

IZA DP No. 163

Why Do Firms Pay an Overtime Premium?

Robert A. Hart Yue Ma

June 2000

Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor

Why Do Firms Pay an Overtime Premium?

Robert A. Hart

University of Stirling and IZA

Yue Ma

University of Stirling

Discussion Paper No. 163 June 2000

IZA

P.O. Box 7240 53072 Bonn Germany

Phone: +49-228-3894-0 Fax: +49-228-3894-180 E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 163 June 2000

ABSTRACT

Why Do Firms Pay an Overtime Premium?*

This paper develops a two-period specific human capital model in which the bargaining parties seek to achieve optimal wage-hours contracts in the face of asymmetrically held information. With a single wage rate, we show that the problem of inefficient separations is so severe that, effectively, no specific training would take place. A wage premium on weekly overtime hours serves to reduce the effects of asymmetric information although it does not completely eliminate inefficiency. For those weekly hours for which a premium is paid, worker compensation exceeds the value of marginal product. There is an optimal automatic compensatory differential rule represented by an inverse relationship between the contractual wage and the overtime premium. Implications of imposing mandatory rules for premium pay and hours of work, as practiced in the United States, are assessed. The model is found to offer insights into important earlier finding in the literature.

JEL Classification: J41, J33

Keywords: overtime premium, specific human capital, asymmetric information, mandatory overtime rules

Corresponding author:

Robert A. Hart Department of Economics University of Stirling Stirling FK9 4LA Scotland UK Email: r.a.hart@stir.ac.uk

^{*} This is a revised version of the paper circulated under the same number until April 2008. We thank Dan Hamermesh, Steve Trejo, seminar participants at the Universities of Hamburg and Mainz as well as the U.K. Education and Employment Economics Group for helpful comments.

1. Introduction

Models based on agency (Lazear, 1981) and firm-specific human capital (Kahn and Lang, 1992) recognise that efficient long-term contracts must set hours as well as wages. Empirically, it is well recognised that unions bargain over both hourly wage rates and the length of working hours (Pencavel, 1991). The contract literature has stopped short, however, of providing explanations of why firms often pay for marginal daily or weekly hours at premium overtime rates. In any given week, roughly 25 percent of U.S. and 40 percent of U.K. male workers undertake paid overtime. Internationally, overtime hours are remunerated at a premium hourly rate, representing the straight time hourly rate plus an added percentage.¹ Why do firms pay this premium? This is clearly an important consideration in a country like the U.K. where the premium is set by collective agreement. But even in the US, where the government mandates a 50 percent premium for weekly hours in excess of 40, over 20 percent of overtime workers receive the premium for hours worked *before* the 40 limit (Trejo, 1993).

We develop a contract model that underpins the payment of an overtime premium. It belongs to the class of human capital models in which the parties have access to private information and that embrace issues of efficient investment and efficient separation.² There are two random components of the contract, returns to productivity and job satisfaction. The firm is better informed about the former and the worker is more knowledgeable about the latter. Period 1 is the investment period and period 2 the postinvestment period. Information is revealed about productivity and job satisfaction at the

¹ See OECD Employment Outlook (1998) for international comparisons of average premium rates.

² See Malcomson (1999) for a review of this work.

beginning of period 2. However, the asymmetric nature of this information necessitates that the parties form *a priori* conditional expectations of the returns at the start of period 1, the contract negotiation stage. Transaction costs of communicating and verifying information preclude contract renegotiation. In a seminal contribution, Hashimoto (1979) re-examined the original wage contract of Becker (1962) in which the parties share the returns to specific investments in period 2. He found that inefficient quits and layoffs can occur. This happens because the firm and the worker only consider their respective shares and each party does not take into account the wealth consequences for the other of a unilateral decision to separate. Thus separation can take place despite the fact that the *joint* return is positive. Even if the contract requires that the party initiating a separation should compensate the other, an approach developed in Mortensen (1978), the problem may remain because one side could act to induce the other to instigate a separation.

To resolve the problem of asymmetric information, Carmichael (1983) introduces a 'third party' into the set up. It cannot achieve the best solution of full efficiency, but it does serve to reduce the inefficiency. There is a steady-state number of period 2 workers who are divided into predetermined numbers of junior and senior jobs. Both types of job require the same amount of training and both are compensated at the second period wage. Additionally, senior jobs earn a bonus. A newly trained worker takes up a junior job at the start of period 2 and then stays in that job until length of job tenure dictates that it is his turn to fill a senior job vacancy. The critical point is that any incentive for the firm to fire a senior worker (or induce a quit) in order to avoid a high wage payment would be reduced because a junior worker would automatically fill the resulting job vacancy. In effect, junior workers act as the third party.

Under this promotion-ladder contract, the worker pays for the initial training and at the end of the training period receives a period 2 wage that is less than the alternative. The bonus payment ensures that the worker is recompensed for the incurred training costs and the low wage as a junior worker. More exactly, the bonus is chosen so that the worker's income is greater than the wage in period 2 by the expected value of the gain from the investment. Two points are worth noting about Carmichael's remuneration system. First, the prospects of amortising investments are improved as the contract ensures fewer inefficient separations and, therefore, longer expected tenure. Second, an interesting feature of the solution is that senior workers will typically be paid more and junior workers less than their respective marginal products. This provides a human capital explanation of pay/productivity outcomes usually linked with agency theory.

Seniority-based payments are an important feature of total compensation for many workers, especially in management and professional jobs. But they do not typically feature in the remuneration of, for example, blue-collar workers. One problem is that, *a priori*, workers might be reluctant to contemplate experiencing a period of relatively low pay - alongside better paid colleagues with the same skills - despite assurances that a spell of seniority will eventually arrive. But can Carmichael's strong contract formulation be derived from other payment structures that do not require different pay levels for equally trained workers?

Suppose we divide period 2 into types of hours rather than types of workers. Specifically, let all hours be paid for at the period 2 wage and let marginal hours receive an added premium. Further, let the wage-hours contract stipulate that this structure of remuneration applies to all equivalently trained workers. Then does the payment of such

a premium produce an efficient solution that is equivalent to that provided by a senioritybased bonus? We show in this paper that it does. We design a contract in which 'overtime' hours of all trained workers are compensated at above marginal product and straight time hours below marginal product. Our wage-hours contract does not require pay differentiation among trained workers. The overtime premium is chosen to ensure that average hourly pay is exceeded by a sufficient amount to compensate the worker by the expected value of the net gain from the initial investment. The firm is prevented from capturing this gain because all fully trained workers are guaranteed this payment structure. Further, joint wealth maximisation ensures that the parties will agree to work longer weekly hours in period 2. This is the intensive margin equivalent of longer tenure as a means of securing investment amortisation.

But we extend beyond establishing these parallels to a seniority-based bonus. We show that this way of viewing overtime and its payment structure may indeed be a fruitful avenue to explore. Thus, we consider the consequences for the contract if the payment of a premium on marginal hours is, as in the U.S., subject to statutory rules. In other words, what would be the consequences of the premium being set exogenously by an outside agency – typically the government – rather than by contractual agreement? It turns out that our results provide strong theoretical support for Trejo's so-called fixed job model (Trejo, 1991). In line with this author's U.S. empirical findings, we show that that the design of the optimal contract within our framework is such that a rise in the mandatory overtime premium will be offset by a reduction in the straight-time hourly wage rate.

2. Underlying framework

Our wage-hours model differentiates between an initial period in which both work and specific training are undertaken and a post-investment period where the investment affects productivity. The analysis is conducted in terms of the marginal worker who initially receives spot market wage earnings in a perfectly competitive labour market. Thus, prior to specific training in the initial period, the particular wage-hours combination available to the worker is determined by the market. The training endows the worker with job-specific skills and so in the second period he is differentiated from other workers in the spot market. The generation of a surplus in the training period allows the parties to set a wage-hours combination in the second period that differs from the market-equivalents.

The worker's pre-entry endowment of general human capital is worth w_a in the spot market and this is not augmented within the firm. Specific training is undertaken at a fixed (i.e. hours-independent) weekly cost, *C*. In period 1, the worker has hourly productivity $MPL_1 = w_a - C/h$, where h is first-period weekly hours. The expected value per unit of specific human capital is *M* so that specific training is expected to raise hourly productivity to $w_a + M$.

The parties negotiate the contract at the beginning of period 1 and there is no subsequent renegotiation. The contract contains an agreed value of investment return M: it may be simple to verify some of the elements that signal the level of productivity, such as the state of current and future orders for the firm's product. However, transaction costs of communicating and verifying information between the parties prevent agreement over the way in which random elements cause deviations from M. Such costs are

represented by a random variable η which has density function $f(\eta)$ and $E(\eta) = 0$. That is, the realised hourly productivity in period 2 is $w_a + M + \eta$. Due to lack of agreement over η , the firm responds unilaterally to the realised value of η at the end of period 1. The worker assesses the degree of job satisfaction θ in the firm, relative to potential outside opportunities, at the end of period 1. Again, transaction costs prevent a mutually agreed value of θ and only the worker responds to its realised value. The density function of θ is $q(\theta)$ with $E(\theta) = 0$. It is assumed that $Cov(\eta, \theta) = 0$. *Ex post*, information is private and cannot be exchanged and so separation decisions are made independently.

The probability of a worker deciding to quit is

$$Q = Q(\theta^*) = \int_{-\infty}^{\theta^*} q(\theta) d\theta$$
 (1)

while the probability of the firm wanting to fire a worker is

$$F = F(\eta^*) = \int_{-\infty}^{\eta^*} f(\eta) d\eta$$
(2)

where θ^* is the level of job satisfaction that leaves the worker indifferent about leaving and η^* is the level of productivity that leaves the firm indifferent over employing the worker. Without loss of generality, the discount rate is set to zero.

The worker works *h* and *H* weekly hours in periods 1 and 2 respectively, with the corresponding disutilities represented by D(h) and D(H). If H > h then the question arises as to how the marginal increase in hours is to be compensated. We examine the consequences of introducing an overtime premium payment $k \ge 1$ with respect to these hours.

The parties' joint wealth consists of the returns arising from three mutually exclusive and exhaustive events, weighted by the probability of their occurrence. A twoperiod time line of worker and producer surplus is illustrated in Figure 1. The worker may be fired or not-fired at the end of the first period. In the event of the worker not being fired, separation may occur due to a quit decision or the employment relationship may continue. In all three outcomes the first period surplus consists of wage earnings net of training cost and work disutility (w_a .h - C -D(h)). If the worker is fired or voluntarily quits, the second period surplus to the worker is given by the market value w_a .h -D(h); in these instances, the firm itself cannot obtain second period surplus. If the worker remains with the firm, second period surplus differs from the first period due enhanced productivity and job satisfaction as well as to the fact that second-period hours may differ from those in the first period.

Formally, the expected joint wealth *W* is expressed:

$$W = F.[w_{a}.h - C - D(h) + w_{a}.h - D(h)]$$
 (the worker is fired)
+ (1-F).Q.[w_{a}.h - C - D(h) + w_{a}.h - D(h)] (the worker quits)
+ (1-F).(1-Q).{w_{a}.h - C - D(h) + w_{a}.H + M.H + H.E(\eta/\eta > \eta^*) + H.E(\theta/\theta > \theta^*) - D(H)]} (the worker stays). (3)

Information concerning job satisfaction and productivity cannot be exchanged *ex post*. To maximise (3), we follow the two-step approach of Carmichael (1983). First, we solve for the optimal η , θ pair to maximise the objective function. Second, we assign second-period wages and hours to the first-order condition from step 1 to obtain the optimal wage-hour contract. From the Results Appendix (a) (see (A7) and (A8)), we

establish a necessary condition for achieving a constrained optimal contract under the assumption that both parties are risk neutral. The contract is offered to ensure that workers will quit whenever job satisfaction is too low; i.e.

$$\theta < \theta^* = -\{M + E(\eta/\eta > \eta^*) - w_a h/H + w_a + [D(h) - D(H)]/H \}.$$
(4)

Equivalently, the firm will fire the workers whenever productivity is too low; i.e.

$$\eta < \eta^* = -\{M + E(\theta|\theta > \theta^*) - w_a h/H + w_a + [D(h) - D(H)]/H\}.$$
(5)

These conditions imply that the party wishing to separate must be made to internalise the entire expected losses from the separation.

3. Earnings and Overtime

The earnings structure in period 2 is given as follows:

$$Y_2 = w_2 \cdot h + k \cdot w_2 \cdot (H - h) \tag{6}$$

where Y_2 is earnings in period 2, w_2 is the straight-time wage in period 2 and $k (\geq 1)$ is the overtime premium.

In establishing an optimal wage-hours contract, the objective of the parties is twofold. First, they seek to minimise sub-optimal separations. Second, since the size of the surplus is in part dependent on hours worked per-period, they wish to attain optimal second-period working time. These two objectives are not independent. Both the wage rate and hours affect productivity and so the process of minimising separations is dependent on is predicated on achieving optimal values of *both* w_2 and *H*.

(a) No overtime premium

What if overtime hours in (6) are compensated at the straight-time wage rate (ie. k = 1) so that earnings become $Y_2 = w_2 H$?

From the first-order conditions for maximising joint wealth in (3), derived in Results Appendix (a) (see (A12) and (A14)), we obtain

$$-\{E(\theta|\theta > \theta^*) - w_a h/H + [D(h) - D(H)/H]\} = w_2$$
(7)

and

$$w_2 = w_a + M + E(\eta/\eta > \eta^*) > w_a.$$
 (8)

As it stands, this particular wage-hours contract does not produce efficiency on either of the labour margins. Through (8), $w_2 = w_a + M + E(\eta/\eta > \eta^*)$; that is the worker is paid his expected productivity. The worker is fired by the firm if the wage is greater than actual productivity, that is $w_2 > w_a + M + \eta$. Thus, the critical value of η at which the worker is fired is $\eta^* = w_2 - M - w_a$. But substituting for w_2 from (8) implies that $\eta^* = E(\eta/\eta > \eta^*)$ which only occurs when η^* is the upper bound of η or when almost all workers are fired. On the other hand, using (4) and (8), the critical value of θ at which the worker quits is $\theta^* = w_a h/H - w_2 - [D(h) - D(H)]/H$. Combining this quit rule with (7) gives $\theta^* = E(\theta/\theta > \theta^*)$ which occurs when almost all workers quit.

What is the nature of the problem with the contract as it stands? The worker cannot observe his own productivity. Whenever the firm sets a firing threshold for the random component of the productivity, the worker's wage demand will reflect the expected compensation for this component conditional on not being fired. This implies that the conditional expectation is always higher than whatever threshold set by the firm. Consider the wedge between the threshold and the worker's conditional expectation. The firm will incur a loss if the *realised* return to the random component falls in the range of the wedge. The firm cannot fire the worker since the return is above the threshold and incurs a loss because the return falls short of the agreed conditional wage. In this circumstance, the only solution for the firm is to set the threshold to its maximum value. The wedge effectively disappears because expected compensation and the threshold now coincide and the firm breaks even on the marginal worker. The problem is that this involves setting a threshold at a level where most workers would be fired since their productivity will fall below this standard.

What is the problem from the worker's viewpoint? The worker receives an employment package consisting of a determined wage and stochastic job satisfaction. Again, asymmetric information dictates that, *a priori*, a conditional expectation of the random component of the return is required. The condition is that satisfaction must exceed the threshold that leaves the worker indifferent over whether or not to remain in the firm. Now, the worker would lose if the realised return to satisfaction lies between the threshold and the value of the conditional expectation. Again, the threshold must be set so that it equals the value of the conditional expectation. But this entails an inefficient quit rule since large numbers of workers will fall below this level of threshold due to an argument parallel to that under the firm's fire rule.

In essence, the contract with k = 1 is roughly equivalent to a spot market solution since almost all workers will quit/be fired before the second period, effectively resulting in no training taking place. The key to overcoming the problem is how to establish an 'efficient wedge' between each threshold and its respective conditional expectation that prevents the firm from making a loss when the realised return falls within the range of the

wedge. We proceed to show that choosing an optimal overtime premium that is in excess of unity offers such a solution.

(b) The payment of an overtime premium

We now consider the payment of a wage premium which renders pay on marginal hours to be greater than hourly straight-time pay; that is k > 1 in (6). We begin by establishing that this serves to counter the contract inefficiency outlined above.

With k > 1, the solutions which maximise joint wealth in (3) are given by (A20) of Results Appendix (a):

$$-\{M + E(\theta|\theta > \theta^*) + w_a \cdot w_a \cdot h/H + [D(h) - D(H)]/H\} = -[w_a + M \cdot k \cdot w_2]$$
(9)

and, using (A22) from Results Appendix (a),

$$-[M + E(\eta/\eta > \eta^*)] = -[w_2.(\gamma + 1) - w_a]$$
(10)

where $\gamma = (k-1).(1 - h/H) > 0.$

If the firm solves (9) and (10) simultaneously, it will attain an optimal wage-hours (w_2^*, H^*) contract, conditional on the level of premium, *k*.

From Appendix Results (d), we show that there exists a k $(1 < k \le \min [kmax_1, kmax_2])$ that maintains

$$w_2^* > w_a \tag{11}$$

and

$$H^* > h. \tag{12}$$

In order to see how the inefficient quit rule of the previous section (i.e. for k = 1) has been overcome, we combine (9) with the quit rule $\theta^* = w_a h/H - w_2 - [D(h) - D(H)]/H$ to obtain $\theta^* = E(\theta/\theta > \theta^*) - w_2.(k-1).h/H$. Note that in the previous model formulation (i.e. with k = 1) the critical value of θ at which the worker quits is given by $\theta^* = E(\theta/\theta > \theta^*)$ which occurs when almost all workers quit. In the new equivalent result, the overtime premium provides a wedge that serves to satisfy the necessary condition for *ex ante* joint wealth maximisation. The introduction of the premium allows the attainment of a similar optimal solution with respect to firing: combining (10) with the fire rule $\eta^* = w_2$ - $M - w_a$ we obtain $\eta^* = E(\eta/\eta > \eta^*) - \gamma.w_2$.

How does the overtime premium serve to reduce the inefficiency inherent in the firing rule in the previous single-wage scenario? To understand this, we return to the wedge - that is the gap between the productivity (firing) threshold and the conditional expected value of random productivity. With a single wage rate for all hours in period 2, we found that the firm makes a loss if realised productivity falls within the range of the wedge. With a two-step earnings schedule, the firm can avoid this loss. We establish below that the optimal premium lies above expected productivity and the wage rate below. Therefore, the premium can be set so that, if the realised return falls in the range of the expected return to the worker in respect of straight time pay. As a result, it is not necessary to set the threshold to the maximum of the random component. This implies that the overtime premium serves as an additional instrument that helps to avoid the need to undertake excessive firing. A parallel argument pertains to quits, with the premium again serving to reduce inefficient separations.

From (10), it can be shown that

$$k.w_2 > MPL_2 = w_a + M + E(\eta/\eta > \eta^*) > w_2$$
(13)

or premium pay is higher than the expected marginal product of labour in period 2 (MPL_2) while the straight-time wage is lower than MPL_2 .

The precise role of the optimum overtime premium can be explained in terms of the worker's optimal hourly pay in period 2. From (6), average hourly earnings in period 2 is given by

$$y_2 = Y_2/H = w_2 + w_2.(k - 1).(H - h)/H$$

where k - 1 is the overtime mark-up. Now, it is easy to show that the mark-up (k - 1) must be chosen so that when applied to overtime hours (H - h), the average hourly payment (y_2) must exceed what the firm can save by firing the worker $(k.w_2)$ by the expected value of the gain from the investment.³

The payment of a wage premium k > 1 for hours (H - h) allows the parties to achieve an optimal wage-hours contract that consists of second-period wages and hours that are higher than their first-period equivalent values. By contrast, there are inefficient quits and layoffs if the premium is set to unity, that is if a single wage is paid in the second period. This would seem to imply that an increase in k is associated with

³ To show this, we can express the gain from the investment as:

Gain = {[$w_a + M + E(\eta/\eta > \eta^*) + E(\theta/\theta > \theta^*)$].H - D(H)} - [$w_a.h - D(h)$] or, re-arranging, Gain = $w_a.(H - h) + [M + E(\eta/\eta > \eta^*) + E(\theta/\theta > \theta^*)].H + [D(h) - D(h)]$. From, (10) we have $y_2 = w_a + M + E(\eta/\eta > \eta^*)$. Finally, using (9), we have $y_2 - k.w_2 =$ Gain.

reductions in separations. This is confirmed formally in Results Appendix (a) (see (A24) and (A25)). It is shown that there exists a $kmax_3 > 1$ such that if $k \in (1, kmax_3]$ then

$$\partial \theta^* / \partial k < 0 \tag{14}$$

and

$$\partial \eta^* / \partial k < 0 . \tag{15}$$

Figure 2 illustrates the effect of an increased *k* on the quit/layoff decisions.

On the worker's side, the result in (14) is quite straightforward. Since overtime is remunerated at a premium rate in excess of marginal product, it is unsurprising that a rise in k will lead to a higher *H* and a lower quit probability. But, from (15), these outcomes appear to fall short of providing contract optimality from the firm's viewpoint. While k >1 provides a desirable instrument towards achieving a contractual solution, the firm would be reluctant to raise the proportion of premium payments within total labour costs in the absence of offsetting cost reductions. In other words, this would act merely to raise the worker's surplus. However, the optimal wage-hours contract contains an in-built compensating differential. We find from (13) that

$$\frac{\partial w_2}{\partial k} < 0$$
 (16)

or there is an inverse relationship between the contractual wage and the overtime premium. Therefore, there exists an "optimal automatic compensating differential rule". A compensating differential reaction in the straight-time wage to an increase in k is embedded in the contract solution and therefore (15) holds. The full solution to the contract is as follows. Let $k^* = min\{kmax_1, kmax_2, kmax_3\}$. Then (9), (10), (11), (12), (14) and (15) will all hold. Given k^* , solving (12) and (13) simultaneously gives us w^* and H^* . And then (w_2^*, H^*, k^*) forms the efficient contract. The solutions are illustrated graphically in Figure 3.

To the extent that these overtime outcomes are feasible, the model predicts that, as with straight-time hourly wages, hours should rise with job tenure. We note that the life cycle models of Ghez and Becker (1975) predict that hours supplied to the market place would be positively related to the price of time over the life cycle. Consistent with this approach is the expectation that workers would contribute their highest average hours when they are most productive, that is in the middle years of their work experience.

As illustrated in Figure 5 (a), the payment of an hours premium in order to effect efficient turnover produces the result that the worker is paid more than marginal product in period 2. Suppose that weekly compensation is based on x number of straight time hours and y number of overtime hours. The optimal payments configuration is to pay the former hours at below and the latter hours above their respective marginal products. There are two parallel payments for overtime and straight time hours which are set for the entire 2nd period and for all workers.

4. Effects of statutory overtime regulations

The wage-hours contract developed in the previous section involves the parties reaching agreement on optimal values of the straight-time wage, the overtime premium and the length of per-period hours. In several OECD economies, statutory overtime pay restrictions limit negotiated degrees of freedom as far as setting the premium is concerned (OECD Employment Outlook, 1998). The usual policy objective behind overtime legislation is to encourage employment expansion by raising marginal cost on the intensive relative to the extensive labour margin. Among the most stringent, comprehensively binding, and simplest sets of legislative measures is the US Fair Labor Standards Act (FLSA). Under FLSA rules, covered workers must be paid an overtime wage of at least one and one-half times the straight-time wage for weekly hours in excess of 40. Let these hours and premium constraints be denoted, respectively, by \overline{h} and \overline{k} (i.e. $\overline{h} = 40$, $\overline{k} = 1.5$). Here, we discuss several implications of our model in relation to this Act.

If we rule out short-time working, Figure 4 illustrates three cases between actual straight-time hours worked in period 1 (*h*), maximum permitted straight-time hours before a premium must be paid (\overline{h}), and optimal second period hours (*H**). All period 2 contracts under the legislation - indicated as either Case A or Case B or Case C - occur at or below the 45⁰ line given that period 2 hours are greater than period 1 hours. Here, we present a non-exhaustive discussion different possible contracts.

Case A

The simplest case in terms of the previous analysis concerns the outcome $h < H^* \le \overline{h}$. Period 2 hours are above those in period 1 but below the mandatory maximum that trigger overtime regulations. As illustrated in Figure 5, the optimal contract would be one whereby the internal premium k^* is paid on hours H^* -h. The implication is that it would be optimal to pay a premium on marginal hours (i.e. H^* -h) despite no compulsion so to do. This outcome is consistent with the interesting observation that a significant proportion of US firms pays an overtime premium even when weekly hours are *below* the statutory minimum set under FLSA rules. Trejo (1993) presents 1985 Current Population

Survey data which reveal that over 20 per cent of workers sampled receive premium pay for hours less than the 40 hour limit; in fact, for 10 per cent of workers, the premium started at 34 hours or less.⁴ Moreover, both union and non-union workers display similar percentages in these respects.

Case B

A commonly occurring case is that of $H^* > \overline{h} \ge h$ and $\overline{k} > k^*$.⁵ These relative outcomes fit with the work-sharing policy-logic discussed above. Two related labour market issues arise.

(i) Compensating differentials

Trejo (1991) presents U.S. and Bell and Hart (2003) U.K. empirical evidence that firms mitigate rises in mandatory premium payments by reducing straight-time wages. This response accords with the so-called 'fixed-job' model whereby firms and workers adhere to agreed compensation packages by lowering straight-time wages in response to increases in overtime premium pay or overtime eligibility. Our model provides theoretical justification for such behaviour by the two parties. Equation (13) in the previous section shows that an automatic compensating differential rule is integral to the optimising behaviour suggested by the wage-hours model. A rise in *k* due to legislation

⁴ Although due to (an under-recorded) anomaly resulting from the fact that some workers receive overtime pay according to a daily standard, Trejo reports that the figures may exaggerate the frequency with which workweek standards other than the 40-hour standard may occur.

⁵ If $k^* \ge \overline{k}$, then the analysis is effectively the same as under Case A.

will result in an agreed offsetting lowering of w_2 . Worker compensation and the outside opportunity will be left unaffected with no incentive to expand employment.

(ii) Non-compliance

Based on late-1970s data, Ehrenberg and Schumann (1982, Ch. 5) indicate that a 10 percentage rate of non-compliance with the FLSA regulations would constitute a conservative estimate (see also Trejo, 1993). If the parties seek this illegal strategy, what type of non-compliance might appear to be most mutually advantageous? The firm and its workers may reach an implicit agreement at the beginning of period 1 to work "unpaid" hours in period 2 to the extent that the value of k, averaged over paid and unpaid overtime hours, is equal to $k^{*.6}$ One such agreement along these lines is illustrated in Figure 6 (a). Here, the parties agree to work paid and unpaid hours in period 2 that solves

$$k^* \cdot w_2^* \cdot (H^* - h) = (\overline{h} - h) \cdot k^* \cdot w_2^* + \overline{k} H_p + 0 \cdot H_u$$
(17)

where, in period 2, H_p is paid-for and H_u is unpaid-for hours. Under this arrangement, the firm pays for hours in excess of period 1 straight-time hours and up to \overline{h} at the optimal premium rate and tackles the problem of $\overline{k} > k^*$ by manipulation of the length of unpaid overtime. This latter exercise determines H_p and H_u as

⁶ Unpaid overtime is quantitatively very significant in the U.K. While there are no national-level overtime regulations in this economy, premium pay and the maximum level of straight-time pay may be set at industry-level - such as in the engineering industry - and effectively act as mandatory rules for relevant firms. The mandatory overtime-related reason for working unpaid hours could be one of several explanations of this phenomenon, as discussed by Bell and Hart (1999).

$$H_p = \frac{k^*}{\bar{k}} (H^* - h) \tag{18}$$

and

$$H_u = \left(1 - \frac{k^*}{\bar{k}}\right)(H^* - h). \tag{19}$$

As with Case A, the firm pays a premium before mandatory hours, \overline{h} , are reached. An alternative strategy to achieve k^* is to agree to commence paying the premium $k = \overline{k}$ for hours above h.⁷

In the absence of a downward adjustment of w_2 , or some form of non-compliance, $\overline{k} > k^*$ implies that either \overline{h} or \overline{k} become binding. The former outcome is illustrated in Figure 6 (b) where no overtime would be undertaken.

Case C

An analysis parallel to Case B applies here; for example, non-compliance incentives are again relevant.

5. Conclusions

Overtime premium pay is an important variable in labour market economics and macroeconomics because, in many instances, it represents the marginal cost of labour input. Existing reasons for the payment of a premium include custom and practice and

⁷ Under this agreement, $k^*.w_2^*.(H^*-h) = (\overline{h} - h).w_2^* + \overline{k}.H_p.w_2^* + 0.H_u$, in which case $H_p = [h - \overline{h} + k^*(H^*-h)]/|\overline{k}|$ and $H_u = H^* - [h - \overline{h} + k^*(H^*-h)]/|\overline{k}|$.

heterogeneous worker preferences (see Hart, 2004). Here, we offer a new approach to understanding the use of premium payments based on a more general wage contract formulation. Our model allows for changes in labour inputs on the extensive (workers) and intensive (working hours) margins. It also incorporates information asymmetries between workers and firms. We show that it is in the interest of the firm and its workforce to increase *both* wages and hours once investments have been sunk. However, the payment of a wage premium for additional hours is required in order to achieve an optimal wage-hours contract. Our results are consistent with a number of recent important empirical findings concerning the operation of overtime working.

- Bell, D N F and R A Hart. "Unpaid work." Economica 66, (1999), 271-90.
- Bell, D N F and R A Hart "Wages, hours, and overtime premia: evidence from the British labor market", Industrial and Labor Relations Review, 56, (2003), 470 480.

Becker, G. Human capital, (1962), New York: Columbia University Press.

- Carmichael, L. "Firm-specific human capital and promotion ladders", *Bell Journal of Economics* 14, (1983), 251-58.
- Ehrenberg, R G. Fringe benefits and overtime behavior, (1971), Lexington, Mass.: Heath.
- Ehrenberg, R G and P L Schumann. *Longer hours or more jobs?* Cornell Studies in Industrial and Labor Relations, no. 22, (1982), Ithaca, NY: Cornell University.
- Ghez, G R and G S Becker. *The allocation of time and goods over the cycle*, 1975, New York: National Bureau of Economic Research.
- Hall, R E and E P Lazear. "The excess sensitivity of layoffs and quits to demand", (1984), *Journal of Labor Economics* 2, 233-57.
- Hart, R A. *The Economics of Overtime Working*. 2004. Cambridge: Cambridge University Press.
- Hashimoto, M. "Bonus payments, on-the-job training and lifetime employment in Japan." Journal of Political Economy 87 (1979): 1086-1104.
- Hashimoto, M and B T Yu, "Specific capital, employment contracts and wage rigidity", *Bell Journal of Economics*, (1980), 11, 536-49.
- Kahn, S and K Lang, "Constraints on the choice of hours", *Journal of Human Resources* 27, (1992), 661 678.
- Lazear, E P. "Agency, earnings profiles, productivity, and hours restrictions", *American Economic Review* 71 (1981): 606-20.
- Malcomson, J M. "Individual employment contracts",<u>Handbook of Labor Economics</u>, volume 3/4 (eds. Orley Ashenfelter and David Card), (1999), Elsevier Science.
- Mortensen, D. "Specific capital and labor turnover", *Bell Journal of Economics* 9, (1978), 572-86.
- OECD Employment Outlook, chapter "Working hours: latest trends and policy initiatives", (1998), Paris: OECD.
- Pencavel, J H, Labor markets under trade unionism, (1991), Oxford: Blackwell.

- Trejo, S J. "The effects of overtime pay regulation on worker compensation.", *American Economic Review*, 81, (1991), 719-40.
- Trejo, S J. "Overtime pay, overtime hours, and labor unions." *Journal of Labor Economics*, 11, (1993), 253-78.

Figures

Figure 1: Time line of worker and producer surplus







Figure 3: Wage, hours, and premium profiles



Figure 6: Case B contracts



Results Appendix

(a) Joint wealth maximisation

In order to illustrate the optimisation with respect to the joint wealth defined in (3) in the main text, we begin by presenting a number of useful lemmas.

Lemma 1

$$\frac{d}{du}\int_{a(u)}^{b(u)}g(x,u)dx=\int_{a(u)}^{b(u)}\frac{\partial g}{\partial u}dx+g(b,u)b'-g(a,u)a'$$

Lemma 2

$$dQ/d\theta^* = q(\theta^*)$$
 and $dF/d\eta^* = f(\eta^*)$

Proof. From (1), (2) and lemma 1.

Lemma 3

$$1 - Q = \int_{\theta^*}^{\infty} q(\theta) d\theta \quad \text{and} \quad 1 - F = \int_{\eta^*}^{\infty} f(\eta) d\eta$$

Lemma 4

$$E(\theta|\theta > \theta^*) = \frac{1}{1 - Q} \int_{\theta^*}^{\infty} \theta q(\theta) d\theta$$
$$E(\eta|\eta > \eta^*) = \frac{1}{1 - F} \int_{\eta^*}^{\infty} \eta f(\eta) d\eta$$

Lemma 5

$$\frac{dE(\theta|\theta > \theta^*)}{d\theta^*} > 0 \text{ and } \frac{dE(\eta|\eta > \eta^*)}{d\eta^*} > 0$$

Proof

From lemma 1 and 4,

i

QED.

$$\frac{dE(\theta|\theta>\theta^*)}{d\theta^*} = -\frac{\theta^*q(\theta^*)}{1-Q} + \frac{q(\theta^*)\int_{\theta^*}^{\infty}\theta q(\theta)d\theta}{(1-Q)^2} = \frac{q(\theta^*)}{(1-Q)^2}\int_{\theta^*}^{\infty}(\theta-\theta^*)q(\theta)d\theta > 0$$

Similar proof applies to $dE(\eta/\eta > \eta^*)/d\eta^*$.

QED.

[wa.h - C - D(h)] in W:

To maximise W defined in (3) in the main text, we first collect the item

$$W = [wa.h - C - D(h)]$$
(defined as X)
+[F+(1-F)Q][wa.h - D(h)] (defined as Y)
+(1-F)(1-Q){wa.H + H.M + H.E(\eta/\eta > \eta*) + H.E(\theta/\theta > \theta*) - D(H)} (defined as Z)
= X + Y + Z (A1)

The first order conditions to max W in (A1) are:

$$\partial W/\partial \theta^* = 0$$
 and $\partial W/\partial \eta^* = 0$

To derive $\partial W / \partial \theta^* = 0$ firstly:

$$\partial X / \partial \theta^* = 0 \tag{A2}$$

$$\partial Y/\partial \theta^* = (1-F).[dQ/d\theta^*].[wa.h-D(h)] = (1-F).q(\theta^*).[wa.h-D(h)]$$
(A3)

(from lemma 2)

Rewrite Z into:

$$Z = (1 - F)(1 - Q) \left[wa.H + M.H + \frac{-\infty}{1 - F} H + \frac{\int_{-\infty}^{\theta^*} \theta q(\theta) d\theta}{1 - Q} H - D(H) \right]$$
(A4)

(from lemma 4)

Substituting (1) and (2) from the main text into (A4), we have:

$$Z = \int_{\theta^* \eta^*}^{\infty} [wa.H + M.H + \eta.H + \theta.H - D(H)]q(\theta)f(\eta)]d\theta d\eta$$
(A5)

Differentiating (A5) with respect to θ^* , we have:

$$\frac{\partial Z}{\partial \theta^*} = -\int_{\eta^*}^{\infty} [wa.H + M.H + \eta.H + \theta.H - D(H)]q(\theta^*)f(\eta)d\eta$$
$$= -(1-F).q(\theta^*).\{wa.H + M.H + H.\theta^* + H.E(\eta|\eta > \eta^*) - D(H)\}$$
(A6)

Therefore,

$$\frac{\partial W}{\partial \theta^*} = \frac{\partial X}{\partial \theta^*} + \frac{\partial Y}{\partial \theta^*} + \frac{\partial Z}{\partial \theta^*}$$

= - [-wa.h + D(h) + wa.H + M.H + H. \theta^* + H.E(\eta/\eta > \eta^*) - D(H)].(1-F).q(\theta^*) = 0 (A7)

By symmetry,

$$\partial W/\partial \eta^* = -[-wa.h + D(h) + wa.H + M.H + H.\eta^* + H.E(\theta|\theta > \theta^*) - D(H))].(1-F).f(\eta^*) = 0$$
(A8)

Information concerning job satisfaction and productivity cannot be exchanged *ex post*. A necessary condition for achieving a constrained optimal contract derives from (A7) and (A8). The contract is offered to ensure that workers will quit whenever job satisfaction is too low, i.e.

$$\theta < \theta^* = -\{M + E(\eta(\eta > \eta^*) - w_a h/H + w_a + [D(h) - D(H)]/H\}$$
(A9)
(from (A7))

Equivalently, the firm will fire the workers whenever productivity is too low, i.e.

$$\eta < \eta^* = -\{M + E(\theta(\theta > \theta^*) - w_a h/H + w_a + [D)(h) - D(H)]/H\}$$
(A10)

(from (A8))

[Equations (A9) and A(10) are (4) and (5) in the main text, respectively.]

These conditions imply that the party wishing to separate must be made to internalise the entire expected losses from the separation. Next we show that it is possible to formulate a wage-hour contract in terms of the parameters of η^* and θ^* to satisfy (A9) and (A10).

The firm will fire the worker if return is below the second period wage, that is if

$$w_a + M + \eta < w_2$$

or

$$\eta < \eta^* = -(w_a + M - w_2) \tag{A11}$$

Substituting (A10) into (A11) gives:

$$-\{M + E(\theta(\theta > \theta^*) - w_a.h/H + w_a + [D(h) - D(H)]/H\} = -(w_a + M - w_2)$$

or

$$-\{E(\theta E(>\theta^*) - w_a h/H + [D(h) - D(H)/H]\} = w_2$$
(A12)

(This is equation (7) in the main text.)

The worker will quit if second period return is less that the outside opportunity, that is

if

$$H.w_2 + H.\theta - D(H) < w_a.h - D(h)$$

or

$$\theta < \theta^* = w_a \cdot h/H - w_2 - [D(h) - D(H)]/H$$
 (A13)

Substituting (A9) into (A13) gives:

$$w_2 = w_a + M + E(\eta/\eta > \eta^*) = MPL_2$$
(A14)

where MPL_2 is the expected productivity of the worker.

[(A14) is equation (8) in the main text.]

If the firm solves (A12) and (A14) simultaneously, it will achieve an optimal wagehour (w_2^* , H^*) contract. Can working hours allocated in the post-investment period serve to induce efficient turnover? That is, do we obtain $\partial \theta^* / \partial H < 0$ and $\partial \eta^* / \partial H < 0$? Substituting (A14) into (A12) and multiplying H on both sides produces

$$-H.E(\theta|\theta > \theta^*) + w_a.h - D(h) + D(H) = wa.H + M.H + H.E(\eta|\eta > \eta^*).$$
(A15)

Given η^* , differentiating (A15) gives

$$\partial \theta^* / \partial H = - \left[E(\theta | \theta > \theta^*) - D'(H) + w_a + M + E(\eta | \eta > \eta^*) \right] / \left[H.dE(\theta | \theta > \theta^*) / d\theta^* \right] . \tag{A16}$$

If the marginal disutility of working D'(H) is less than the sum of the second period wage rate and the expected job satisfaction, then

$$\partial \theta^* / \partial H < 0 \tag{A17}$$

(by Lemma 5).

By symmetry,

$$\partial \eta^{*} / \partial H = - \left[E(\theta | \theta > \theta^{*}) - D'(H) + w_{a} + M + E(\eta | \eta > \eta^{*}) \right] / \left[H.dE(\eta | \eta > \eta^{*}) / d\eta^{*} \right].$$
(A18)

Again, if the marginal disutility of working D'(H) is less than the combined value of the second period wage and expected job satisfaction, then

$$\partial \eta^* / \partial H < 0$$
. (A19)

(by Lemma 5).

Therefore, within the marginal disutility constraints, equations (A17) and (A19) state that an increase in hours in the post-investment period will reduce the probability of separation.

An efficient contract with a marginal hours' premium

Now we consider that additional working hours are compensated at a premium, k, where k > 1.

Firstly, if the firm decides to fire a worker, then equations (A11) and (A12) should be modified as the marginal product is now $k.w_2$. Thus, (A12) becomes

$$-\{M + E(\theta|\theta > \theta^*) + w_a - w_a \cdot h/H + [D(h) - D(H)]/H\} = -[w_a + M - k \cdot w_2]$$

(This is equation (9) in the main text)

or

$$-\{ E(\theta|\theta > \theta^*) - w_a \cdot h/H + [D(h) - D(H)]/H \} = k \cdot w_2 .$$
(A20)

The premium also alters the worker's decision to quit:

$$w_{2}h + k.w_{2}(H-h) + \theta.H - D(H) < w_{a}h - D(h)$$

or

$$\theta < \theta^* = w_a \cdot h/H - w_2 \cdot h/H - k \cdot w_2 \cdot (1 - h/H) - [D(h) - D(H)]/H.$$
(A21)

Substituting (A9) into (A21) gives:

$$-[M + E(\eta | \eta > \eta^*)] = -[w_2.(\gamma + 1) - w_a]$$
(A22)

where $\gamma = (k-1).(1 - h/H) > 0.$

(This is equation (10) in the main text.)

That is:

$$w_2 = [w_a + M + E(\eta/\eta > \eta^*)]/(\gamma + 1) < MPL_2 = w_a + M + E(\eta/\eta > \eta^*).$$
(A23)

If the firm solves (A20) and (A22) simultaneously, it will attain an optimal wage-hour (w_2^*, H^*) contract, conditional on the level of premium, *k*.

Now we turn to discuss the implications of this optimal contract.

First, as we discussed in the main text, we know that if k=1 then $\partial \theta^*/\partial H < 0$ and $\partial \eta^*/\partial H < 0$ from (A17) and (A19). By the continuity assumption, there exists a $kmax_2 > 1$ such that if $k \in (1, kmax_2]$ then $\partial \theta^*/\partial H \le 0$ and $\partial \eta^*/\partial H \le 0$. That is, the working hours allocated to the trained workers can induce efficient turnover under the premium scheme.

Second is to establish that the premium can further induce efficient turnover; i.e.

$$\partial \theta^* / \partial k < 0 \tag{A24}$$

and

$$\partial \eta^* / \partial k < 0 . \tag{A25}$$

Re-write (A24) and (A25) as:

$$\partial \theta^* / \partial k = (\partial \theta^* / \partial H) . (\partial H / \partial k) \tag{A26}$$

and

$$\partial \eta^* / \partial k = (\partial \eta^* / \partial H) . (\partial H / \partial k) . \tag{A27}$$

We know that $\partial \theta^* / \partial H < 0$ and $\partial \eta^* / \partial H < 0$ for small k > 1. Hence to show (A24) and (A25) hold we need only to demonstrate that

$$\partial H/\partial k > 0 \tag{A28}$$

for some k>1. The logic is that if increasing k from k = 1 would increase working hours H and increasing H will reduce sub-optimal separation, then a premium k>1 would improve efficiency.

In Results Appendix result (c) it is shown that if k=1 then dH/dk > 0. Hence by the continuity assumption, there exists a $kmax_3 > 1$ such that if $k \in (1, kmax_3]$ then dH/dk

> 0. This establishes (A24) and (A25), as well as (14) and (15) in the main text.

(b) First- and second-period hours

For any $h \in (0, h_0)$ (where h_0 satisfies $w_a(h_0) - D(h_0) = 0$), equation (7) in the main text has two solutions H^* and H^{**} which satisfy $0 < H^{**} < h < H^*$ if D(0) = 0, D'(.) > 0 and D'(.) > 0.

Proof

Equation (7) is the same as (A12) in Results Appendix (a). By substituting (A14) into (A12) we obtain (A15). Re-write equation (A15) as

$$w_{a} \cdot h - D(h) = H \cdot [w_{a} + M + E(\eta / \eta > \eta^{*}) + E(\theta / \theta > \theta^{*})] - D(H)$$

Let $w_b \equiv w_a + M + E(.) + E(.)$ and so we have $w_b > w_a > 0$. Define $g_1(h) \equiv w_a \cdot h - D(h)$ and $g_2(H) \equiv w_b \cdot H - D(H)$. Since $g_1''(h) = D''(h) < 0$ and $g_1''(H) = D''(H) < 0$ both $g_1(h)$ and $g_2(H)$ are concave (see Figure A1). Also, $g_2(x) - g_1(x) = (w_b - w_a)x > 0$ for any x > 0.



Hence $H_0 > h_0 > 0$ where $g_2(H_0) = g_1(h_0) = 0$. From Figure A1, for any $h \in (0, h_0)$ there are two solutions H^* and H^{**} for equation $g_1(h) = g_2(H)$ and $0 < H^{**} < h < H^*$. QED.

(c) The relation between second-period hours and the overtime premium

We show here that (A28), $\partial H/\partial k > 0$, holds provided that k and the disutility of working are not too large.

Substituting (A23) into (A20) produces

$$(1+\gamma).[-E(\theta/\partial > \theta^*) + w_a.h/H - D(h)/H + D(H)]/H] = k.[w_a + M + E(\eta/\eta > \eta^*)]$$
(C1)

where $\gamma = (k-1).(1 - h/H) > 0.$

Differentiating (C1) gives

$$A.[-h/H^{2}.dH + dk.(1-h/H) + (k.h/H^{2}).dH] + (1+\gamma).B.dH = G.dk$$
(C2)

where

$$A = -E(\theta | \theta > \theta^*) + w_a \cdot h/H - D(h)/H + D(H)/H = k \cdot w_2 > 0$$
(C3)

$$B = w_{a} \cdot h/H^{2} + D(h)/H^{2} + D'(H)/H - D(H)/H^{2} > 0 \text{ (if } D(H) \text{ is not too large)}$$
(C4)

$$G = w_a + M + E(\eta/\eta > \eta^*) > 0$$
(C5)

From (C2):

$$\frac{\partial H}{\partial k} = \frac{G - A\left(1 - \frac{h}{H}\right)}{(1 + \gamma\gamma). + A\left(\frac{-h + kh}{H^2}\right)} > 0$$
(C6)

where

$$G - A(1-h/H) = wa.h/H + M + E(\eta/\eta > \eta^*) + E(\theta/\theta > \theta^*) + D(h)/H - D(H)/H > 0$$
(C7)

if D(H) is not too large. Both (C4) and (C7) would be satisfied if marginal disutility is constant (i.e. a linear disutility function) or the increase in disutility from period 1 to period 2, D(H) - D(h), is less than the wage earning of period 1, *wa.h*. (d) Existance of k that maintains (i) $w_2^* > w_a$ and (ii) $H^* > h$.

We show here that there exists a *k* that maintains (i) $w_2^* > w_a$ and (ii) $H^* > h$, i.e. both equations (14) and (15) in the main text hold..

From (8), we know that if k=1 then

$$w_2^* > w_a. \tag{D1}$$

By assuming continuity, there exists a $kmax_1 > 1$ such that if $k \in [1, kmax_1]$ then w_2^* > w_a from (13). So the well-known result that the contractual wage rate increases with tenure is retained. Similarly, we also know that if k=1 then $\partial \theta^*/\partial H < 0$ and $\partial \eta^*/\partial H < 0$ from (A17) and (A19). By the continuity assumption, there exists a $kmax_2 > 1$ such that if $k \in [1, kmax_2]$ then $\partial \theta^*/\partial H \le 0$ and $\partial \eta^*/\partial H \le 0$. That is, the working hours allocated to the trained workers can induce efficient turnover. The implied premium profile is shown in Figure 3. Parallel to the developments in the previous section, it can be shown that

$$H^* > h \tag{D2}$$

for any *k* if $1 < k \le \min[kmax_1, kmax_2]$.