 2a



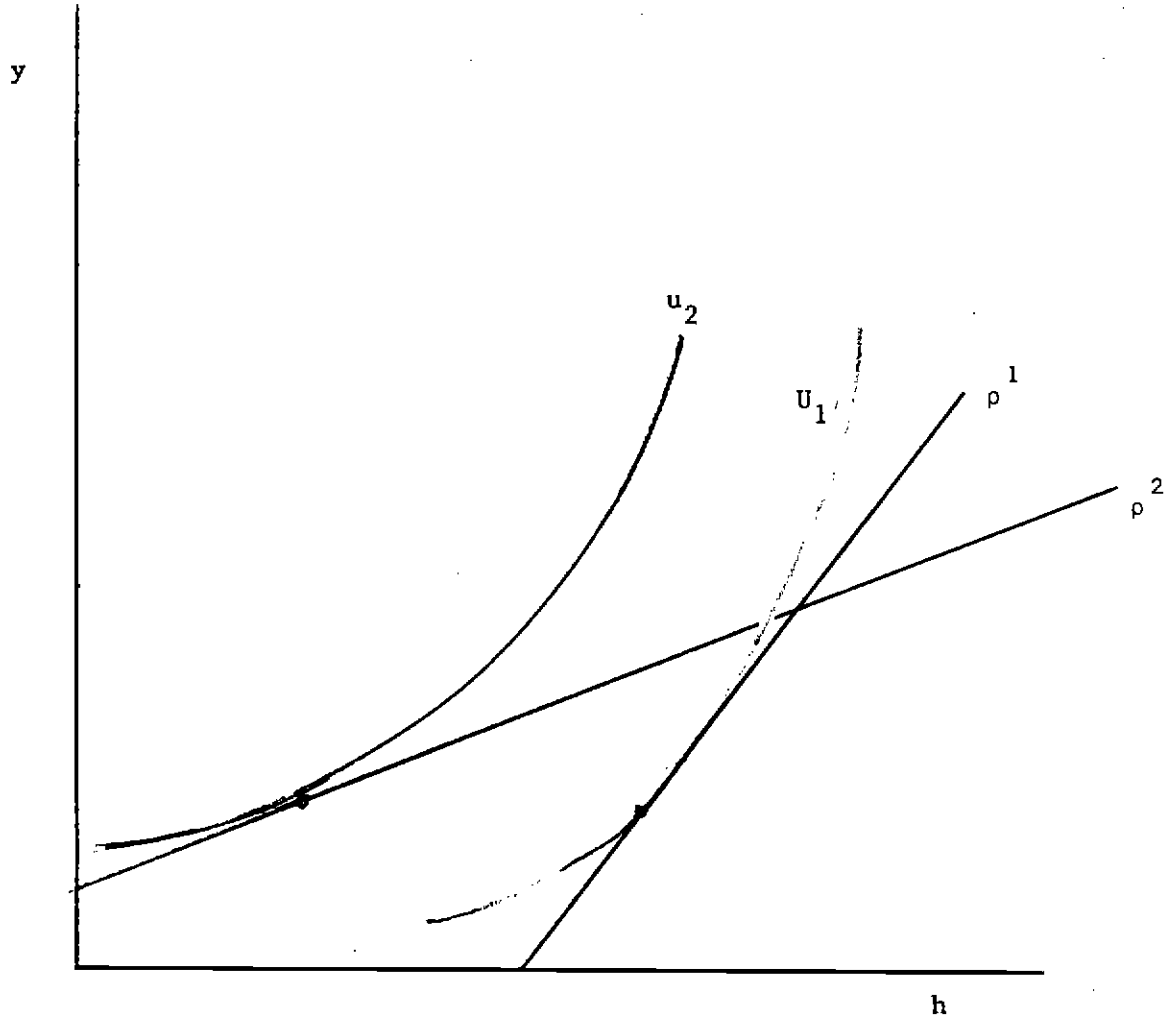
Contracts E_2' and E_1 satisfy the self-selection constraints, but there is underemployment in the bad state (to force truth telling, employment is restricted in bad state).

FIGURE 2b



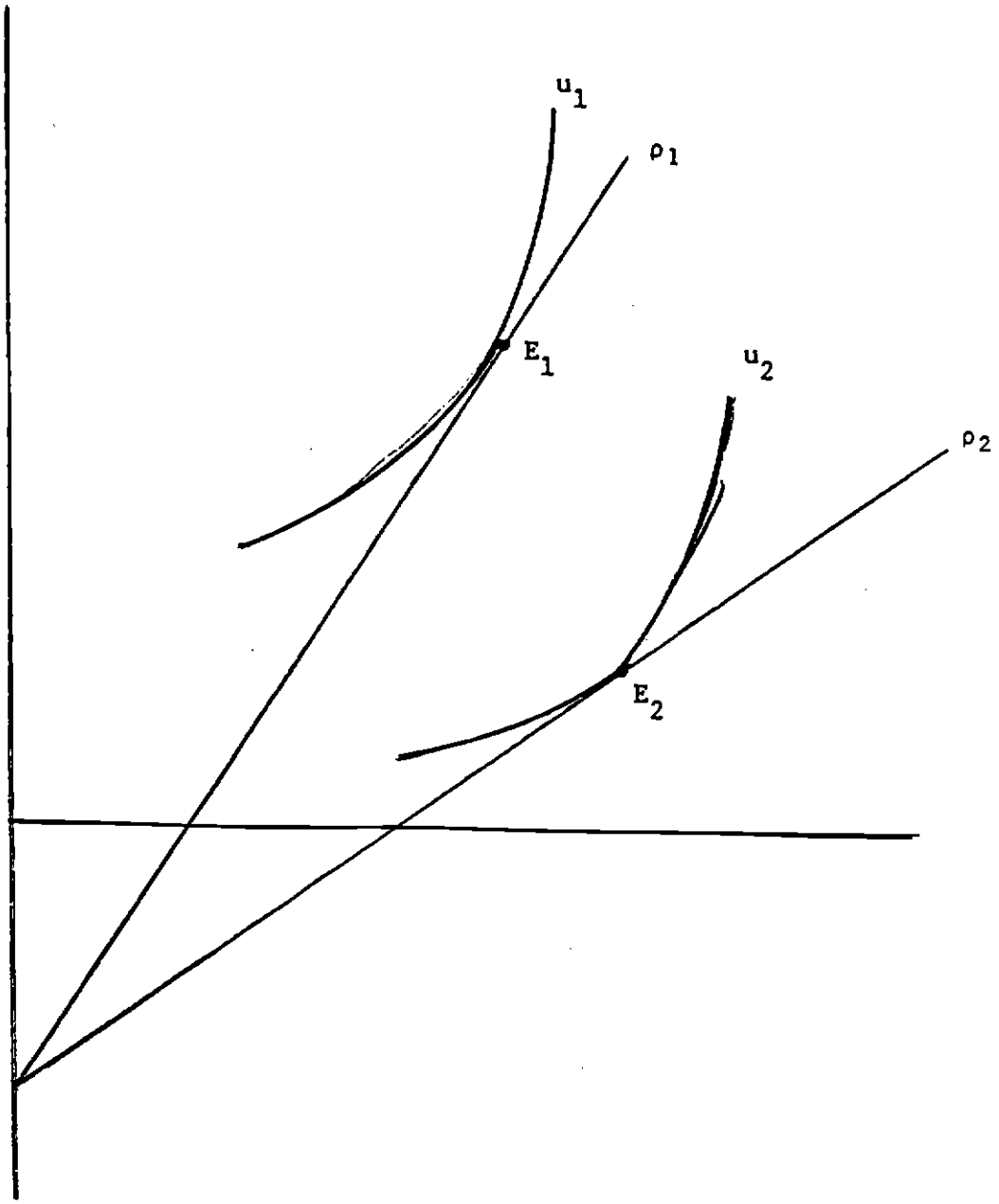
Contracts E_2 and E_1' satisfy the self-selection constraint, but there is overemployment in the good state (to force truth telling, employment is expanded in good state.)

FIGURE 3



If $U_{yh} \leq 0$, there will always be over-employment.

Figure 4



With infinitely risk averse firms (and finitely risk averse individuals) in the good state, the firm always prefers E_2 to E_1.

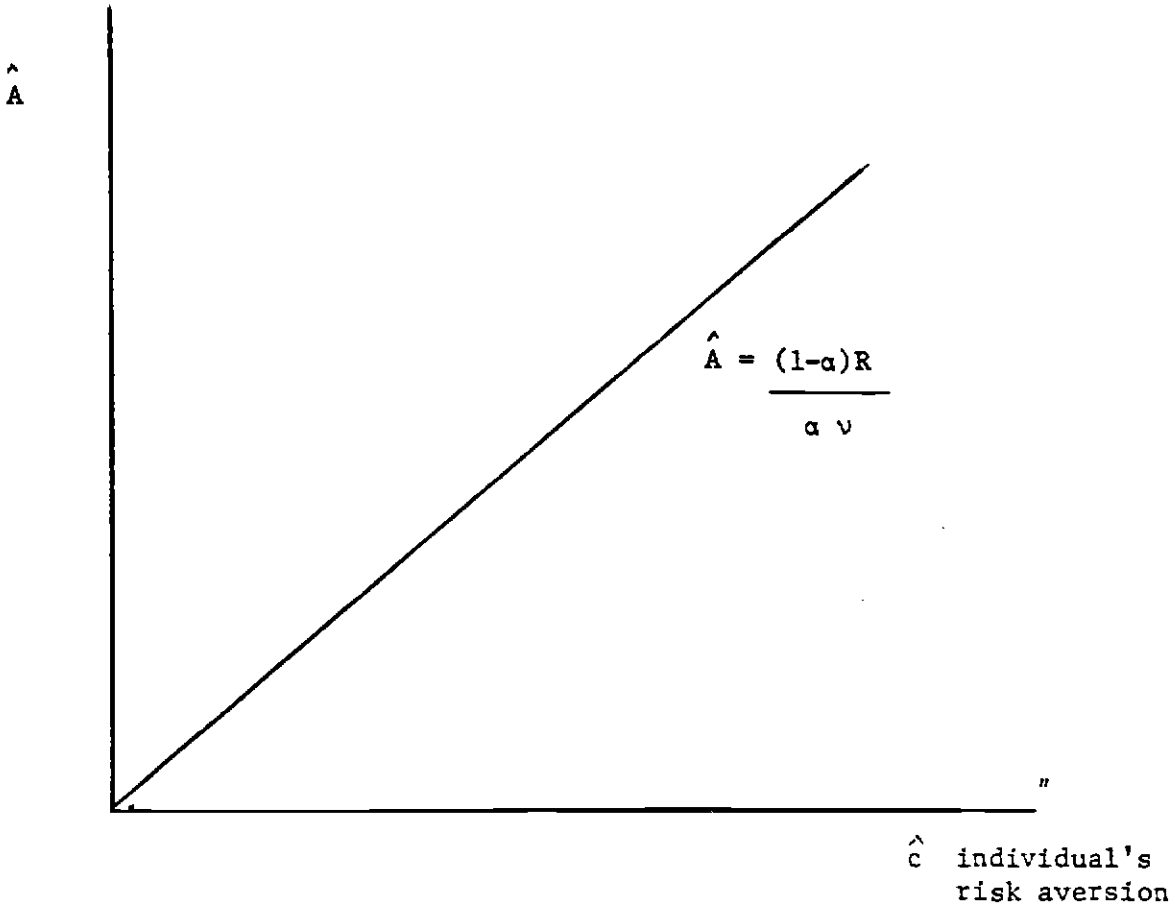
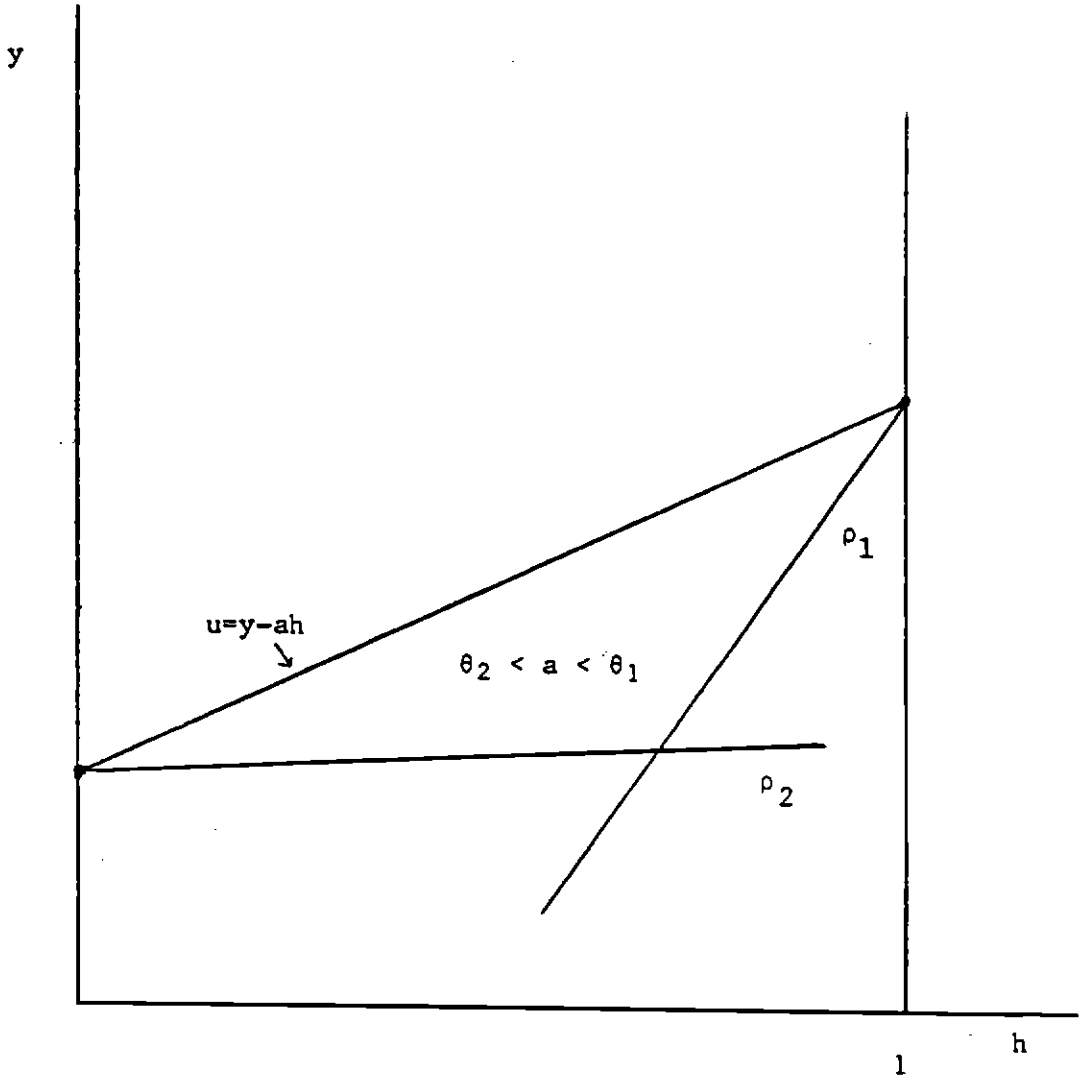


Figure 5

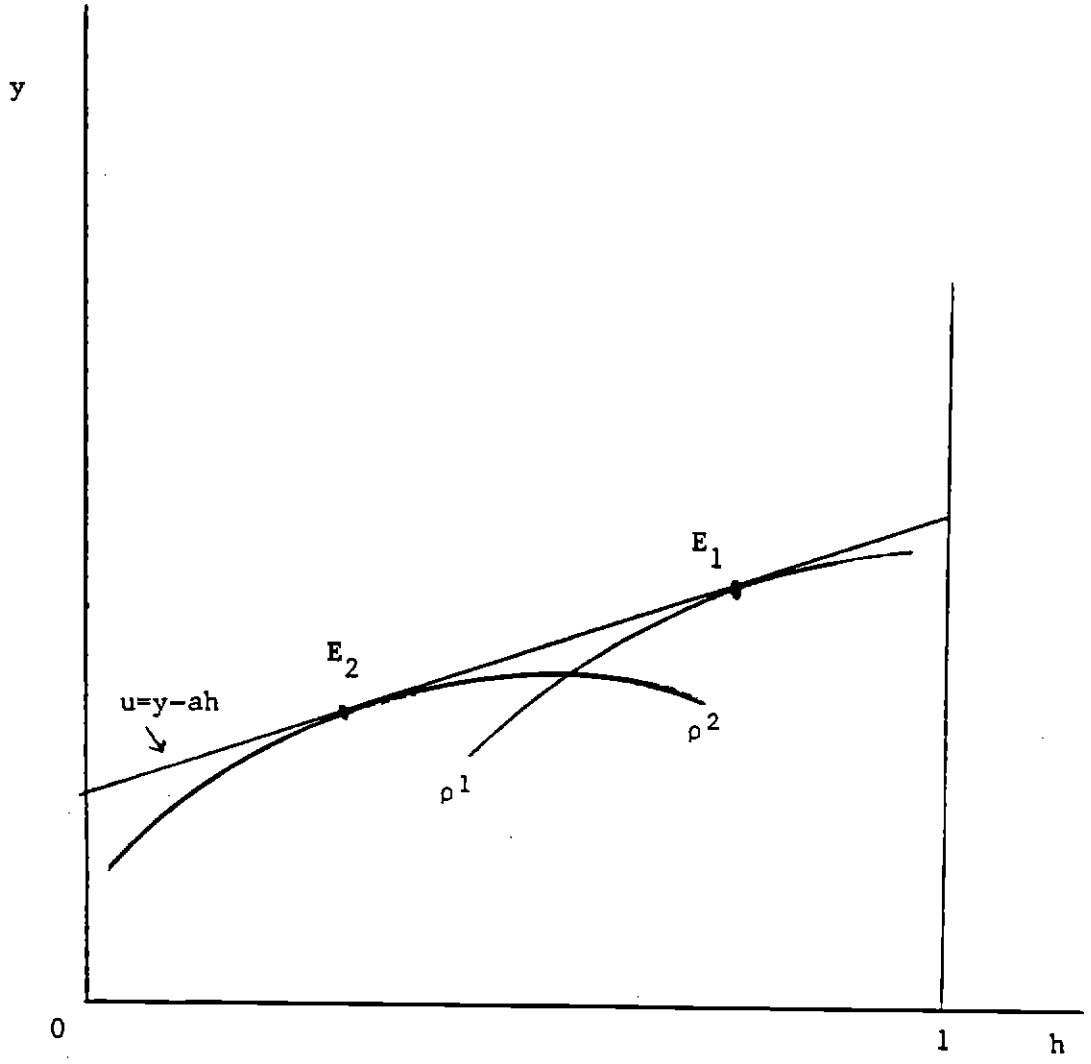
Critical value of firm relative risk aversion ($\hat{A}$), below which the only equilibria entail over or full employment. Increasing workers elasticity of marginal disutility of labor and reducing the workers degree of risk aversion make under-employment equilibria more likely.

FIGURE 6a



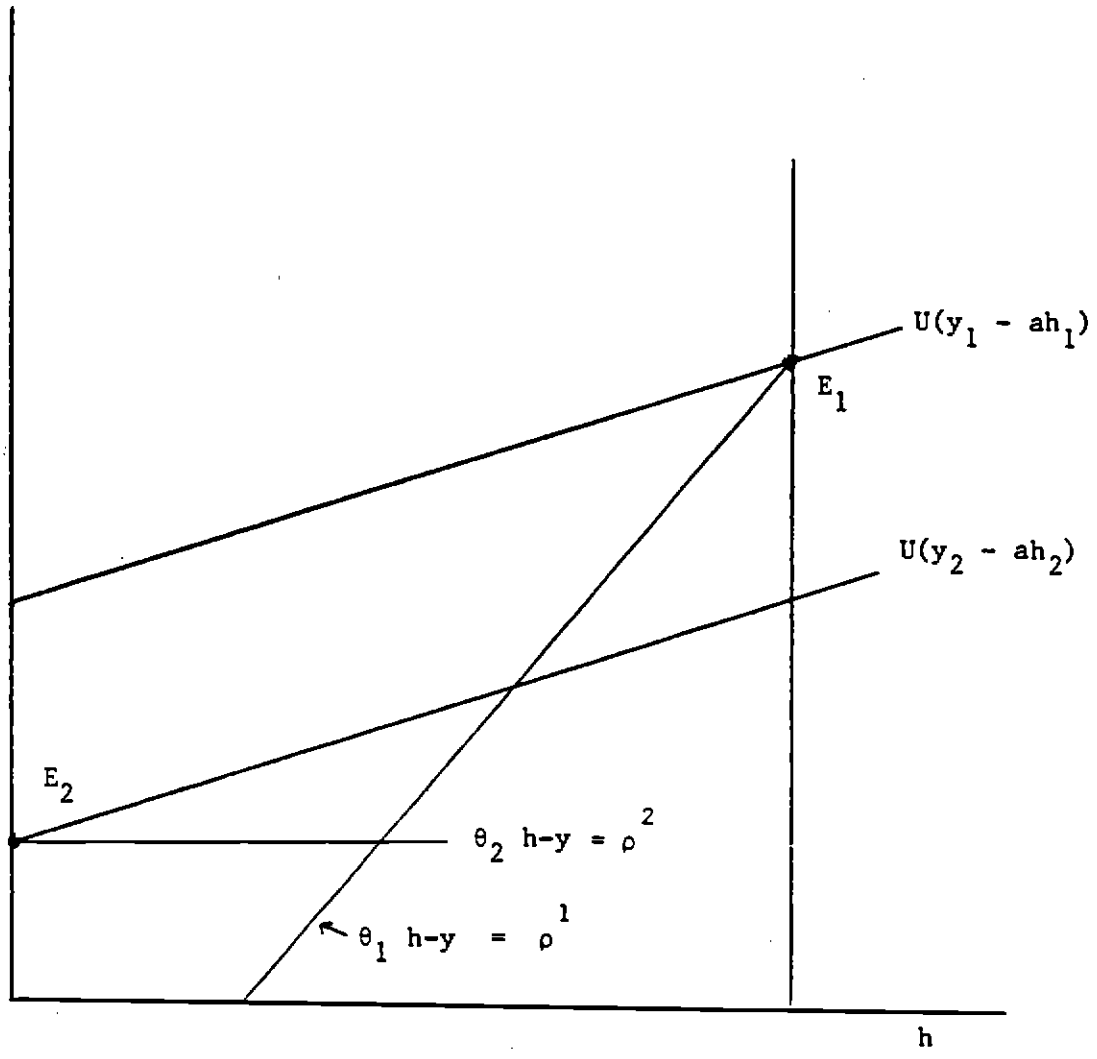
If leisure and consumption are perfect substitutes and firms are risk neutral, self-selection constraints are always satisfied in first best equilibrium.

FIGURE 6b



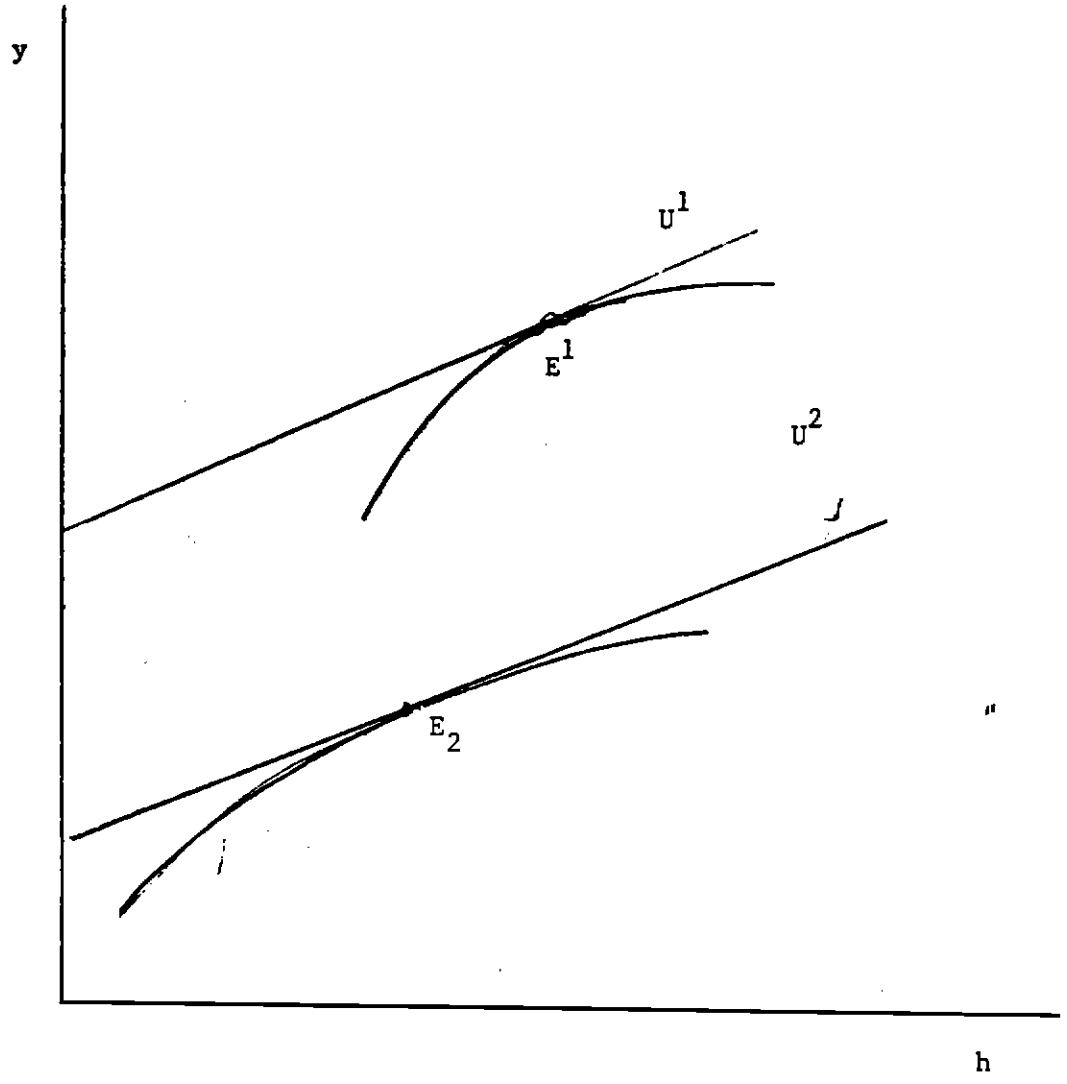
The self-selection constraints are not binding, even with a non-linear technology.

FIGURE 6c



Self-selection constraints are still satisfied if firms are risk averse.

FIGURE 6d



If firms are risk averse and there is diminishing returns (to the number of hours worked by each worker), then there may exist an underemployment equilibrium.

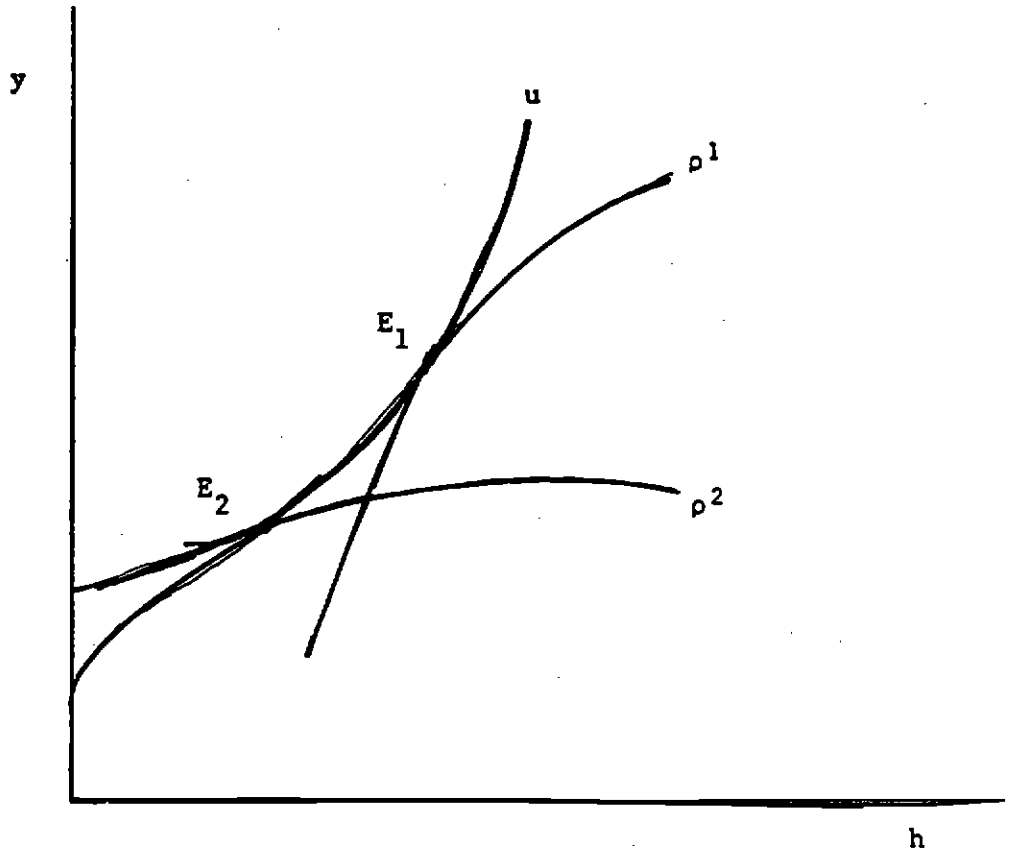
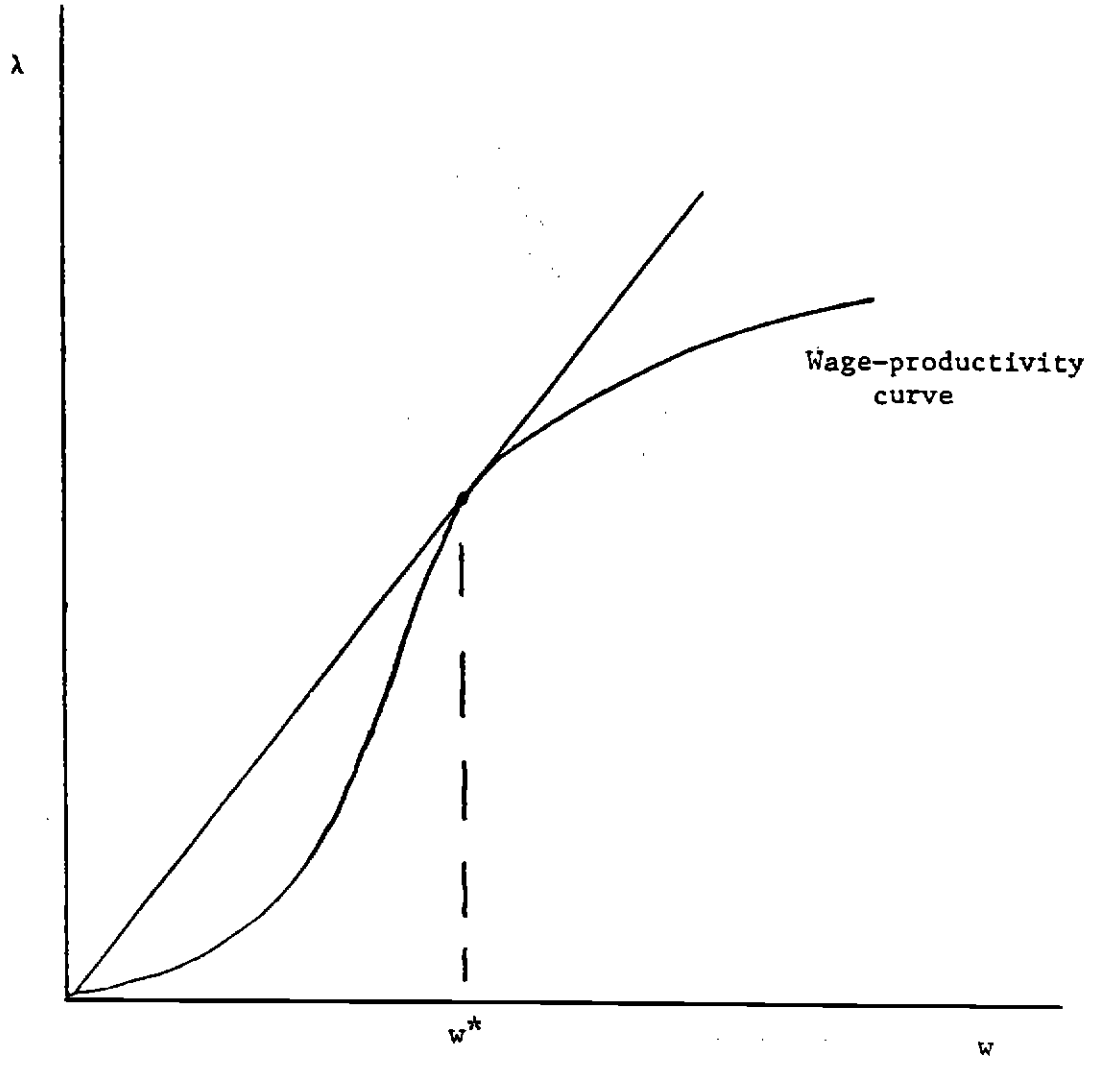


Figure 6e

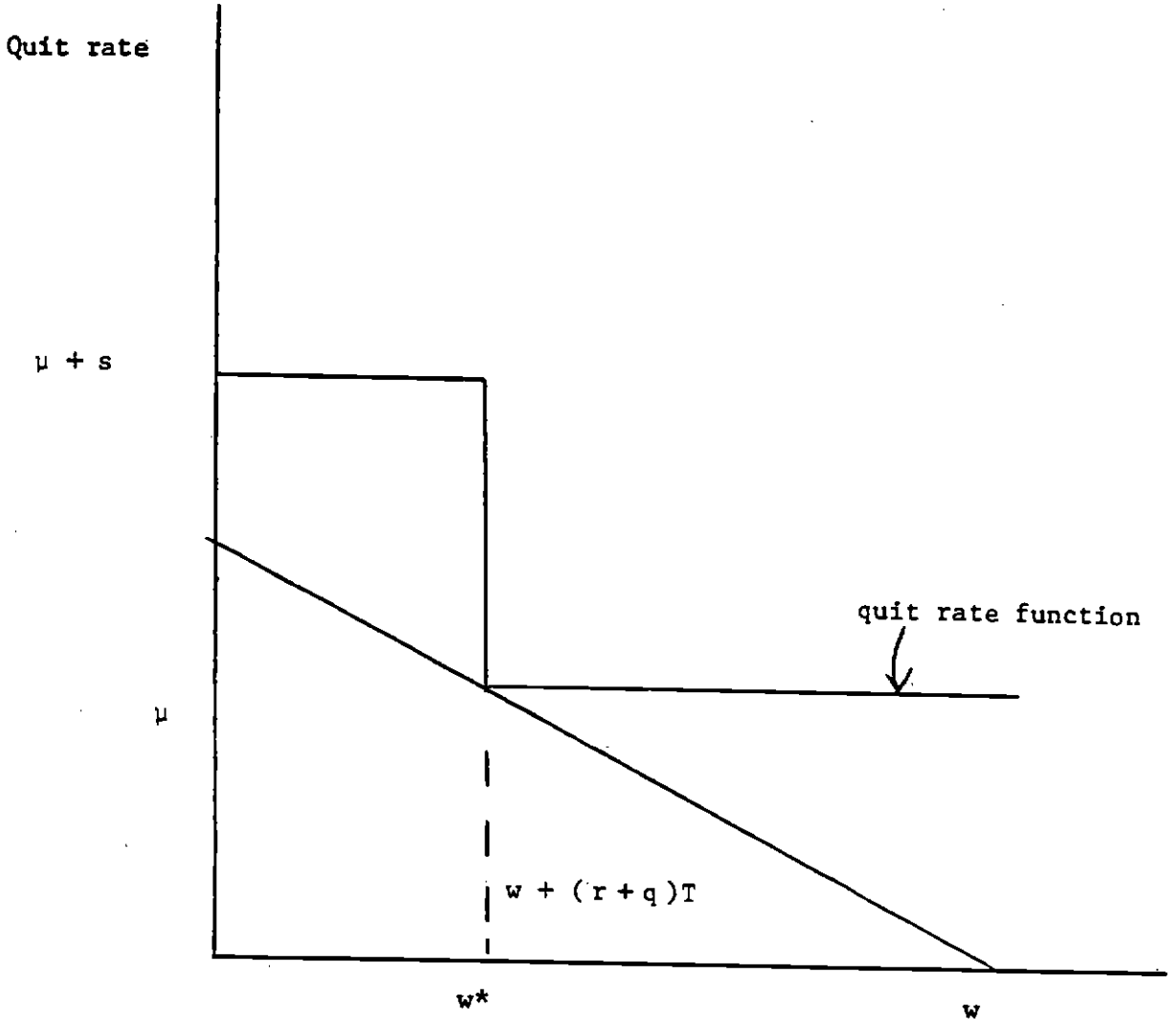
If individuals have utility functions of the form U = U(y-v(h)) and firms are risk neutral, then the first best equilibrium always satisfies the self-selection constraints.

FIGURE 7



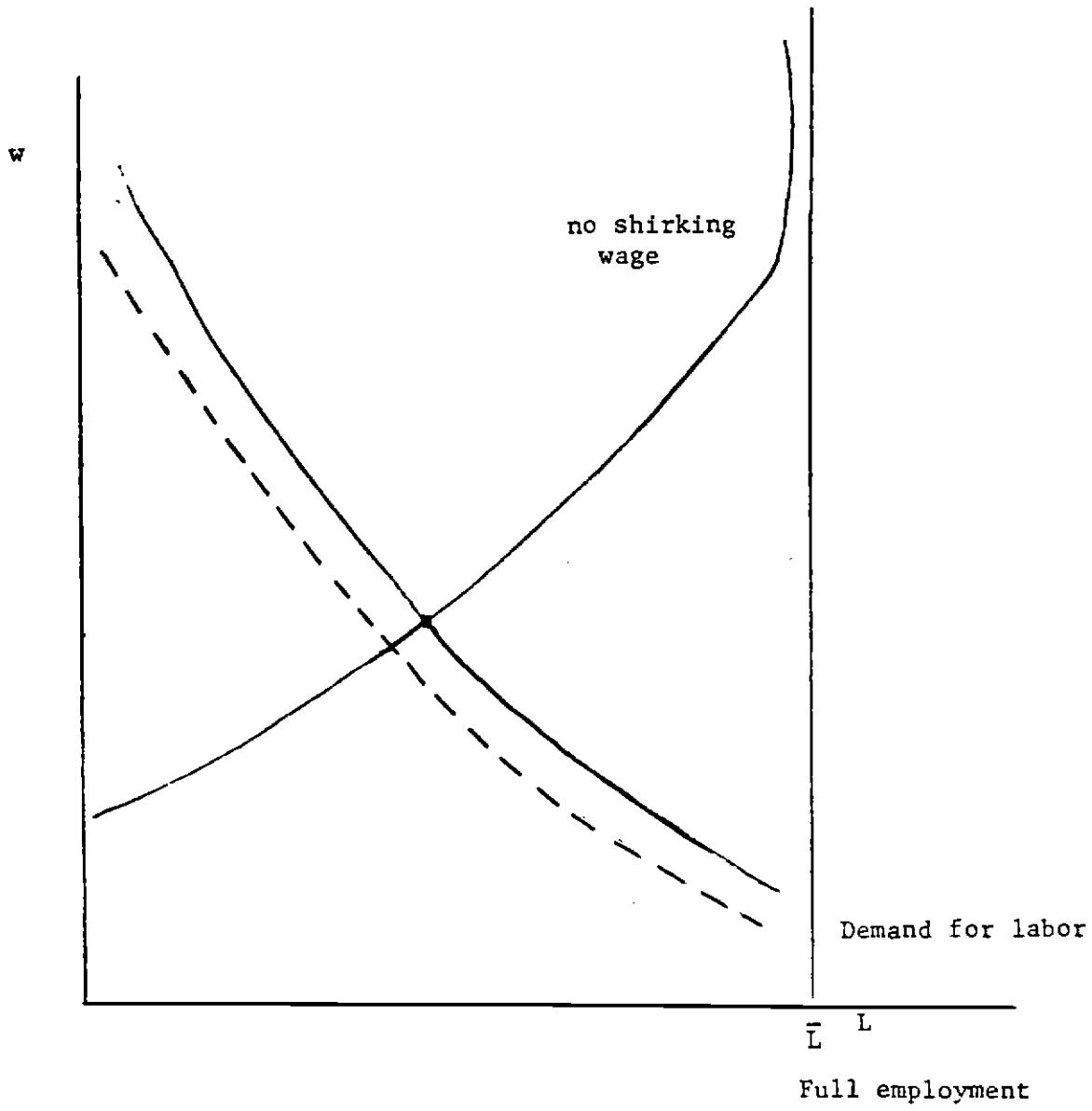
w* is the efficiency wage. At w*, wage costs per-efficiency unit are minimized.

FIGURE 8



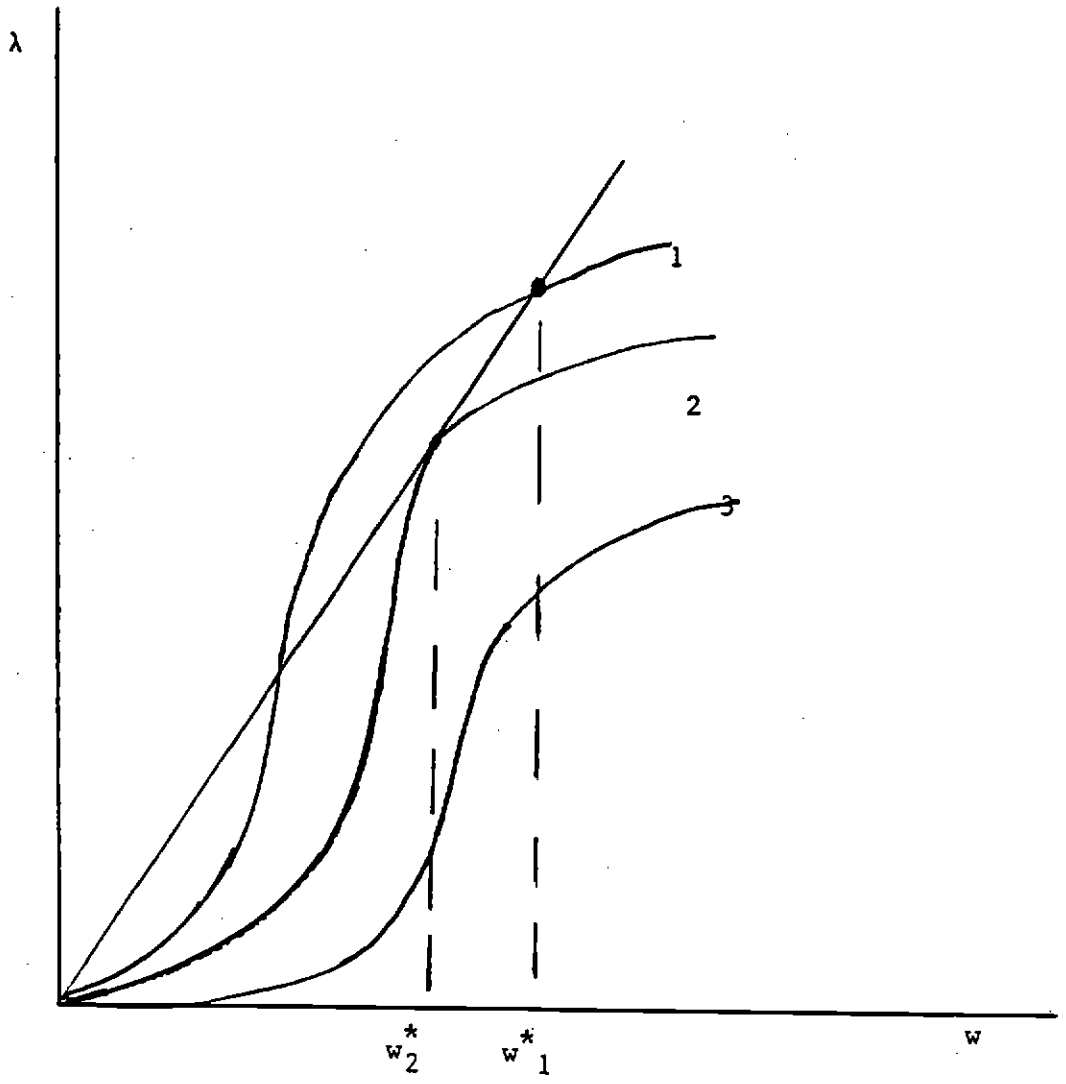
w* is the efficiency wage. At w*, total labor costs (including turnover costs) are minimized.

FIGURE 9



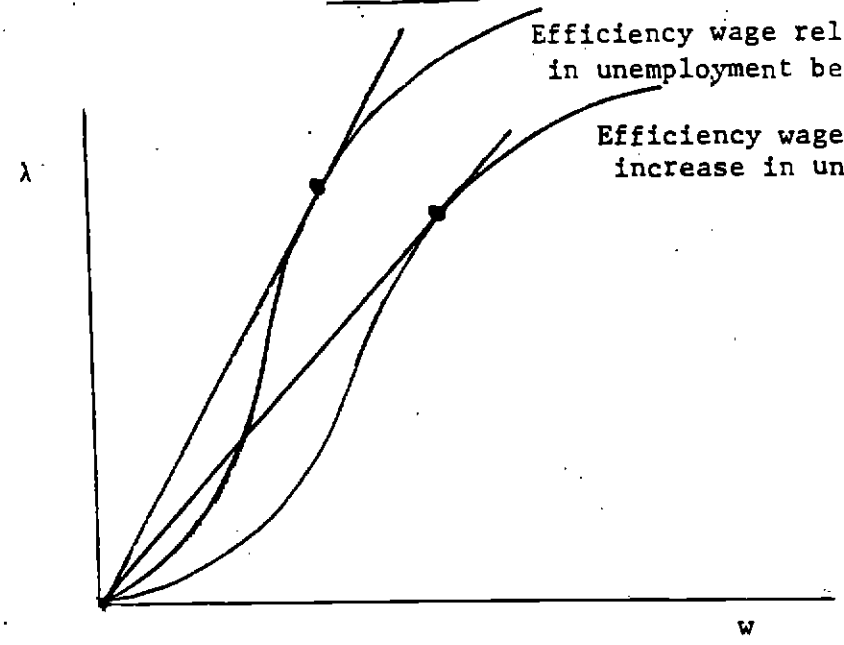
If monitoring is costly, the only equilibria entail unemployment. A decrease in the demand for labor will result in lower real wages and more unemployment.

FIGURE 10



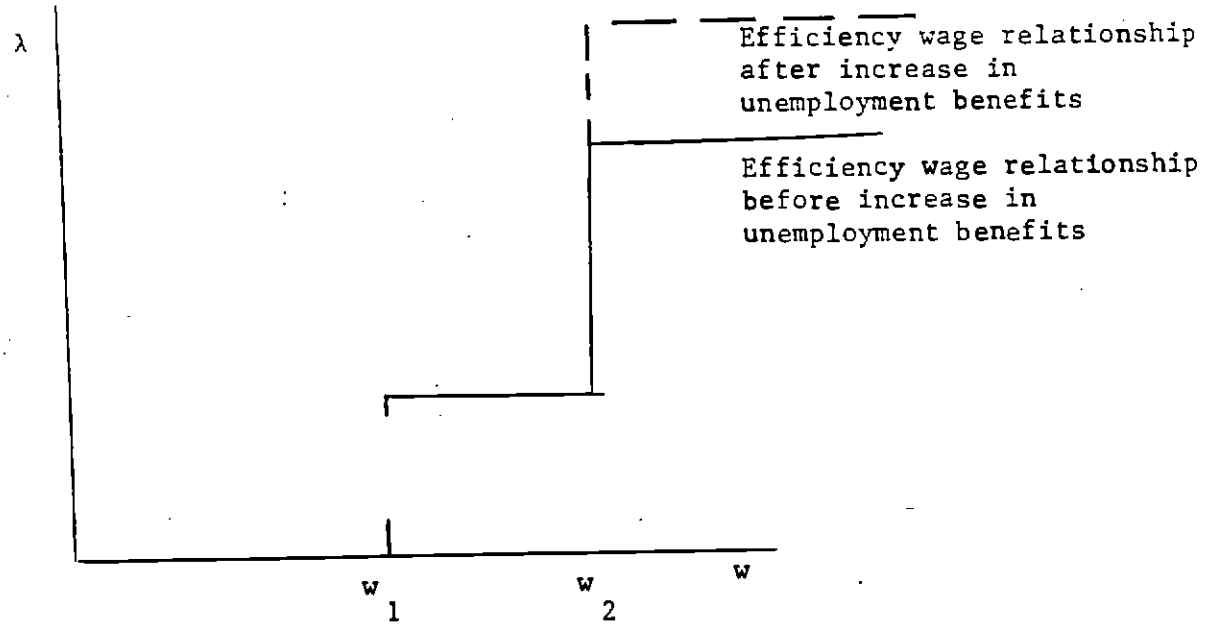
Consequences of differing wage productivity functions: some groups will be fully employed (group 1), some partially employed (group 2) and some completely unemployed (group 3).

FIGURE 11



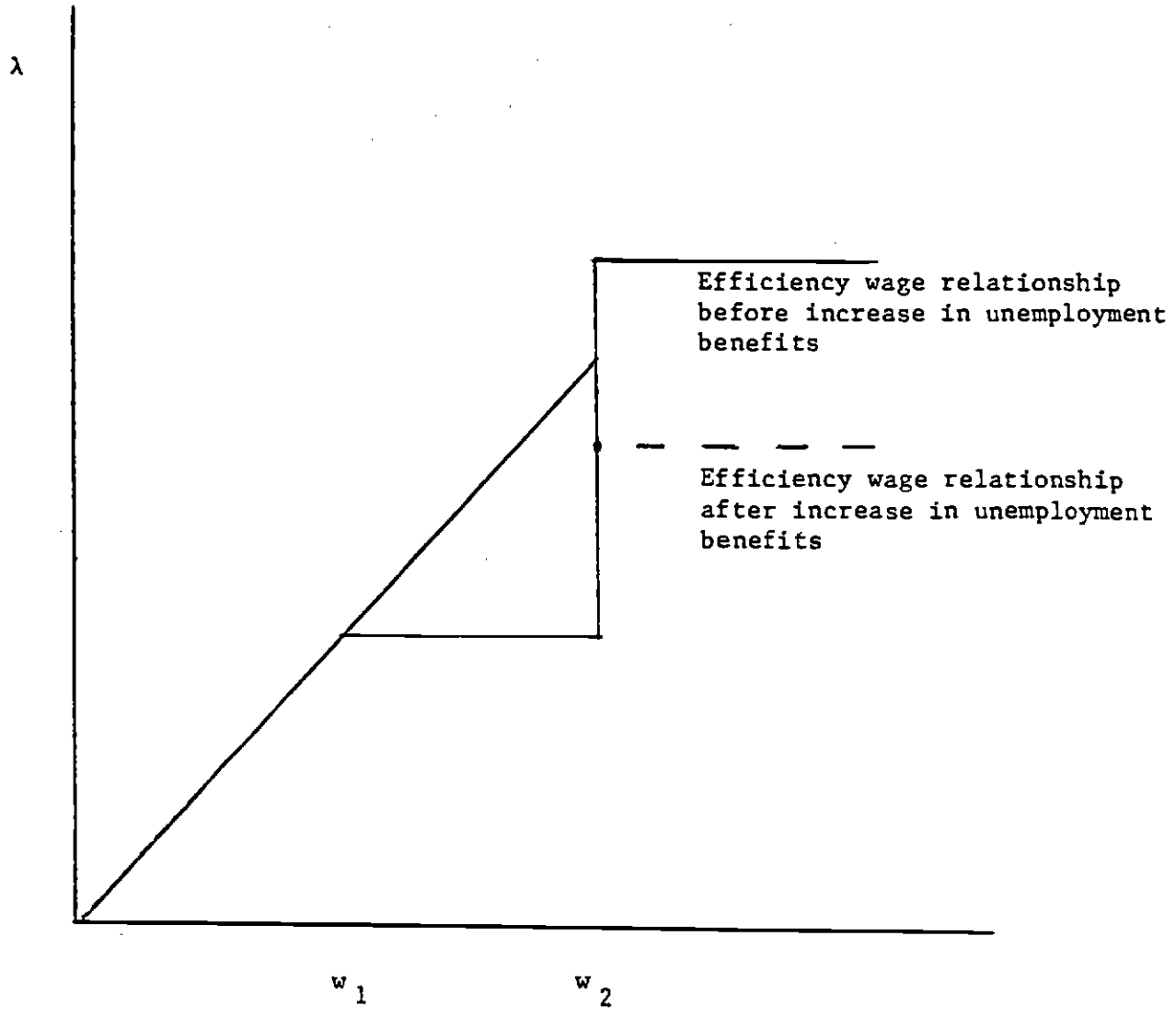
An increase in unemployment compensation may change wages and the demand for labor.

FIGURE 12



Two group case: an increase in unemployment compensation leaves wages unchanged but reduces unemployment.

FIGURE 12b



Two group case: an increase in unemployment compensation lowers the wage and reduces unemployment.