Discussion Paper No. 442

CONTRACTURAL MODELS OF THE LABOR MARKET

by

Bengt Holmstrom *

September 1980

* This paper was presented at the ASSA meetings in Denver, September 5-7, 1980 under the title "Long-term Contracts and Labor Market Equilibrium." Financial support from the Center for Advanced Study in Managerial Economics and Decision Sciences is gratefully acknowledged. I would like to thank Dale Mortensen and Edward Prescott for comments on an earlier manuscript.

In this paper I will discuss some approaches to modelling labor markets as contractually mediated. The central thesis of such models is that due to market imperfections - absence of contingent claims for labor and income - wage mediated auction markets will in general not be sustainable. There will be an opportunity for firms and workers to make a joint long-term contract which improves the welfare of both, and causes the auction market to collapse. This view, first suggested by implicit contract theory (Azariadis, Baily), opens up quite new perspectives on the operation of labor markets. Most importantly, one finds that in the contractual paradigm wage and marginal product may differ so that what appears to be disequilibrium in the short run may be a consistent equilibrium in the long run. Whether or not this can account for involunary unemployment I will comment upon later. It relates to my main concern, which is with the extent to which complex contingent contracts can be enforced, and the implications this has on the nature of equilibrium.

I will start by surveying the main results of implicit contract theory in the light of a rather general model of contractual equilibrium (Section I). Weaknesses of implicit contract theory will lead us to question the assumption of enforceable state-contingent contracts. Section II looks at the alternative of non-contingent fix-wage contracts whereas Section III presents a simple analysis of reputation as a means of enforcing more complex contingent contracts.

I. Implicit Contracts

Since it has been widely believed that implicit contracts can emerge only in markets where it is costly to move, let me start by showing that this is an artifact of the commonly used one period model. Assume initially that the labor market is cleared through sequential wage auctions and consider for

simplicity two periods only. Current wage is w_0 and next period wage w_1 is uncertain with known distribution $G(w_1)$. A worker's expected utility of participating in the market is assumed to be $u(w_0) + \int u(w_1)dG(w_1)$, where $u(.)$ is an atemporal risk averse utility function. Firms are assumed risk neutral. The expected wage bill for one worker is $w_0 + \int w_1 dG(w_1)$.

The claim is that firms can depart from the auction outcome to the benefit of both parties. To show that, consider a contract which pays the worker w' in the first period and guarantees w' in the second period as well. If market wage in the second period exceeds w', the firm has to follow suit or else the worker quits so the contract is called off. Choose w' so that

$$(1) \qquad u(w_0) + \int_0^{w'} u(w_1)dG(w_1) = u(w')(1 + \int_0^{w'} dG(w_1)),$$

i.e. so that the contract offers the worker the same prospect as the auction market. Dividing (1) by $(1 + \int_0^{w'} dG(w_1))$ and applying Jensen's inequality, gives

$$(2) \qquad w_0 + \int_0^{w'} w_1 dG(w_1) > w'(1 + \int_0^{w'} dG(w_1)),$$

establishing the claim.

Note that the role of two periods is to allow the firm to collect a premium in the first period ($w_o - w' > 0$) for the insurance it provides in the second period. The suggested contract looks like an option and (2) indicates that the selling price includes a risk premium which makes the contract favorable to the firm. However, the argument can be amended for risk averse firms. Also, identifying the worker's utility function is inessential. What

is essential though is that beliefs are not too dispersed and most importantly, that the firm will not lay off or exchange the worker when $w_1 < w'$. This last point I will return to.

Once all firms recognize the value of long-term contracts, wage auctions get replaced by contractual auctions in which the market will be equilibrated through the expected utility contracts offer. Since the market still is incomplete there will be a need to reopen it each period. What is envisioned to happen in future markets will influence current contract design and vice versa so the natural notion of equilibrium is one based on rational expectations. One can show such an equilibrium to exist. It is closely related to the pioneering models of Radner and Hart, with the distinction that here securities are created endogenously by firms and by assumption can be bought only through labor attachment at the respective firm.

Let me now turn to some properties of an equilibrium in implicit contracts, basing my discussion on an explicit treatment of a two-period version of the model sketched above (Holmstrom). With risk neutral firms and homogeneous labor there will be a unique optimal state-contingent contract that firms offer to their workers. If firms differ in their risks, contracts will differ and different firms may therefore pay different wages. In general contracts will involve layoffs in bad states and if the market turns favorable some workers may quit. Wage-wise an optimal contract will look like the earlier described option. As long as market forces do not push up equilibrium expected utility levels, retained workers will enjoy a constant wage, but with increased labor demand, wages will rise (due to legal constraints on involuntary servitude). It should be stressed that the implied downward rigidity of wages relate to individual contracts rather than aggregate wage levels. There is no presumption that new generations will receive the same

wage as old ones - in general they will not - and a laid off worker who has to find a job elsewhere is normally forced to take a wage cut. Thus, optimal contracting creates endogenously seniority classes within an otherwise homogenous labor force. Therefore, in aggregate the wage level will be flexible both up and down, but at a more sluggish pace than wage auctions would imply. This sluggishness is further increased by rights to be recalled (at previous wage) before any new workers are employed;a provision which will be part of an optimal contract.

Regarding the employment part of contracts I note that the model accomodates both quits and layoffs. Layoffs will occur for the same reason as in Azariadis' original treatment, namely the outside opportunity a worker has. But somewhat more acceptably this outside opportunity could be another firm rather than an exogenous benefit (e.g. household income). Essential for understanding the determination of employment is the fact that since the implicit contract is ex ante efficient and state-contingent it is also ex post efficient (as a necessary condition for ex ante efficiency). It follows immediately that a worker who can produce more within the firm than outside, will not be laid off. A more specific condition is obtained by equating the marginal rates of substitution between wage and employment probability (defined as the percentage of workers retained), which gives:

$$(3) \qquad pf'(nr) = w' - \frac{u(w') - u(w_1)}{u'(w')} ,$$

where p = output price (random), f(.) = production function, n = labor pool, r = proportion retained, and the other variables have been defined earlier. The right hand side is decreasing in w' and achieves its maximum for $w' = w_1$ (since $w' > w_1$ is required or else the worker quits). Thus, we find that

actually labor will be retained not only beyond the point where marginal product equals contract wage but market wage! Two important conclusions follow: wage and marginal product generally differ and a divergence of the two does not signal disequilibrium (which casts some doubt on recent fix-price modelling); and, secondly, there will be less rather than more unemployment in the contractual model compared to wage auctions.

The last conclusion has been viewed as a failure of implicit contract theory to explain involuntary unemployment, and if one defines involuntary unemployment as unexploited opportunities to trade, indeed it is. However, it is not clear to what extent opportunities are left unexploited in the real economy when one looks at the labor market in isolation; what we observe taking place could be consistent with (3), with inefficiencies being due to faults in the coordination between product and labor markets instead.

In order to alter the conclusion about the level of unemployement it is necessary to bring in some elements of incomplete or asymmetric information (otherwise contracts will be ex post optimal). More importantly, such modificatons will help to patch a logical inconsistency of implicit contract theory, namely the lack of severance payments with fully insured income. One simple change in informational assumptions is that w_1 cannot be observed by the firm and it will instead have to act on the conditional expectation $E(w_1|p)$. Then it is clear that even if it wanted to, the firm could not guarantee the worker a constant income and in some states (recessions) the level of unemployment would fall below that of an auction market (the rule would be as (3) with $Eu(w_1|p)$ replacing $u(w_1)$). Another change, explaining incomplete severance payments, would take note of the fact that laid off workers would have no incentive to search if their income was fully insured. With search, layoffs would result in some unemployment rather than pure

transfers as in the simple model.

The third change, which I will pay some more attention to, calls into question the firms assumed honesty. Generally, it is hard (for a single worker at least) to observe the marginal product of the firm and hence see if the required rule (3) is being followed. It appears therefore tempting for the firm to deviate from this behavior and insofar that this can be expected to happen, implicit contracts are rendered infeasible. Though I will indicate in the last section how a concern for reputation in the labor market may induce the firm to behave as if an implicit contract was written and honored, the doubt about enforceability of contracts prompts us to look at alternative contracts, which are not contingent on states that cannot be observed.

II. Fix-Price Contracts

The simplest non-state contingent contract is one in which the firm guarantees the worker a nominal wage, but adjusts employment at its own discretion. Let me first analyze whether such contracts can be expected to arise endogenously as a mutually beneficial arrangement compared to the wage auction outcome. If so, we would again have an explanation for the break down of wage auctions and have a legitimate reason to study fix-wage contracts further. I look at a simple model.

Consider an industry in which there is a large number of identical risk neutral firms characterized by the production function $f(n)$, $f'(1) = 1$. The output price fluctuates randomly between 1 and $p\varepsilon(0,1)$, which both occur with equal frequency. Workers supply their unit of labor inelastically and have an atemporal utility function $u(w)$. The number of workers per firm (appropriately scaled) is $n=1$. In a wage auction therefore, wage will equal output price (1 or p). Expected utility from such a market is

$Eu = \frac{1}{2}(u(1) + u(p))$, and expected profit $E\pi = (1 + p)(f(1) - 1)$.

Suppose now a firm would offer its workers a fixed wage w and lay off the unprofitable ones when output price is p. Letting r be the number (and proportion) retained, we have

$$pf'(r) = w \tag{4}$$

In order for this contract to be favorable (compared to the auction outcome) both to the worker and the firm we must have:

$$\bar{Eu} \equiv \frac{1}{2}(1 + r)\, u(w) + \frac{1}{2}(1 - r)\, u(p) \geq Eu, \tag{5}$$

$$\bar{E\pi} \equiv \frac{1}{2}(f(1) - w) + \frac{1}{2}(pf(r) - wr) \geq E\pi. \tag{6}$$

In (5) I use the assumption that the firm is small so a laid off worker can take a job in the spot market at wage p. Rewriting (5) and (6) (using (4)) gives:

$$r(u(w) - u(p)) \geq u(1) - u(w), \tag{7}$$

$$(1 + p)/p - f'(r)(1 + r) \geq f(1) - f(r). \tag{8}$$

It is easy to see that (4), (7) and (8) can be simultaneously satisfied; for instance, if $w \leq (1 + p)/2$, then (4) implies that (6) (hence (8)) is satisfied for all f and we can take f and u so that (7) holds. To understand what factors determine when (4)-(6) will hold, we can look at comparative statics. Assume that the equations above hold. Then they will remain intact

either when the absolute risk aversion of u is increased or if f becomes more kinked in the sense that $f'(r)$ and $f(1) - f(r)$ decrease (f becomes flatter between r and 1 and steeper between 0 and r). These conditions accord with intuition, since it is easy to understand (and check from (4)-(6)) that when the worker is risk neutral or when the production function exhibits constant returns to scale, there can be no gains to a fix-wage contract. Regarding changes in p, the effect is ambiguous. Note, however, that if there is a firm for which productivity does not fluctuate at all, it will always pay this firm to offer a fix-wage contract.

The assumption that the market is frictionless is rather unfavorable for a fix-wage scheme. Robert Gordon was first to suggest that a fix-wage scheme may be a response to the temptation a firm would have to lower the wage (claiming low marginal product) at will. Such behavior is, of course, limited by the worker's outside opportunity, but with costly labor mobility the firm's opportunity to exploit workers would quickly become rather substantial. My intention with the model above was to show that even without such transaction costs one can make a case for a fix-wage arrangement, and to indicate what factors will influence the benefits thereof.

In general, fix-wage arrangements will imply unemployment in excess of voluntary levels. Yet it may be an outcome of equilibrium and moreover an efficient arrangement subject to informational constraints, since as recent work on efficiency under asymmetric information emphasizes, it is false to apply standard ex post efficiency measures in the type of situation described.

The incomplete information paradigm has some apparent advantages over implicit contract theory. Nominal rather than real wages are rigid and this is not dependent on the unrealistic assumption of firm risk neutrality. With risk neutrality one is left wondering why a firm does not pay severance to

laid off workers and thereby insure income (the model above begs the same question; a justification would require a state in which severance is infeasible). Indeed, the role of unemployment benefits is quite unclear - why do firms not provide those privately? When wages are fixed due to incentive considerations, firms can be assumed risk averse, explaining incomplete severance and the incentive to pool risks through jointly paid unemployment benefits. One may also expect that optimal levels of such benefits, the degree to which they should be experience rated, and any additional severance payments could be determined. These questions appear fruitful for future research.

However, one should keep in mind that if there would be large gains to state contingent contracts, we would expect to see proxies for those states enter into contracts (a host of public data would be available for that) or observe direct monitoring of requisite states. But we do not, and before we understand why, enthusiasm over the imcomplete inforamtion paradigm, in particular its implication for involuntary unemployment, should be controlled.

III. Reputation.

I turn to the final point: can a concern for reputation lead the firm to act as if it honored an implicit contract? As we know from the growing literature on incentives in agencies(and more generally from the theory of repeated games), multi-period considerations will allow a rather richer set of opportunities to combat averse effects of informational asymmetries. To illustrate how, I will again use a simple example.

Consider a firm operating in an economy that can be in one of two states s=0 or s=1. The firm's revenue function is $(1 - \alpha s) f(n)$, where n is labor input. A suggested interpretation is that s=1 corresponds to a recession and the parameter α indicates how sensitive the firm is to the recession. The

sensitivity parameter stays fixed over time. It is initially unknown to workers, and this creates the opportunity for the firm to build a reputation by signalling the value of α through layoff behavior.

Workers work only in one period for which they sign a fix-wage contract as in the previous section. The wage demand will depend on their assessment of the probability of layoff, labelled as the firms reputation $\rho\varepsilon(0,1)$. In order to attract workers the firm has to pay $w(\rho)$, which is defined from the expected utility expression:

$$u(w(\rho))p + u(\bar{w})(1-p)(1-\rho) = v, \tag{9}$$

where $\bar{w}$ is income as unemployed, $p = Pr(s=o)$, and v is the expected utility offered by the market (assumed constant over time). Reputation ρ is a function of how the firm behaved in the previous recession. If a percentage r was laid off, this will be interpreted as indicating that the firm's type is $\alpha(r)$ (a signalling function to be determined in equilibrium), which in turn will give a prediction for how the firm will behave in the next recession $\rho(\alpha)$. Since α does not change over time in the simple version presented here, equilibrium will have $\rho(\alpha(r)) = r$, that is, the firm will lay off the same amount in each recession and the predictions obtained by assuming that the firm will repeat its behavior will become self-fullfilling in equilibrium. Therefore, I take reputation formation to progress as: $\rho_t = \rho_{t-1}$, if $s_t = 0$, $\rho_t = r_t$ if $s_t = 1$, where t is a time index.

Let the firm's discount factor be δ , and $V(\rho)$ be the optimal discounted expected profit function given current reputation ρ. Then

$$V(\rho) = \max_{n,r} \{(f(n) - w(\rho)n)p + ((1-\alpha)f(nr) - w(\rho)nr)(1-p) - \delta V(\rho)p + \delta V(r)(1-p)\}. \tag{10}$$

Here the firm is viewed as deciding the proportion retained r in case of a recession at the time it hires labor. Note that $\pi(n,w,r) \equiv (f(n) - wn)p + ((1-\alpha)f(nr) - wnr)(1-p)$, is the one period expected profit and that an implicit contract that is optimal solves $\max_{n,r} \pi(n,w(r),r)$.

In a stationary state we want $r = \rho^*$ to solve (10), where ρ^* is the stationary retention value. Thus,

$$(11) \qquad V(\rho^*) = \frac{1}{1-\delta}\,\pi(n^*,w(\rho^*),\rho^*),$$

where $n^* = \operatorname{argmax}_n \pi(n, w(\rho^*),\rho^*)$. For ρ^* we get a condition by noting that it must not benefit the firm to move to any other stationary layoff policy. Thus,

$$(12) \qquad \rho^* = \operatorname{argmax}_r \left\{(1-\alpha)f(n^*r) - w(\rho^*)n^*r + \frac{\delta}{1-\delta}\,\pi(n^*,w(r),r)\right\}.$$

This is the main equation of interest. Myopic behavior (as assumed in the previous section) results if $\delta = 0$. Then layoffs will equate current marginal product to wage. But when $\delta > 0$, so that there is a concern for the future, reputation will force the firm to set layoffs below the myopically optimal value; ($\frac{d}{dr}\,\pi(n^*,w(r),r) > 0$ for myopic r). Thus, reputation will provide for increased employment insurance. On the other hand, unless $\delta=1$, ρ^* will not be as high as in an implicit contract (which just maximizes $\pi(n^*,w(r),r)$ over r), since there is a cost of signalling through the excess employment. The optimal level of retained workers ρ^* will therefore lie between the myopic solution and the implicit contract solution, and be closer to the latter the closer δ is to 1 (the case of no discounting).

The paradigm suggested above is, of course, extremely simple and stylized, but I take it to indicate that reputation may have the power to enforce implicit contracts (or nearly so) even when the worker cannot verify the state of nature.This deserves further study. One would hope to learn how sophisticated the contingencies can be in reputation contracts, what influence the frequency of events has on reputation formation and could wage variation (when desirable) be supported by reputation as well. Intuition suggests that only rather simple, regularly occuring contingencies can be included in reputation contracts and that wages cannot be varied in such contracts because the wage-part of a contract is a zero-sum game whereas the employment-part is not, but the art of modelling reputation is still too primitive to confirm this intuition.

IV. Concluding remarks.

Implicit contract theory has been influential in suggesting a contractual view of labor markets, which carries the promise of a much improved understanding of how labor markets operate. The theory itself is incomplete at some crucial points. The main question taken up here concerns the enforceability of state-contingent contracts. On one hand, if we abandon the assumption that finely state-contingent contracts can be enforced, it leads to a fix-wage model which at least casually looking displays more reasonable features than the implicit contracts. On the other hand, the previous section indicates that a concern for reputation leads to behavior which appears as if implicit contracts were enforced. A consolidate view is that wages will be nominally downward rigid largely due to enforcement problems (rather than risk sharing), whereas employment rules will, at least to some degree, reflect a concern for reputation and therefore come closer to what would be implied by

implicit contracts, perhaps close enough to explain why more complex contingent contracts that would be feasible to write are not written.

REFERENCES:

C. Azariadis, "Implicit Contracts and Underemployment Equilibria," J. Polit. Econ., 1975, 83, 1183-1202.

M.N. Baily, "Wages and Employment under Uncertain Demand," Rev. Econ. Stud., 1974, 41, 37-50.

R. Gordon, "Recent Developments in the Theory of Inflation and Unemployment," Journal of Monetary Economics, 1976, 2, 185-219.

O. Hart, "On the Optimality of Equilibrium when the Market Structure is Incomplete,"Journal of Economic Theory, 1975, 11, 418-443.

B. Holmstrom, "Equilibrium Long-Term Labor Contracts," DP No. 414, Center for Mathematical Studies in Economics and Management Science, Northwestern University 1979.

R. Radner,"Existence of Equilibrium of Plans, Prices and Price Expectations in a Sequence of Markets," Econometrica, 1972, 40, 283-303.