Bargaining versus efficiency wages in a dynamic labor market:

A synthesis

By

Jon Strand

Department of Economics

University of Oslo

Box 1095, Blindern, 0317 Oslo, Norway

Tel: (+47)22855147, fax: (+47)22855035

E-mail: jon.strand@econ.uio.no

August 1998

Abstract

We construct a model integrating the efficiency wage model of Shapiro-Stiglitz (1984) with the matching-bargaining models of Diamond, Mortensen and Pissarides (DMP). Firms and workers form pairwise matches, workers may shirk on the job, and the wage is set in an asymmetric Nash bargain over the surplus created by nonshirking. The wage is then always higher, and employment lower, than in both the corresponding Shapiro-Stiglitz and DMP models. When firms determine workers' efforts unilaterally, efforts are inefficiently low and lower than in both the Shapiro-Stigitz and DMP models. Bargaining over both wages and effort raises effort, possibly to the first-best level. The overall equilibrium allocation may then be more or less efficient than in the Shapiro-Stiglitz model, but always less efficient than in the DMP model.

1. Introduction

Current theoretical modelling of unemployment in dynamic labor markets involving nonunionized workers is dominated by two paradigms. The first is the efficiency wage-shirking model, following the seminal paper by Shapiro and Stiglitz (1984), whereby unemployment is required at equilibrium to ensure that workers put up a required effort, given that effort cannot be continuously monitored and firing shirking workers is used as a mechanism for enforcing effort. This model generally disregards any effects of possible market frictions. The second is the matching-bargaining model, building on important contributions by in particular Diamond (1982), Mortensen (1978, 1982) and Pissarides (1988, 1990) (hereafter DMP). Here effort enforcement problems are neglected, while instead market frictions or mobility costs prevent the instantaneous matching of workers searching for jobs with firms searching for workers, and each firm-worker match gives rise to a specific surplus to be divided between the two parties. Both models give rise to involuntary unemployment at equilibrium.

The current paper represents an attempt to reconcile and merge these two modelling approaches. We assume that there are many (identical) workers and firms, each firm has many jobs, and creating a new job opening involves a fixed (recruiting, hiring or training) cost H, incurred by the firm. Once the worker is employed in the job, the worker and the firm are assumed to bargain over a specific current surplus. There are three main differences from the DMP structure. First, we simplify by abstracting from frictions in the matching process, such that one type of agents (active jobs) is always perfectly matched. Secondly, we extend the DMP analysis by assuming that the firm cannot continuously monitor the output produced by each worker, but as in Holmstrom (1982), only observe the average output over all active jobs. Thirdly, we assume that worker effort, and thus productivity, is endogenously determined, and either set unilaterally by the firm or as part of a worker-firm bargain. The second and third modifications may give workers an incentive to shirk, i.e., not put up the required effort. In the same way as Shapiro and Stiglitz (1984) (but different from Holmstrom (1982)) we assume that the firm may monitor worker effort, at discrete and stochastic points of time. The firm can in principle punish each worker after shirking has been discovered, and we disregard group punishments of the type studied by Holmstrom. In our model "shirking" involves a minimum effort which leads to a positive surplus in the firm-worker relationship. Given free entry of jobs (and thus zero expected profits from opening one), a firm then will not fire a worker who is caught shirking. Consequently, and departing from Shapiro-Stiglitz, firings are not part of a subgame perfect strategy and cannot be used as an equilibrium mechanism for enforcing effort. Instead, we assume that workers and firms bargain over the specific surplus created by worker nonshirking. This bargaining solution is affected by the (credible and "inside") option of the worker to shirk, after the bargain has been struck.

In section 2 we assume asymmetric Nash bargaining over the wage only, with relative bargaining strengths β and 1-β to the worker and firm respectively. The required effort is then decided unilaterally by the firm. In this bargain the utility of the worker under the shirking option defines a minimum below which the worker's utility cannot fall, and the parties bargain over the surplus in excess of this level. In the (out-of-equilibrium) event of shirking, the two parties would be assumed to renegotiate the wage given contiuned future shirking.This solution is shown to be inefficient in two different ways. First, it yields a wage in excess of the level derived in the Shaprio-Stiglitz framework, and an inefficiently low level of employment when all firm-worker pairs in the economy behave in the way specified by our model. Secondly, the allocation <u>within each worker-firm match</u> is inefficient since the chosen work effort requirement is inefficiently low.

In section 3 we extend this model by assuming that the parties may bargain over both the wage and the required level of effort, with the same bargaining strengths as before, and with the outcomes of the bargaining solution in section 2 as the parties' security levels. Such a solution yields higher levels of both the wage and worker effort, than the solution studied in section 2. Employment is also higher, since firms' profits are higher for given parameters exogenous to each firm. The higher effort requirement however leads to greater incentives for workers to shirk. We must then distinguish between two separate cases, namely one where the wage-effort bargaining solution is unconstrained by such considerations, and one where it is incentive constrained. The unconstrained case applies when workers' bargaining strength is above some minimum level. We then show that equilibrium effort is always first best, while overall employment may or may not be higher. When the worker nonshirking constraint directly affects the chosen wage-effort combination, effort is generally suboptimal but above the level in the corresponding Shapiro-Stiglitz model.

Section 4 sums up main conclusions and points out some directions for future research.

2. Bargaining over wages only

2.1 Basic assumptions

Consider an economy with identical and infinitely lived workers. When unemployed each worker obtains a constant utility b (from leisure, unemployment benefits or home production). When employed he earns a constant wage w_1 (as long as he is not caught shirking), and puts up a subjective effort e_1. This effort includes the disutility associated with turning up for work, and the disutility of productive effort. Whether or not a worker shows up for work is perfectly observable. Once the worker has shown up, however, the worker's productive effort is not perfectly observable. We will assume that regardless of monitoring, a worker will always choose to put up a positive effort and thus contribute positively to the firm's output. This can be explained in at least two ways. First, a zero productive effort might be conspicuous and a minimum effort necessary in order for “shirking” not to be immediately detected. Secondly, a zero effort may not be optimal for the worker once he has turned up for work. It may be more boring, and less rewarding, to drink coffee, read the newspaper and chat with coworkers all day, than doing at least some minimum amount of required chores. We correspondingly assume that the worker’s effort is given by a strictly convex function e(x), where x is his output. e(x) has a minimum, call it e_0, for some positive $x = x_0$. Given that he “shirks”, the worker would then rationally select $e = e_0$. Define, for $e \geq e_0$, $x = f(e)$ as the inverse of e(x). Then $f'(e) > 0$ and $f''(e) < 0$. Moreover, f’(e) goes to infinity when e goes to e_0 from above. This implies that the firm will always require work effort e_1 in excess of e_0, from a “nonshirking” worker.

Assume that each firm employs one worker and incurs a hiring cost H upon employment.[1] Job matches break up at a constant and exogenous rate s. Workers’ efforts are monitored at rate q, meaning that the interval between two monitoring times is exponentially distributed with exogenous parameter q (endogenizing q as in e.g. Ordover and Shapiro (1984) would not change the analysis fundamentally as long as optimal q is positive).

We assume that the wage paid to the worker is determined in a bilateral bargain between the worker and the firm, over the surplus created by the worker’s above-minimum effort. The option of workers to reduce efforts below the agreed level without the firm immediately detecting this, puts workers in a strategically advantageous strategic position versus the firms, which has no similar strategy available.[2] We will in the following assume that the default option in the bargain

[1]In reality the firm has many jobs whose output may be measureable only jointly. Our formal analysis however focusses on the individual job. This is unproblematic as long as total firm output increases linearly in the number of jobs filled.

[2]Note that firms cannot immediately detect whether a worker disagreement strategy (shirking) is being played, while workers can immediately detect whether a firm disagreement strategy (failure of wage payments) is being played.

implies that workers shirk and firms make the agreed-upon wage payments to workers. This bargaining solution is established on "day one" of the worker's tenure with the firm. Since the conditions facing the worker and the firm are stationary through time, the wage is at a given level w, as long as the worker is presumed by the firm not to shirk and is not caught shirking.[3] The relevant net surplus is created by the effort put up by the worker beyond the minimum level e_0. Assume that the worker and firm have bargaining strengths β and $1-\beta$, respectively. To derive the bargaining solution, we must first derive the payoffs to the two parties in the "default option", where the worker actually shirks. The firm's output is then $f(e_0)$. Given free entry among firms and identical jobs, the value to the firm of posting a job vacancy is zero. Given that the worker creates a net positive match value while "shirking", the firm has no ex post incentive to fire a worker caught shirking.[4] From this follows the logical conclusion that a threat to fire workers cannot be used to enforce worker effort.

In the analysis to follow in this section we start by deriving the renegotiated bargaining solution for a shirking worker, in subsection 2.2. The bargaining solution for a nonshirking worker is derived in subsection 2.3, market equilibrium in subsection 2.4, and the constrained optimal effort requirement in subsection 2.5.

2.2 The renegotiated bargaining solution

---

[3]This corresponds to the assumptions made by Shapiro and Stiglitz (1984). Alternative formulations of non-shirking contracts may imply that workers are required to pay an entrance fee when taking up employment with the firm, or be offered a lower initial wage. Such schemes may under some circumstances be incentive compatible for firms (see e.g. MacLeod and Malcomson (1989)). Such alternatives would however imply that workers' bargaining power is reduced; see the final section for a further discussion.

[4]As will be seen, whenever the value of output when shirking is less than the value to the worker of quitting and looking for a new job, the renegotiated bargain implies that worker will never quit (nor fired or forced to quit) even when the worker shirks, and is expected to shirk, in the continuation.

Define the discounted lifetime value to the worker of being in the unemployed state by U, the value of a negotiated contract for a nonshirking worker by E(N), and the value of a renegotiated contract for a shirking worker by E(R). Let r be the current rate of interest (common to workers and firms), s the exogenous rate of job loss, and h the rate at which unemployed workers find jobs (which is viewed as exogenous by each individual worker).[5] In steady state and with identical firms and workers (and thus bargaining solutions), the value of joining any firm will be E(N). We then have the asset equation rU = b + h[E(N) - U], which solves for U as

(1) $$U = \frac{1}{r+s}[b + h\ E(N)].$$

Given E(R) $\geq$ U, we have a similar asset equation r E(R) = w_0 - e_0 + s[U - E(R)], yielding

(2) $$E(R) = \frac{1}{r+s}(w_0 - e_0 + s\ U).$$

Define the value to the firm of a renegotiated contract with a worker who was caught shirking and continues to shirk, by V(R). Since the value to the firm of not having the worker is zero, we have the asset equation rV(R) = f(e_0) - w_0 -sV(R), implying

(3) $$V(R) = \frac{1}{r+s}[f(e_0) - w_0].$$

Assume an asymmetric Nash bargaining solution for w_0, over the joint net surplus E(R) - U + V(R), with relative bargaining powers β and 1-β to the worker and the firm. This implies the following expression for w_0:

(4) $$w_0 = \beta f(e_0) + (1-\beta)(e_0 + r\ U).$$

[5]Thus active job search is disregarded. Taking active search into consideration would change little in what follows below.

In this bargain U is taken as exogenous, determined in the market as a whole, and given by the (preliminary) expression (1). Inserting from (4) into (3) now also yields:

$$(5) \qquad V(R) = \frac{1}{r+s}(1-\beta)[f(e_0) - e_0 - r\ U].$$

A meaningful renegotiated bargaining solution requires V(R) > 0, i.e., $f(e_0) - e_0 > rU$. Thus net output when shirking must be greater than the current flow value of unemployment to a worker. If not there is nothing to gain by retaining a shirking worker who, as here, is expected to continue shirking in the future.

2.3 The initial wage bargaining solution

To derive the bargaining solution for a worker joining the firm, denote the wage in such a solution by w_1. The value E(N) of nonshirking is derived in analogous fashion to E(R) above,

$$(6) \qquad E(N) = \frac{1}{r+s}(w_1 - e_1 + s\ U),$$

where e_1 is the work effort level when the worker does not shirk, and where it is (rationally) assumed that the worker will never shirk in the future. From (1) and (6) we derive the following solutions for E(N) and U in terms of w_1:

$$(7) \qquad E(N) = \frac{1}{r(r+s+h)}[(r+h)(w_1-e_1) + s\ b]$$

$$(8) \qquad U = \frac{1}{r(r+s+h)}[h(w_1-e_1) + (r+s)b].$$

The value of shirking is given from the asset equation $rE(S) = w_1 - e_0 + s[U - E(S)] + q[E(R) - E(S)]$, where a worker is controlled at rate q, and that once shirking is discovered, the wage is renegotiated and the asset value E(R) enjoyed from then on. The solution for E(S) in terms of U and E(R) is given by

$$(9) \qquad E(S) = \frac{1}{r+s+q}[w_1 - e_0 + s\,U + q\,E(R)].$$

For the firm, similar values, of an initially engaged worker working and shirking, denoted V(N) and V(S) respectively, are derived as

$$(10) \qquad V(N) = \frac{1}{r+s}[f(e_1) - w_1]$$

$$(11) \qquad V(S) = \frac{1}{r+s+q}[f(e_0) - w_1 + q\,V(R)].$$

Asymmetric Nash bargaining over w_1, with relative bargaining weights β and 1-β to the worker and the firm, implies β[V(N)-V(S)] = (1-β)[E(N)-E(S)]. In deriving this solution, the parties take V(R), U and E(R) as exogenous.This leads to the following expression for w_1:

$$(12) \quad w_1 = \beta[\frac{r+s+q}{q}f(e_1) - \frac{r+s}{q}f(e_0) - (r+s)V(R)] + (1-\beta)[\frac{r+s+q}{q}e_1 - \frac{r+s}{q}e_0 - s\,U + (r+s)E(R)].$$

Using V(R), U and E(R), we find the following alternative expression:

$$(13) \quad w_1 = \frac{\beta(r+s+h)}{r+s+\beta h}[\frac{r+s+q}{q}f(e_1) - \frac{r+s}{q}f(e_0)] + \frac{(1-\beta)(r+s)}{r+s+\beta h}[b + \frac{r+s+h+q}{q}e_1 - \frac{r+s+h}{q}e_0].$$

(13) is derived under the assumption that w_1 is the same in all firms. It may be called a “reduced form” expression for the equilibrium wage w_1, as a function of parameters which are all

exogenous in the model, except for the rate of hiring of unemployed workers, h. (13) can be viewed as a generalization of both the efficiency wage w(S) derived in Shapiro and Stiglitz (1984), and the DMP bargaining solution. It is straightforward to see that w_1 = w(S) when β=0, i.e., workers have no bargaining power.[6] When $\beta > 0$, w_1 > w(S) for a given h. We also find that the DMP solution is obtained as $q \to \infty$ in (13).

Existence of a wage bargaining solution requires that we can find solutions where the firm's current profit, $\pi_1 = f(e_1) - w_1$, exceeds the flow cost of entry, (r+s)H. Current profit is given by

$$(14) \qquad \pi_1 = \frac{(1-\beta)(r+s)}{r+s+\beta h}[f(e_1)-b-e_1-\frac{r+s+h}{q}(e_1-e_0)] - \frac{\beta(r+s+h)}{r+s+\beta h}\frac{r+s}{q}[f(e_1)-f(e_0)].$$

When h is sufficiently large (and $\beta > 0$), $\pi_1 < 0$ and a solution cannot exist. The same holds (independent of h) when β is sufficiently large (close to one). Since h is endogenous and nonnegative, and π_1 falls in h, feasibility requires that π > (r+s)H for h = 0. This leads to the condition

$$(15) \qquad (1-\beta)[f(e_1)-b-e_1] - \frac{r+s}{q}[\beta(f(e_1)-f(e_0))+(1-\beta)(e_1-e_0)] > (r+s)H.$$

In the continuation we will assume that (15) holds. It puts an upper bound on H, and in addition e.g. implies a requirement that β cannot be close to one. Note also that as long as E(S) > U, the workers option to quit cannot have an influence on the bargaining solution. This can be shown to imply the following condition:

[6]Note that in Shapiro and Stiglitz's formulation, $e_0 = 0$; our formuation in this case can be viewed as a generalization to cases where $e_0 > 0$.

(16) $$w_1 > e_0 + b + \frac{(s+q)h}{(r+s)r}(e_1 - e_0).$$

(16) always holds when h is sufficiently low. Since w_1 however is bounded above by (13) when h increases, while the right-hand side of (16) is not, (16) can never hold when h is sufficiently great, and thus the unemployment rate is sufficiently low. The outside option of quitting must then always be greater than the "inside option" of shirking. One might then argue, in the same way as is customary in the DMP literature, that the quitting option constitutes the relevant threat point in the worker-firm bargain. In the following we will simply assume that (16) holds, and that the "inside option" consequently is relevant.

We may sum up this result in the following proposition:

Proposition 1: In the wage bargaining model with continuous renegotiation and identical firms and workers, relative bargaining strength β to workers, and (16) holds, the equilibrium wage is given by (13). An equilibrium solution with positive employment exists given that (15) holds.

The derived bargaining solution is based on an assumption that the security levels of the two parties are determined from a "disagreement" point at which worker shirk unilaterally while maintaining the relationship. This is fundamentally different from the approach taken in the DMP model, where the security levels of the parties rather are their outside options, determined as the parties' utilities in the event of the worker leaving the firm.

2.4 Effort requirement determination under wage bargaining

So far we have assumed that the worker effort required by the firm, e_1, is exogenous. With bargaining over the wage only, e_1 may instead be set unilaterally by the firm, and prior to the bargain over the wage. n setting e_1, the firm will take the rule for w_1, from (14), as given. Assume still that H is fixed. Then any given firm will set e_1 to maximize $\pi_1 = f(e_1) - w_1$, where w is given from (14). This leads to the following first-order condition for the firm:

$$f^{/}(e_1) = \frac{(1-\beta)(r+s+q)}{(1-\beta)q - \beta(r+s)}. \tag{17}$$

Here e_1 is a decreasing function of β. Thus a higher bargaining power to the worker leads the firm to require less effort. Intuitively, a greater worker bargaining power increases the wage for any given required effort level, and reduces the firm's marginal net return from requiring more effort. At equilibrium, the marginal gross return $f'(e_1)$ must then be raised, and e_1 reduced.

A condition for existence of an equilibrium effort solution in (17) is $(1-\beta)q - \beta(r+s) > 0$. Note also that as β increases and $(1-\beta)q - \beta(r+s)$ tends to zero, f' tends to infinity and e_1 tends to e_0.

The following result is now obvious:

Proposition 2: In the wage bargaining model, as described in proposition 1, effort is given by e_1 in (17), and is suboptimal.

This result and its implications will be elaborated on further in section 3 below. e_1 is suboptimal, since a first-best solution for worker effort must imply $e=e_B$ given by $f'(e_B) = 1$ (which e.g. is the effort that would be expended by a worker who owns the firm).

Note that in the Shapiro-Stiglitz model effort was viewed as exogenous by both firms and workers. Using our model above, however, the equilibrium effort under their assumptions in other respects is derived simply setting $\beta=0$ in (17). This yields $e = e^*$, given by

(18)
$$f'(e^*) = \frac{r + s + q}{q}.$$

Whenever $\beta>0$ and q is finite, this expression is generally lower than the corresponding expression in (17), but greater than the first-best level of one. Thus $e_B > e^* > e_1$.

(17) reveals that the basic condition for efficient effort to be chosen unilaterally by the firm is $q = \infty$; i.e., effort is monitored perfectly. Thus effort is in general inefficient in the context of the Shapiro-Stiglitz model, but efficient in the context of the DMP model. Intuitively, the threat point of the worker is affected by what effort is set in the former, but not in the latter, model. In the latter, the overriding concern for the firm, in setting the worker effort requirement, is maximizing the joint surplus of the two parties, which is then split in given fractions. This leads to a first-best allocation of effort.

2.5 Equilibrium employment and firm entry

We now wish to derive the equilibrium level of employment and the number of firms in the market as a whole, for the model of sections 2.1-2.4. We must then specify the process by which matching between workers and firms takes place. Two alternatives are particularly relevant:

1. The case of frictionless (or perfect) matching.Then <u>either</u> a firm searching for workers instantaneously finds a worker, <u>or</u> a worker searching for firms instantaneously finds a firm. Assume that firms always have to put up a fixed cost, H, upon hiring a worker, while workers bear no hiring costs. Assume first that workers are matched instantaneously. But this must imply $h = \infty$, which from (14) must imply that $\pi_1 < 0$. Thus instantaneous matching of workers is never compatible with equilibrium. This result is identical to the conclusion drawn by Shapiro and

Stiglitz (1984). At equilibrium, the rate of unemployment is $u = s/(s+h)$, which must be positive (and thus h finite) at an equilibrium solution where workers' efforts are enforced. Perfect matching thus implies that firms are matched instantaneously, and that workers must suffer some unemployment.

2. The case of match frictions (or imperfect matching) in the sense of Pissarides (1990). Denoting the number of vacant jobs as a fraction of the labor force by v, this approach implies that the rate at which vacant jobs are filled is given by $q(\theta)$, where θ denotes v/u, and the mean duration of a vacancy is $1/q(\theta)$. The function $q(\theta)$ has the property that $q'(\theta) < 0$, where q tends to zero (infinity) as θ tends to infinity (zero). Moreover, the rate of matching of workers, $\theta q(\theta)$, which corresponds to h above, tends to zero (infinity) as θ tends to zero (infinity), and the mean duration of unemployment is given by $1/\theta q(\theta)$. In such cases there will be some unemployment both among workers and active firms.

It can be argued that the imperfect-matching assumption is somewhat artificial and of limited interest in the context of a model of homogeneous labor such as ours. Firms will be exactly indifferent with respect to whom to hire, and as long as there is some (nonnegligible) unemployment, there should always be some worker available for immediate hiring. Rather, the imperfect-matching model should more realistically apply to a labor market where workers are heterogeneous, and where firms therefore are likely to spend time locating a suitable worker for the vacant position.

Our approach here is to emphasize the simpler, perfect-matching, case. Little is lost in terms of generality of results. We will discuss some implications of the more general case with frictions in the concluding section.

With perfect matching, searching firms are matched instantaneously. Given an exogenous hiring cost H, the zero-profit condition implied by free entry of jobs then leads to the condition $\pi/(r+s) = H$, which implies, using (14),

$$(19) \quad (1-\beta)[f(e_1)-e_1-b] \;-\; \beta\frac{r+s+h}{q}[f(e_1)-f(e_0)] \;-\; (1-\beta)\frac{r+s+h}{q}(e_1-e_0) \;=\; (r+s+\beta h)H.$$

Given exogenous e_i and H, (20) solves for the one endogenous variable h. In particular we find

$$(20) \qquad \frac{dh}{dH} = -\frac{q(r+s+\beta h)}{\beta[f(e_1)-f(e_0)]+(1-\beta)(e_1)-e_0)+\beta q H} < 0.$$

Thus h drops when the hiring cost increases. Intuitively, when H increases, equilibrium firm profits must increase, and thus the wage decrease. Then the equilibrium value of h must be reduced, since w is increasing in h. Denote equilibrium employment by L, and the total labor force by N, assumed exogenous. We have L = [h/(s+h)]N, which increases in h. Thus overall employment is reduced when firms' hiring costs increase. Considering a partial increase in the bargaining power of workers, represented by β, we find

$$(21) \qquad \frac{dh}{d\beta} = -\frac{q[f(e_1)-e_1-b] \;+\; (r+s+h)[(f(e_1)-f(e_0)) \;-\; (e_1-e_0)] \;+\; q\,h\,H}{\beta[f(e_1)-f(e_0)] \;+\; (1-\beta)(e_1-e_0) \;+\; \beta\,q\,H} < 0.$$

A greater bargaining strength to workers reduces h and thus employment. One implication is that since we have the equivalent of the Shapiro-Stiglitz model whenever β=0, employment must be lower here.

3. Bargaining over both wage and effort

3.1 Introduction

A basic assumption in section 2 was that only wages were subject to bargaining, while effort was decided unilaterally by the firm. Given worker-firm bargaining it may however seem artificial to restrict attention to wage bargaining only. Effort requirements must in any case be specified and agreed on by the firm and the worker in each individual relationship. This opens up for the natural issue of simultaneous bargaining over wages and effort. Since the firm's required effort, from (17), is inefficient in the firm-worker relationship, there is a potential for improving on the allocation, and making both the worker and the firm better off, through such simultaneous bargaining.

In general, a bargaining solution over effort and wages will lead to a level of effort exceeding the level in (17). This leads to greater incentives for workers to shirk on the job. Any bargaining solution derived in this section must correspondingly obey the worker non-shirking constraint $E(N) - E(S) \geq 0$, where E(N) and E(S) are appropriately modified. In subsection 3.2 we will assume that the derived (unconstrained) bargaining solution obeys this constraint. In subsection 3.3 below, we will instead assume that the wage-effort bargaining solution is constrained by the worker non-shirking constraint.

In the continuation, we denote the wage as given from (12) by w(1) (which will generally be different from w_1 above). Denote the equilibrium wage with simultaneous wage-effort bargaining, by w(2), and the corresponding equilibrium effort level by e_2. An important conceptual issue is what should be the appropriate surplus over which such bargaining takes place. In this section and the next, we will assume that the relevant surplus measure is the surplus <u>in excess of the utilities</u>

to the two parties, as given by the wage bargaining solution in section 2. The idea is that the firms could be assumed to propose the possibility of bargaining over effort and wages, instead of bargaining over only wages, under the threat of reverting to a solution with bargaining over only wages and unilateral firm determination of effort. I.e., during a conflict concerning the determination of effort, the firm is assumed to determine effort unilaterally and according to (17).

In deriving w(1) one must now take into consideration that the wage earned in alternative firms is w(2). This leads to the following general expression:

$$(22)\quad w(1) = \beta[\frac{r+s+q}{q}f(e_1) - \frac{r+s}{q}f(e_0)] + (1-\beta)[\frac{r+s+q}{q}e_1 - \frac{r+s}{q}e_0] + \frac{1-\beta}{r+s+h}[h(w(2)-e_2)+(r+s)b].$$

Denote the utilities for the firm, in the solution of section 2 and in the current solution, respectively, by V(N(1)) and V(N(2)), and the corresponding utilities for the worker by E(N(1)) and E(N(2)). Note that the levels of V(N(1)) and E(N(1)) differ from the levels V(N) and E(N) in section 2, since w(1) differs from w_1 in (12). The reason is that w(2) enters the expression for w(1) in (22), while in (12) not w(2) but w_1 itself enters. Thus (for given h) w(1) > w_1. The expressions for the utility gains are

$$(23)\quad V(N(2)) - V(N(1)) = \frac{1}{r+s}[f(e_2) - f(e_1) - (w(2) - w(1))]$$

$$(24)\quad E(N(2)) - E(N(1)) = \frac{1}{r+s}[w(2) - w(1) - (e_2 - e_1)].$$

Here e_1 is assumed to be given from (17).

3.2 Efficient wage-effort bargaining

An efficient bargaining solution with respect to w(2) and e_2, over this surplus with relative bargaining strengths β and 1-β, can (whenever it exists) be derived maximizing the generalized Nash surplus $(E(N(2))-E(N(1))^{\beta} (V(N(2))-V(N(1))^{1-\beta}$, with respect to w(2) and e_2. This yields the following two equations:

(25) $$\beta[V(N(2)) - V(N(1))] = (1-\beta)[E(N(2)) - E(N(1))]$$

(26) $$f'(e_2) = 1.$$

(26) implies the following:

Proposition 3: In an efficient wage-effort bargaining solution, given that it exists, worker effort is first-best optimal.

The rule (26) is simply the first-best rule for e_2. Thus bargaining over both the wage and effort in the firm-worker relationship implies an effort that is efficient, and higher than under unilateral firm effort determination. Intuitively, e_2 maximizes any joint individual worker-firm surplus. This is a robust result, independent of the bargaining strengths of the two parties, and of their security levels.

(25) together with (22) imply the following solution for w(2):

$$(27)\quad w(2) = \frac{\beta(r+s+h)}{r+s+\beta h}[f(e_2)+\frac{r+s}{q}(f(e_1)-f(e_0))] + \frac{(1-\beta)(r+s)}{r+s+\beta h}[b+e_2+\frac{r+s+h}{q}(e_1-e_0)].$$

Comparing (27) to (13), we find for given h and β:

$$(28)\quad d\,w \equiv w(2) - w_1 = \frac{\beta(r+s+h)}{r+s+\beta h}[f(e_2) - f(e_1)] + \frac{(1-\beta)(r+s)}{r+s+\beta h}(e_2 - e_1).$$

Moreover, since worker utilities u(2) and u(1) in these two cases are given by w(2)-e_2 and w_1 - e_1 respectively, we find, again for given β and h:

$$(29)\quad d\,u \equiv u(2) - u(1) = \frac{\beta(r+s+h)}{r+s+\beta h}[f(e_2) - e_2 - (f(e_1) - e_1)]$$

This expression is positive by virtue of e_2 being optimal. The corresponding expression for the firm's current profit is

$$(30)\quad \pi(2) = \frac{(1-\beta)(r+s)}{r+s+\beta h}[f(e_2)-b-e_2-\frac{r+s+h}{q}(e_1-e_0)] - \frac{\beta(r+s+h)}{r+s+\beta h}\frac{r+s}{q}[f(e_1)-f(e_0)].$$

Comparing (28) to (14), we find that excess profits in the current solution, as compared to the wage bargaining solution (for given β and h) are given by

$$(31)\quad d\pi \equiv \pi(2) - \pi_1 = \frac{(1-\beta)(r+s)}{r+s+\beta h}[f(e_2) - e_2 - (f(e_1) - e_1)].$$

Here dπ > 0, for the same reason as why du > 0. The results just derived can be summed up as follows:

Proposition 4: Assume β and h given, and an efficient wage-effort bargaining solution exists. Compare an economy where all firm-worker pairs bargain over both wages and effort, to another economy where there is bargaining over wages only. Then wages, workers' utilities, and firms' profits are all greater in the economy with wage-effort bargaining.

These results may seem intuitively straightforward. Since the wage-effort bargaining solution is clearly more efficient than the pure wage bargaining solution, there will be a greater surplus to share between the parties, and thus both firms and workers should gain when this surplus is shared. Note however that when all firm-worker pairs bargain, E(N(1)) is greater than the value of employment in the wage bargaining solution, E(N) given by (6). Thus when also other firms offer such contracts based on wage-effort bargaining, the security level of the worker, relevant for the current bargain, is raised. Proposition 4 shows that at equilibrium, this increase in worker bargaining power is never so great that not also firms makes a net gain.

The results in proposition 4 are derived for given β and h. But we know by the assumption of free entry among firms, that $\pi(2) = \pi_1 = H/(r+s)$ at equilibrium. Note that π(2) is decreasing in h. This must imply the following:

Proposition 5: Assume that an efficient wage-effort bargaining solution exists, and compare an economy where all firm-worker pairs bargain over both wages and effort, to an economy where all bargain over only wages, and where β is the same in both. Then employment is greater in the economy with efficient wage-effort bargaining.

This follows because h must be greater in the economy with wage-effort bargaining. Using (31), the rate of employment, given by h/(s+h), must then be greater.

A consequence of propositions 3 and 5 is then that (for a given β) the number of employed workers and output per employed worker both will be greater under efficient wage-effort bargaining than under wage bargaining only. Thus clearly, output will be greater, for two independent reasons. Moreover, the wage and worker utilities will be higher.

We have so far taken for granted that an efficient wage-effort bargaining solution exists. It is clear that whenever firms' profits are positive in the pure wage bargaining solution in section 2 above, profits must be positive here as well. Note however that the derived solution does not directly involve the worker non-shirking constraint. In general, since the worker effort requirement is greater now, workers' incentive to shirk is also greater. We may show the following:

Proposition 6: Assume that for a given β, a wage bargaining solution exists in section 2 above. Then an efficient wage-effort bargaining solution exists, provided that the following condition is fulfilled:

$$(32) \quad \beta[f(e_2) - e_2 - (f(e_0) - e_0)] \geq \frac{r+s}{q}[e_2 - e_1 - \beta(f(e_1) - e_1 - (f(e_0) - e_0))].$$

Proof: First, note that whenever a wage bargaining solution exists, firms must make profits in excess of (r+s)H at that solution, given h = 0. Since by proposition 4, profits are higher for a given h at an efficient wage-effort bargaining solution, this is the case here as well. Thus there must exist some finite h > 0 which yields profits equal to (r+s)H. Next, consider the incentives of workers to put up effort. From the non-shirking constraint E(N(2)) ≥ E(S(2)), (32) can be derived. Thus whenever (32) holds, workers will put up effort e_2 given an efficient wage-effort bargaining solution. Q.E.D.

(32) requires that the rent accruing to workers in a wage-effort bargaining solution be sufficiently great for shirking not to be preferred, at the (greater) level of effort required under the efficient bargaining solution. This requires that workers' bargaining strength be above some minimum level. In particular, since the right-hand side of (32) is always positive, this inequality cannot hold when β is close to zero.

We also see that whenever e_2 is sufficiently close to e_1, an efficient wage-effort bargaining solution must always exist whenever $\beta > 0$. Intuitively, the worker incentive compatibility constraint (represented by the efficiency wage) is then not much more strict than the corresponding constraint for $e = e_1$. Since $\beta > 0$ implies that w(1) is strictly above the efficiency wage corresponding to e_1, w(2) must be above the efficiency wage corresponding to e_2, for e_1 and e_2 sufficiently close. Put otherwise, the extra rent implied by the worker bargaining strength β is sufficient to ensure incentive compatibility.

3.3 Incentive-constrained wage-effort bargaining solutions

Consider now the situation where (32) does not hold and an efficient bargaining solution is not implementable. The non-shirking constraint given $e = e_2$ must then formally be introduced with equality as part of the constrained optimal solution. The form of this constraint which is relevant in a bargaining context is derived directly from the non-shirking constraint in this case, as

$$(33) \qquad w(2) \geq \frac{r+s+q}{q}e_2 - \frac{r+s}{q}e_0 - s\,U + (r+s)E(R).$$

(33) will be fulfilled with equality at a constrained optimal solution. Maximizing the generalized Nash product $[E(N(2))-E(N(1))]^{\beta}\,[V(N(2))-V(N(1))]^{1-\beta}$ with respect to e_2 under such a constraint yields the following solution for e_2 :

(34) $$f'(e_2) = 1 + \frac{r+s}{q}\left[1 - \frac{\beta}{1-\beta}\frac{V(N(2))-V(N(1))}{E(N(2))-E(N(1))}\right].$$

The exact solutions for V(N(2)) and E(N(2)) may now be obtained implicitly, inserting the solution value for e_2 from (34) into the respective expressions. Here the constraint (33) effectively forces w(2) up relative to the solution given in (27), and at the same time forces e_2 down. The last result follows because the expression inside the square bracket in (34) is generally between zero and one, at a constrained efficient bargaining solution. Denoting the last main term inside the square bracket in (34) by α, we have in the unconstrained case above that $\alpha = 1$, corresponding to the first-best effort level given by $f'(e_2) = 1$. In the current constrained case we must have $\alpha \in (0,1)$. Note that the fraction of the bargaining surplus that goes to workers now is given by $\beta/[\beta+\alpha(1-\beta)]$, which increases when α is reduced. The effective bargaining power of workers is thereby increased. We have the following conclusion:

Proposition 7: In a constrained efficient wage-effort bargaining solution where the worker non-shirking constraint binds, worker effort will be set intermediate between the first-best level, and the level in the (appropriately modified) Shapiro-Stiglitz model. The effective bargaining power of workers is then raised above the Nash bargaining parameter β.

A final interesting issue implies comparing the efficiency of an economy with wage-effort bargaining, to an economy characterized by pure efficiency wages, corresponding to the appropriately modified Shapiro-Stiglitz model, here, the case of pure wage bargaining and the limit solution as $\beta \rightarrow 0$. This comparison can be made for different levels of β. It is then clear that for sufficiently high β, the Shapiro-Stiglitz model must yield a more efficient solution. For lower levels of β, firms' profits for given h are still always lower in the current case, and thus equilibrium employment lower. But effort is more efficient here (regardless of whether the wage-effort bargaining solution is unconstrained or constrained), and output per worker greater. This implies a possibility that the overall equilibrium solution is more efficient under such a bargaining solution, than under efficiency wages (or "wage posting") a la Shapiro and Stiglitz.[7]

4. Conclusions and final comments

We have studied a model where pairwise worker-job matches are formed and each worker bargains with the firm over the resulting surplus. Workers may shirk on the job for some period without being discovered, and this affects the outcome of the bargaining solution since the option of shirking while earning the bargained wage defines the worker's threat point in the bargain over the wage only. Worker shirking implies some positive effort, and thus a positive surplus to be shared. A threat by the firm to fire a shirking worker is then not credible, and renegotiation is assumed to take place in such circumstances.

We first in section 2 considered asymmetric Nash wage bargaining. When effort requirements are set by the firm, they are inefficiently low, while wages (for given effort) are higher than in either the pure moral hazard or DMP model. Both employment and worker effort (decided by the firm) are lower, and as a consequence the overall allocation less efficient, than in either of the two "pure" models.

[7] Note that here, fewer workers will be employed but each will have a higher wage, relative to Shapiro and Stiglitz. Thus gains will be distributed less evenly across workers under wage-effort bargaining here.

In section 3 we assumed bargaining over both the wage and effort. When workers' bargaining power is "not too low", the resulting bargaining solution may be unconstrained, and it then yields a first-best level of match-specific effort. The bargaining solution may however be constrained by the worker nonshirking constraint when the effort requirement is raised. Effort will then not be fully efficient, but higher than under pure moral hazard. In general the overall allocation may be more efficient under wage-effort bargaining than under pure moral hazard, despite the wage being higher and employment lower.

Overall, the model yields a number of insights. First, for given work effort, we integrate the DMP wage bargaining model with the Shapiro-Stiglitz efficiency wage model, and demonstrate that each may arise as a special case of the more general model (the DMP model as q tends to infinity, and the Shapiro-Stiglitz model as β tends to zero). We show, not surprisingly, that there is a "compounding" effect of adding bargaining to the moral hazard framework, in the sense that the resulting wage is higher than in either of the two "pure" models. Secondly, we integrate the issue of effort determination into this combined wage bargaining-efficiency wage framework. We point out that unilateral firm effort determination is generally always efficient in the DMP model, but inefficient when (as here) workers' bargaining strength is affected by the effort requirement. There are then incentives to bargain over effort as well, given initial bargaining over the wage, thus making effort higher and more efficient. Thirdly, we show that bargaining may lead to efficiency improvements in an economy with worker shirking. Such efficiency gains are due to the more efficient effort levels that become achievable under bargaining. The overall allocation may be more efficient even though the wage is higher and employment lower than under the assumption of shirking only.[8] Fourthly, we argue that our model may provide a better theoretical

[8] A different attempt to integrate efficiency wage theory with the DMP model of search and matcing has recently been made by Mortensen and Pissarides (1998, section 4.6). They however focus on the implications of efficiency wages as the mechanism for determining wages, within a labor market with an imperfect matching

foundation for the Shapiro-Stiglitz model itself, in view of the extensive criticism of this approach in the literature. Some critics (e.g., Carmichael (1985)) claim that involuntary unemployment would disappear in the context of that model, if eachworker upon joining the firm could make a sufficiently great up-front payment, or be forced to accept a sufficiently low initial wage during the period immediately after being hired. Such alternatives however lower the utility of workers and constitute a reduction in their effective bargaining power, and will be effectively resisted when workers have positive bargaining power, as assumed here. This issue will be discussed further in section 4 below.

A number of issues remain to be fully resolved by our analysis. We have assumed that the solution to the worker-firm wage bargain always takes the form of a fixed wage, something that is not self-evident. It can be shown that more efficient equilibria are possible in the Shapiro-Stiglitz model, if the wage is allowed to be reduced during an initial period and later rise, in a manner similar to Lazear (1981).[9] One might here visualize similar bargaining solutions embedding profiles of time-dependent wages. Such solutions would however in general imply lower effective bargaining power to workers (for given β). An analysis of them must be left for future research.

In the paper we have assumed a perfect-matching approach, which we claim can be defended given homogeneous workers and jobs. Under the more traditional imperfect-matching approach taken in the DMP literature equilibrium implies that firms posting a vacant position do not immediately find an unemployed worker. Assume no hiring cost but instead a recruiting cost c per unit of time, for a firm that posts a vacancy which is not yet filled. Call the value of a filled job J. Using the notation of section 2.5 we have the equilibrium relationship

---

technology. Our focus is rather on the interaction of efficiency wages and bargaining in the wage determination process.

[9]See e.g. Strand (1991, 1992).

(35) $$J = \frac{c}{q(\theta)}.$$

Such a formulation implies the following three main modifications relative to our perfect-matching formulation. a) There will now be some vacancy among firms, since $1/q(\theta)$ represents the time it takes on average to fill a vacant position. This time can be compared to the average duration of a job, 1/s, and of unemployment, 1/h. b) An initial increase in firm profits as before increases the rate of worker matching h, but now also reduces the rate of job matching. The effects of an initial increase in firm profits on overall employment are thereby dampened somewhat. c) A "socially efficient" solution implies some idleness, for both workers and active jobs, while in the perfect-matching case, the efficient solution implies full employment for both. Except for these differences, most of the interesting results are retained.

The assumption of homogeneous agents on both sides of the market, in particular, homogeneous workers, can be questioned. Dropping this assumption is likely to have far-reaching consequences which are today not well understood. Heterogeneous labor has been introduced into the Shapiro-Stiglitz model e.g. by Strand (1987) and Albrecht and Vroman (1998), and into the bargaining-matching model by Ellingsen and Rosén (1997), Rioux (1996) and Strand (1997a,b). These are however more examples than general analyses, and further investigations of worker heterogeneity, in similar contexts, must be an important topic for future research.

A crucial final issue is of course the model's empirical relevance. Among other things, the model predicts the following: a) that the wage tends to be higher in employment relationships which are governed by a combination of bargaining and moral hazard, instead of just one of these phenomena; b) that worker productivities tend to be higher under "pure bargaining" than when the bargaining solution which is influenced by moral hazard considerations; and c) that employment relationships tend to be longer-lasting when there are match-specific bargaining surpluses than when there are not, even given a basic problem of worker shirking. A proper validation of the model would require that such questions be answered in the affirmative. Unfortunately, empirical evidence on such issues is hard to come by. Future research should thus be directed at clarifying such relationships, in particular by providing the relevant evidence.

References:

Albrecht, J. W. and Vroman, S. B. (1998), Nash equilibrium efficiency wage distributions. International Economic Review, 39, 183-203.

Carmichael, L. (1985), Can unemployment be voluntary? Comment. American Economic Review, 75, 1213-1214.

Diamond, P. (1982), Wage determination and efficiency in search equilibrium. Review of Economic Studies, 49, 217-227.

Ellingsen, T. And Rosén, Å. (1997), Fixed or flexible? Wage setting in search equilibrium. Working paper, Stockholm School of Economics.

Holmstrom, B. (1982), Moral hazard in teams. Bell Journal of Economics, 13, 324-340.

Lazear, E. P. (1981), Agency, earnings profiles, productivity, and hours restrictions. American Economic Review, 71, 606-620.

MacLeod, W. B. and Malcolmson, J. M. (1989), Implicit contracts, incentive compatibility, and involuntary unemployment. Econometrica, 57, 447-480.

Mortensen, D. T. (1978), Specific capital and labor turnover. Bell Journal of Economics, 9, 572-586.

Mortensen, D. T. (1982), The matching process as a noncooperative/bargaining game. In J. J. McCall (ed.): The economics of information and uncertainty. Chicago: Chicago University Press, 233-254.

Mortensen, D. T. and Pissarides, C. A. (1998), New developments in models of search in the labor market. Working paper, Northwestern University/London School of Economics.

Ordover, J. and Shapiro, C. (1984), Advances in supervision technology and economic welfare: A general equilibrium analysis. Journal of Public Economics, 25, 371-389.

Pissarides, C. A. (1988), The search equilibrium approach to fluctuations in employment. American Economic Review Papers and Proceedings, 78, 363,369.

Pissarides, C. A. (1990), Equilibrium unemployment theory. London: Basil Blackwell.

Rioux, L. (1996), Wage inequality as explained by endogenous labor market segmentation. Working paper, DELTA, Paris.

Shapiro, C. and Stiglitz, J. E. (1984), Equilibruim unemployment as a worker discipline device. American Economic Review, 74, 433-444.

Strand, J. (1987), Unemployment as a discipline device with heterogeneous labor. American Economic Review, 77, 489-493.

Strand, J. (1991), Unemployment and wages under worker moral hazard with firm-specific cycles. International Economic Review, 32, 601-612.

Strand, J. (1992), Employment and wages with deferred wage payments in models of worker moral hazard. Memorandum from Department of Economics, University of Oslo.

Strand, J. (1997a), Wage bargaining with heterogeneous labor and turnover costs: Perfect history screening. Mimeo, Department of Economics, University of Oslo.

Strand, J. (1997b), Wage bargaining with heterogeneous labor and turnover costs: The no-screening case. Mimeo, Department of Economics, University of Oslo.