



**UNIVERSITY OF
STIRLING**

DEPARTMENT OF
ECONOMICS

Wage-Hours Contracts, Overtime Working and Premium Pay

Robert A. Hart

Yue Ma

Stirling Economics Discussion Paper 2008-04

April 2008 (updated in October 2008)

Online at <http://www.economics.stir.ac.uk>

Wage-Hours Contracts, Overtime Working and Premium Pay

ROBERT A HART

Department of Economics,
University of Stirling,
Stirling FK9 4LA,
Scotland, UK
(email: r.a.hart@stir.ac.uk)

and

YUE MA

Department of Economics
Lingnan University
Hong Kong
(email: yuema@ln.edu.hk)

October 2008

ABSTRACT:

This paper offers a contract-based theory to explain the determination of standard hours, overtime hours and overtime premium pay. We expand on the wage contract literature that emphasises the role of firm-specific human capital and that explores problems of contract efficiency in the face of information asymmetries between the firm and the worker. We first explore a simple wage-hours contract without overtime and show that incorporating hours into the contract may itself produce efficiency gains. We then show how the introduction of overtime hours, remunerated at premium rates, can further improve contract efficiency. Our modelling outcomes in respect of the relationship between the overtime premium and the standard wage rate relate closely to earlier developments in hedonic wage theory. Throughout, we emphasise the intuitive reasoning behind the theory and we also supply relevant empirical evidence. Mathematical derivations are provided in an appendix.

Keywords: wage-hours contracts, overtime, premium pay, specific human capital, asymmetric information

JEL CLASSIFICATION NUMBERS: J41, J33

Acknowledgements: We thank two referees, 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 many firms employ overtime hours for which they pay premium rates. Yet overtime working is an important aspect of total working hours determination. In the U.K., and based on the British Household Panel Survey (BHPS), an annual average of one-third of male employees worked paid overtime between 1991 and 2005. Moreover, since this economy experiences virtually no exogenous rules and regulations governing the use of overtime hours, these proportions suggest that there may be considerable advantages to the firm and its workforce in adopting such a working time arrangement. But why is it so popular? There are surprisingly few theories that attempt to provide economic explanations for this working time arrangement.¹

Our explanation of why some firms make use of paid overtime and reward overtime hours at premium rates is embedded in wage-hours contract theory. That is, we concentrate on wages and hours setting based on agreement between employer and employee. Undoubtedly, the British economy offers one of the best examples of a labour market in which paid overtime is subject largely to contractual agreement. As stated by Income Data Services (IDS, 1997), it is generally the case that “the number of hours

¹ Hart (2004, Chapter 5) offers a summary of existing explanations.

(including overtime) that an employee can be expected to work is a matter to be agreed between employer and employee. The terms relating to working hours should be set out in full in the employee's written statement of terms and conditions of employment which must be supplied under the...(Employment Rights Act)".² Interestingly, although overtime is overwhelmingly paid for at wage rates in excess of standard, or basic, hourly rates, there is no legal requirement in Britain that premium rates should apply.

In stark contrast to Britain's *laissez-faire* attitude to paid overtime, the United States government imposes strict overtime controls on most workers. The Fair Labor Standards Act sets standard weekly hours at 40 beyond which marginal hours have to be remunerated at a minimum rate of one-and-a-half times the standard hourly rate. However, evidence provided by Trejo (1993) suggests that 20 per cent of overtime hours is paid at a premium above standard rates before the weekly 40 hour-limit is reached. This points to the likelihood that, irrespective of outside rules and regulations, bargaining parties in some U.S. firms perceive internal advantages in employing weekly overtime hours that are paid for at premium rates.

Our theory builds out from the two-period wage contract models in which the firm and its workers undertake specific human capital in period 1 and then share the surplus during their working relationship in period 2. Management and workers are assumed to be asymmetrically informed about internal and external values of workers' productivities. Due to high transaction costs of communicating and verifying privately-held information

² Actually, the contractual terms may also be implied. "Contractual terms relating to overtime can be express – i.e. written into the contract of employment – or implied. A term will only usually be implied to permit a change in working hours where it is necessary to give business efficacy or where it can be ascertained through custom in the particular industry or past practice between the parties" (IDS, 1997).

contracts are agreed at the start of the working relationship with subsequent re-negotiation precluded. Modellers have concentrated on questions of contract efficiency, and means of improving efficiency, given the strictures of *a priori* bargaining. Essentially, the problem is one of minimizing sub-optimal separations (i.e. quits and layoffs).³ Our contribution is to extend these models to include working time. Even in the absence of overtime working we show that incorporating hours into bargaining agreements helps potentially to improve efficiency. In other words, even simple hours extensions provide insights as to the potential importance of studying wage-hours contracts. We go on to show how further efficiency gains can be achieved if overtime working is introduced into the total hours arrangements.

Theoretical outcomes are considered against the background of a number of empirical findings that are outlined in section 2. Our approach to the theory itself is to attempt to bring out the underlying intuition behind developments, with formal derivations confined to an appendix. Section 3 compares and contrasts the theory of single wage contracts and simple wage-hours contracts. Section 4 provides the model extensions that embrace overtime work and pay. Section 5 compares the practice of using overtime payments rather than bonuses as a means of improving contract efficiency. Section 6 concludes.

³ Hashimoto (1981) and Carmichael (1983) are among the best known examples, and the ones that are most influential to the developments here. See Malcomson (1999) for an excellent review of this and related work.

2. Empirical Background

With an emphasis on the British labour market, we answer seven empirical questions regarding the practice of overtime working. At later stages, our answers are linked to theoretical findings.

(a) Who works paid overtime?

The incidence of paid overtime is far higher among blue collar compared to white collar workers. For example, Hart (2004) shows that within the British male workforce in 2001, 7.5% of managers and 13.1% of professionals worked paid overtime⁴ in contrast to 45.8% of plant and machine operatives. Respective figures for females are 9.1% and 13.1% compared with 27.7%. From the British company case studies reported in IDS (1997) we know that almost all manual workers are eligible for paid overtime while most companies stop paying for overtime among non-manual staff when they reach specified salary levels and grades.

(b) What proportions of eligible workers work overtime?

For purposes of convenience and simplification, most theoretical models of overtime working have assumed that within overtime firms all employees – usually represented as a homogeneous workforce - work overtime. Based on proportions of British employees within a given occupation who work paid and unpaid overtime, Table

⁴ In the statistical survey used below – i.e. the British Household Panel Survey (BHPS) between 1991 and 2005 – we make the distinction between all male workers and all workers excluding managers, professionals and associate professions. While 33% of all male workers are found to work paid overtime during this period, this rises to 43% when managers, professionals and associate professionals are excluded. In fact, only 13% of the latter group work paid overtime.

1 shows that, in reality, this is generally not the case. Where overtime is worked, there is typically less than complete overtime participation in a given occupation. There are undoubtedly many reasons for this, but the observation at least points to the desirability of deriving modelling outcomes that are consistent with this observed partial incidence.

(c) How high are overtime premiums?

Internationally, premiums paid for overtime hours in excess of standard hours represent substantial incremental increases in basic hourly rates of pay. Of the 27 countries covered by the OECD (1998), one-half reported premiums of 50% or more⁵. In many countries, high premiums result from statutory intervention. The United States with a minimum mandatory premium of 50% provides the best known example. But even in Britain where no legislation applies we know from the company case studies of IDS (1997) that a 50% premium occurs with the greatest frequency. Hart (2004) shows that overtime premiums vary between 30% and 40% when averaged over all British overtime workers. .

(d) To what extent is overtime a requirement of the job?

To what extent is there a permanent or systematic recourse to the use of overtime as opposed to reasons involving temporary contingencies like rush orders, cover for illness and labour shortages? Hart (2004) presents British evidence – based on the Workplace Employees Relations Survey (WERS) – that almost one-quarter of employees who report that they work overtime claim that it a requirement of their job. This points to the likelihood that many firms integrate overtime schedules as a permanent component of

⁵ That is 50% of the standard hourly wage rate so that overtime is rewarded at ‘time-and-a-half’ the standard rate.

their working time requirements. Other WERS responses, cross-tabulated the workers who report that overtime is a job requirement, make it clear that individuals are generally favourably disposed towards their jobs and associated work experiences.

(e) Is there a relationship between paid overtime and job tenure?

We develop wage-hours contracts that emphasise the role of specific human capital investment. A useful empirical backdrop is to establish the connection between the probability of undertaking paid overtime and the two human capital-related Mincer variables, job tenure and work experience. Based on the British Household Panel Survey (BHPS) for the years 1991 to 2005, we estimate a probit equation in which the dependent variable takes the value of 1 if an individual worked paid overtime and 0 otherwise. Explanatory variables include quadratics in job tenure (i.e. length of stay in the current job) and work experience (length of labour market experience since completing full-time education, including length of stay in the current job) as well as individual and time fixed effects. Other control variables are described in Table 2 together with estimated coefficients. Estimation is carried out both excluding and including individual fixed effects. We find that the probability of paid overtime rises in job tenure and declines in work experience. The job tenure result is significant at 5% in the probit excluding individual fixed effects and slightly weaker in the fuller specification. We note, however, that tenure coefficients are virtually unaltered when individual fixed effects are controlled for. We carried out the same regression excluding managers, professionals and associate professionals, because this group work little paid overtime (see footnote 4), but this made no difference to the results.

(f) What is the association between working paid overtime and job separation?

Job separations (quits and layoffs) feature prominently in later developments. What is the relationship between separations and overtime working? More specifically, if an individual works paid overtime in a given period does this affect the probability of a job move in the subsequent period? Again using the BHPS for 1991 to 2005, we estimate a probit equation in which the dependent variable takes the value of 1 if an individual changed job in the current year and 0 otherwise.⁶ Explanatory variables include a binary variable indicating whether the individual worked paid overtime in the previous year, with the remaining variables included matching those of the job tenure regression.⁷ Results are presented in Table 3. For all workers, the estimated coefficient on working paid overtime in the previous year is negative but statistically insignificant. When managers and professional workers are excluded this variable displays a significant negative association. At least in respect of non-managers and non-professionals, working paid overtime reduces the probability of subsequent job move.

(g) What is the relationship between the basic wage rate and the overtime premium?

Based on the British New Earnings Survey, Bell and Hart (2003) and Hart (2004) show that there is a clear negative relationship between basic hourly wage rates (i.e. excluding overtime) and hourly premium rates for overtime hours. It turns out that this relationship comprises an essential aspect of our theoretical predictions.

⁶ A job move in BHPS refers to moves both within and between firms.

⁷ Except tenure which is zero at the point of job change.

3. Wage and Wage-Hours Contracts

Our discussion in this and the following two sections attempts to minimise technical detail and maximise intuitive explanation. At various stages we link developments to a more formal exposition contained in the appendix.

In this section and throughout we discuss wage and wage-hours contracts within a two period framework. Following Carmichael (1983), our wage-hours models differentiate 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 common denominators of each and every model – taken as given as we move from model to model – are as follows. In period 1, the (marginal) worker receives specific training. In period 2 the worker is fully trained and no further training takes place. Retirement occurs at the end of period 2. First and second period wage rates are denoted respectively by w_1 and w_2 , with the subscripts 1 and 2 carrying the same meaning on all variables. The value of the alternative wage (or the outside opportunity) is w_a and the value of marginal product is VMP . Weekly hours in the firm are denoted by h and in alternative employment by h_a .

(a) *Wage contracts*⁸

We start with brief resumes of three seminal contributions in the wage literature, concentrating on aspects that are most important to our subsequent working time extensions.

Becker (1962) argues that the firm and its workers share the returns and costs (the surplus) associated with specific training. Sharing consists of the firm paying for part of the training in period 1 and receiving a return in period 2 through paying $w_2 < VMP_2$. The second-period return to the worker is realised through $w_2 > w_a$. The inequalities $VMP_2 > w_2 > w_a$ discourage the firm from laying off trained workers and encourage workers to participate in period 1 training.

Hashimoto (1981) investigates Becker's sharing arrangements in more depth. He develops the theme that sharing is strongly conditioned by the transaction costs of verifying and communicating information with respect to VMP and w_a (see also Hashimoto and Yu, 1980). The critical problem is that the parties may not be equally knowledgeable about the productivities underlying these two variables, the values of which are not revealed until the start of period 2. Hashimoto assumes that the firm observes VMP and the worker w_a . High transaction costs preclude the exchange of information in period 2. Consequently, the two sides agree to set w_1 and w_2 *ex ante* (i.e. before training begins). Those workers who subsequently find that $w_a > w_2$ quit their jobs while the firm lays off workers for whom it turns out that $w_2 > VMP_2$. Workers who remain receive $w_2 \leq VMP_2$. The key point is that this can lead to inefficient separations. Why? A separation would occur under a first best contract iff $w_a > VMP_2$. So, the actual

⁸ See also Hutchens (1989).

quit/layoff decisions are likely to involve separations taking place when the surplus is positive.

An inefficient quit would occur if

$$(1) \quad VMP_2 > w_a > w_2.$$

The worker quits because of a better outside opportunity but the firm wants the worker to remain because it makes a surplus. But since knowledge of w_a and VMP is asymmetrically held and since the transaction costs of communication and verification are prohibitively high, there is no solution that involves the firm, *ex post*, giving up part of its gain to the worker by enough to avoid the separation.

An inefficient layoff would occur if

$$(2) \quad w_2 > VMP_2 > w_a.$$

Again, the surplus is positive, but this time solely in favour of the worker. The inability to re-negotiate the contract because of problems of credible information exchange leads to an inefficient layoff.

Carmichael (1983) suggests a work and pay arrangement that improves on the efficiency of the Hashimoto model.⁹ He introduces a seniority system for period 2 consisting of type 1 jobs and a fixed number of more senior type 2 jobs. Type 1 and type 2 workers are trained to the same standard in period 1 and are equally productive. Type 1

⁹ Carmichael makes the more realistic assumption that both parties are equally knowledgeable about w_a but that only the worker knows the degree of job satisfaction derived from the current job. For continuity of exposition, and because it makes no substantive difference to outcomes, we stick to Hashimoto's informational assumptions in the main text. In the appendix, we link our model more directly to that of Carmichael (1983) and so we use the idea of workers' private information on job satisfaction. Again, we emphasise that this makes no difference to the key findings.

jobs are remunerated at w_2 and type 2 at $w_2 + S$ where S is a seniority bonus. Promotion to type 2 jobs is related to length of tenure. A newly trained worker is assigned to a type 1 job. As tenure lengthens, the worker eventually reaches the head of the promotion queue of type 1 employees, achieving type 2 status when the next vacancy occurs. The precise timing of promotion is uncertain, occurring sometime around the middle of the second period. In respect of contract efficiency, the critical consequence of this automatic promotion rule is that a layoff can only save the firm w_2 . This outcome means that S provides an additional instrument to w_2 with which to achieve contract efficiency. As pointed out by Hutchens (1989), for S to add value, it must have a different effect on separations from that of w_2 . Carmichael shows that an increase in w_2 reduces quits and raises layoffs while an increase in S reduces quits (the incentive to wait to receive a wage greater than VMP_2) but does not affect layoffs (the firm can only save w_2 if a worker is laid off.)

Carmichael shows that

$$(3) \quad w_2 + S > VMP_2 > w_2,$$

that is at least some of the type 2 workers are paid above their marginal products and type 1 workers below their marginal products. What accounts for these inequalities? Under the bonus scheme, both parties have an incentive to agree w_2 such that $VMP_2 > w_2 > w_a$, thereby reducing the inefficient layoffs as represented by inequality (2) in the Hashimoto model. They would agree to this because (i) this improves the incentive for them to stay together and (ii) a relatively low wage can be compensated by a high bonus. This accounts for the second inequality in (3), but what accounts for the first, i.e. a level of seniority pay above marginal product? Suppose there are N_2 workers in period 2 of which

N_B receive a bonus. As long as the firm has a positive surplus, that is $N_2.VMP_2 > (w_2 + S).N_B + w_2.(N_2 - N_B)$, high remuneration due to seniority ($w_2 + S$), even above VMP_2 , can reduce inefficient quits as represented by inequality (1) in the Hashimoto model. This is due to the fact that the expected *ex ante* wage income of a marginal worker is given by $\omega = w_2 + S.N_B/N_2$. A high S would make it more likely that $\omega > w_a$, which offsets the low w_2 .

Under his seniority bonus scheme, Carmichael's compensating rule (3) reduces both the inefficient quit represented by (1) and the inefficient layoff in (2) highlighted by Hashimoto in his model.

(b) Wage-hours contracts

Concentrating on the framework of Carmichael (1983), we now introduce working hours into the picture and show that hours matter in these human capital models.¹⁰ Early pointers are provided in the wage-hours labour demand literature (Brechling, 1965; Ehrenberg, 1971). In labour demand models, a rise in initial training investment induces the firm to increase working hours since investment amortisation is improved both by longer tenure among trained workers and more intensive labour input for given tenure.

The inclusion of working time necessarily alters the representation of the worker's pay and marginal product. Pay is now expressed in terms of weekly earnings net of the disutility of providing weekly hours; that is $y = w.h - d(h)$. As for marginal product, we recognise that it may be functionally related to the length of weekly hours. In fact, we

¹⁰ Formal developments are given in appendix.

would expect that the first derivative of $VMP(h)$ with respect to hours to be $VMP'(h) \leq 0$. Typically, the working time literature assumes VMP is declining in hours due to such influences as fatigue and boredom. However, VMP that is independent of hours changes may not be uncommon in work environments where weekly hours are relatively short or where working time is systematically punctuated by rest periods or where performance monitoring is prevalent. For simplicity, and without losing the essential features of our hours' modelling extensions, we focus on hours-invariant VMP in the main text. In the appendix, we indicate how the results are modified when hours-related VMP is considered.¹¹

Second-period weekly earnings are given by $w_2 \cdot h_2$.¹² Suppose initially that weekly hours are exogenously determined. For example, the firm might adopt the customarily accepted normal hours of the industry to which it belongs. Workers quit if $w_a \cdot h_a - d(h_a) > w_2 \cdot h_2 - d(h_2)$. The firm lays off workers if $w_2 \cdot h_2 > VMP_2 \cdot h_2$. Adopting the same private information assumptions as before, inefficient separations are likely to occur because the first-best separation rule is given by $w_a \cdot h_a - d(h_a) > VMP_2 \cdot h_2 - d(h_2)$. In line with the arguments surrounding inequality (1), an inefficient quit would occur if

$$(4) \quad VMP_2 \cdot h_2 - d(h_2) > w_a \cdot h_a - d(h_a) > w_2 \cdot h_2 - d(h_2)$$

¹¹ In appendix section (i) – (iv), we fully develop the case where hourly productivity VMP is assumed to be (working) hours-invariant. In appendix section (v), we explain why the introduction of hours-related VMP does not qualitatively change the conclusions reached in the simpler set-up.

¹² Assumption concerning first-period hours, marginal product and training cost are outlined in appendix section (i).

or a worker would quit the firm despite a positive surplus. In line with inequality (2), an inefficient layoff would occur if

$$(5) \quad w_2.h_2 - d(h_2) > VMP_2.h_2 - d(h_2) > w_a.h_a - d(h_a).$$

or the firm would fire the worker despite a positive surplus.

What if the parties were to move away from using exogenously determined hours?

As long as the worker's return $y_2 = w_2.h_2 - d(h_2)$ increases with h_2 [i.e. $w_2 > d'(h_2)$], then longer hours increase the return and hence induce a greater incentive for the worker to stay. At the margin, assuming $w_2 > d'(h_2)$, workers for whom

$$(6) \quad w_a.h_a - d(h_a) > w_2.h_2 - d(h_2)$$

held before the increase in h_2 would now be induced to stay by a reversal of this inequality. As for the firm, increasing h_2 involves a cost (weekly earnings are increased) and a gain (weekly marginal product is increased). As long as $VMP_2.h_2 \geq w_2.h_2$ the firm has no incentive to fire. In fact, under the assumption that VMP is hours-invariant, i.e., $VMP'(h_2) \equiv 0$, a change in h_2 has no effect on layoffs.¹³

What is the stopping rule for the h_2 increase? It is undertaken until $y_2 = w_2.h_2 - d(h_2)$ is maximized for the marginal worker subject to the constraint that $w_2.h_2 \leq VMP_2.h_2$. Let the optimal hours for this worker be denoted h_o . If hours are too long, or $h_2 > h_o$, this would reduce $y_2 = w_2.h_2 - d(h_2)$ and we would go back to the inequality in (6) thereby inducing separation.

¹³ See appendix section (v) for a relaxation of this assumption.

How does this wage-hours specification compare with Carmichael's wage model incorporating a seniority bonus? There is one similarity. The hours variable provides a second potential instrument to help effect efficient separations. Conditional on $w_2.h_2 \leq VMP_2.h_2$, it may act as an incentive for workers to stay with the firm without strongly affecting the firm's own layoff decision.¹⁴ There are three differences between the hours and bonus mechanisms. First, the hours instrument applies to all trained workers. Second and related, in the (w_2, h_o) hours contract, the cost of retaining a (marginal) worker in period two is exactly the same as the pay of a marginal worker. These two costs are different in Carmichael's model. Third, Carmichael's automatic compensation rule shown by (3) effectively reduces both inefficient quits and layoffs. But our simple wage-hours contract only reduces inefficient quits, it may well not reduce inefficient layoffs. The second and third of these differences are important because they point to the possibility that there may be room for further efficiency improvements. This is where the use of overtime hours becomes relevant.

4. Overtime hours and premium pay¹⁵

Suppose that the parties are operating under the above (w_2, h_o) hours contract. This does not rule out the possibility that $VMP_2 > w_2$, in which case the firm would prefer longer hours $h_2 > h_o$. This possibility is precluded in the contract as it stands because hours in excess of h_o would reduce y_2 , or

¹⁴ If VMP_2 is hours-invariant, the firm's layoff decision is strictly unaffected and h_2 provides an especially effective instrument.

¹⁵ See appendix section (iv) for formal developments.

$$(7) \quad w_2.h_2 - d(h_2) < w_2.h_o - d(h_o), \text{ for } h_2 > h_o$$

implying that the probability of quitting is increased thereby triggering more inefficient separations.

One possibility of compensating the fall of y_2 for a rise in h_2 beyond h_o , is for the firm to offer overtime pay $k.w_2$ for these marginal hours such that $w_2.h_o + k.w_2.(h_2 - h_o) - d(h_2) > w_2.h_o - d(h_o)$. Using inequality (7), this implies that the overtime premium k would be set such that

$$(8) \quad k > \frac{d(h_2) - d(h_o)}{w_2.(h_2 - h_o)} > 1.$$

The firm must pay an overtime premium $k > 1$ to compensate the worker for the disutility of ‘involuntary’ long hours.

Should the firm pay all, equally productive, trained workers the same per-person overtime hours at a premium $k_2.w_2$ ($k_2 > 1$) such that the gap $h_2 - h_o$ is filled? This is problematic because it would increase marginal pay and hence increase the probability of layoffs. A superior outcome is suggested by Carmichael’s second period two-tier bonus system. In terms of overtime, this translates into guaranteeing a fixed number of trained and longer tenured workers additional overtime hours at a premium rate. A junior trained worker waits in a queue until his turn arrives to work the guaranteed overtime. The firm’s marginal hourly cost $k.w_2$ while the marginal hourly replacement cost is w_2 . As in Carmichael, efficiency is gained because the cost of retaining a marginal worker differs from the pay of the marginal worker.

As with Carmichael's bonus arrangement, is there an incentive for workers to want to work paid overtime? In other words, does this overtime pay scheme also have an automatic compensating rule

$$(9) \quad k.w_2 > VMP_2 > w_2$$

that reduces both inefficient quits as in (4) and inefficient layoffs as in (5) that occur in the simple wage-hours contract? The answer is yes.

Both the firm and the worker have incentive to lower w_2 to a level such that $VMP_2.h_2 - d(h_2) > w_2.h_2 - d(h_2) > w_a.h_a - d(h_a)$, in order to reduce an inefficient layoff under (5). The lower wage w_2 is then automatically compensated by an overtime premium $k > 1$ when a senior worker quits or is fired. This explains the second inequality in (9). Then why is the firm willing to pay a long-tenured workers an overtime premium k such that $k.w_2 > VMP_2$? This is because as long as the firm has a positive surplus, i.e.,

$$VMP_2 [N_2.h_o + N_P (h_2 - h_o)] > w_2 [N_2.h_o + k.N_P (h_2 - h_o)],$$

where N_p is the number of senior workers working overtime for premium pay, a high overtime pay $k.w_2$, even above the VMP_2 , can reduce inefficient quits as represented by (4). This is due to the fact that the expected wage income, net of the expected disutility of hours, of a marginal worker is given by:

$$\begin{aligned} y_2 &= (1-N_P/N_2)[w_2.h_o - d(h_o)] + (N_P/N_2) [w_2.h_o + k.w_2.(h_2 - h_o) - d(h_2)] \\ &= w_2.h_o + (N_P/N_2).k.w_2.(h_2-h_o) - (1-N_P/N_2).d(h_o) - (N_P/N_2).d(h_2). \end{aligned}$$

A high $k.w_2$ would make it more likely that $y_2 > w_a.h_a - d(h_a)$, which offsets a low w_2 .

This explains the first inequality in (9).

Figure 1 illustrates the overtime pay schematic resulting from these developments. It is consistent with the evidence produced in Section 2. First, and generally, we expect firms that systematically and consistently make use of overtime to be most likely to be involved in this contractual arrangement because they are guaranteeing overtime to a fixed number of senior workers (i.e. workers with longer tenure). We know from Section 2 (d) that a significant proportion of workers indicate that overtime working is a job requirement. Second, from Section 2 (a) and Table 1 we know that, where overtime is worked, it is typical that not all workers in a given occupation are overtime workers.¹⁶ Third, since we argue that workers with longer tenure are more likely to work overtime, the model is consistent with the findings in Section 2 (e) and Table 2 that overtime working rises in job tenure. Fourth, while of course not conclusive evidence, it would be unsurprising if premiums that typically represent between a 30% and 50% mark-up of basic rates, as discussed in 2 (c), are found to be above marginal product. In fact, we provide another piece of evidence that is consistent with high returns to paid overtime. If paid overtime is rewarded at above marginal product then we would expect overtime workers would exhibit relatively low probabilities of leaving their current jobs. Results reported in Section 2 (f) and Table 3 provide some support for this expectation.

As reported in Section 2 (g), U.K. empirical work has established a negative relationship between the basic hourly wage rate and the hourly overtime premium rate of

¹⁶ Moreover, the fractions in Table 1 that do not work overtime are generally too large to be accounted for by workers in their early tenure who are undergoing training.

pay. This is consistent with the compensating rule of our wage-hours contract. We have established that, embedded in the contract solution [see (A23) in appendix], we have

$$(10) \quad \frac{\partial w_2}{\partial k} < 0$$

or there is an inverse relationship between the contractual wage and the overtime premium. Lowering w_2 increases profit to the firm but also increases the probability of the worker quitting. Hence the wage stopping rule is where the marginal profit to the firm equals to marginal loss of an extra unit reduction of w_2 . Similarly, an increase of k reduces the profit to the firm but increases the probability of the worker staying, which in turn enhances the firm's profit. Hence the premium stopping rule is where the marginal loss of the firm equals to marginal profit of an extra unit increase of k .

This wage-premium trade-off is an especially important outcome since it links to a wider theoretical and empirical literature. Based on the seminal paper of Lewis (1969), the theory of hedonic wages also establishes this negative wage-premium relationship. Essentially, the parties agree optimal compensation packages based the worker's objective of finding earnings/hours combinations that maximise utility and the firm's profit maximising motivation that establishes optimal workers/hours combinations.¹⁷ In an important policy application, Trejo (1991) shows that if an outside agent (i.e. the government) were to increase the size of k by mandate then the parties would simply agree to decrease w_2 so as to leave their agreed compensation package intact. Attempts to increase employment on the extensive margin by imposing more costly overtime on the

¹⁷ The best source for theoretical developments is Kinoshita (1987).

intensive margin are essentially negated by such an automatic adjustment reaction.¹⁸ Our contract model provides an alternative theoretical approach that provides the same mechanism.

5. An Overtime Premium or a Bonus?

The foregoing overtime premium arrangement has the same mechanism as the seniority bonus S of Carmichael. An increase in k reduces quits (the incentive to wait to receive a wage that is greater than w_2) but – due to the delayed overtime eligibility assumption - does not affect layoffs. In fact the two schemes are mathematically equivalent. For example, a firm may introduce a seniority bonus S , instead of k , to achieve the same objective. That is, a senior worker may receive a package of $w_2.h_2 + S = w_2.h_o + k.w_2.(h_2-h_o)$ as an equivalent contract (w_2, h_2, S) , where $h_2 > h_o$ is specified in the contract and S is not explicitly linked to hours. However, at least for the class of workers who typically work overtime, the use of overtime premiums rather than bonuses to reward longer tenured workers has two very strong advantages.

The first advantage relates to the type of worker who works paid overtime. As we have seen in Section 2 (a), paid overtime is typically related to blue-collar work. Such workers often work alongside colleagues who possess the same or very close skills.¹⁹ Even under *a priori* contractual agreements, paying a bonus for no extra effort to more senior workers who are equally trained may well be deemed by management to lead to

¹⁸ Trejo (1991) presents U.S. evidence for such a reaction. See also Bell and Hart (2003) for U.K. evidence.

¹⁹ Of course, such blue-collar workers can also possess significant firm specific human and organisational capital that is not easily transferable to other work environments.

potentially serious shop floor industrial relations problems. Essentially, reward and effort are relatively more visible within blue collar work environments than in more management-related occupations.²⁰ Where it is clear that higher pay is linked to more effort in the form of longer per-period hours then there will be a greater perception of fairness.

The second advantage relates to the earlier human capital literature that includes hours as a choice variable. The parties invest in specific human capital in period 1. It is in their interest to maximise the returns to such investment in period 2. The relevant extensive market action to this end involves encouraging longer tenure. The comparable intensive margin action involves encouraging longer per-period hours. A delayed overtime payment system for senior trained workers provides a scheme that is *transparently* consistent with this latter objective.

6. Conclusions

Overtime working is an important consideration in labour market economics and macroeconomics because for many workers it represents the marginal cost of labour input. There have been very few attempts to provide an economic rationale for the use of overtime hours and associated premium pay. In fact, the assumption of ‘custom and practice’ is probably the most prevalent rationale. Here, we offer explanations based on an important earlier wage contract literature that emphasises specific human capital and asymmetric information. Our wage-hours model allows for changes in labour inputs on

²⁰ Those engaged in managerial and/or professional occupations are likely to undertake more complex and multi-faceted job tasks that would tend to be individual-specific and less widely understood by work colleagues. Bonus payments may therefore be a less contentious reward in respect of seniority.

the extensive (stock of workers) and intensive (working hours) margins. 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. Even without the use of overtime, we illustrate how jointly bargaining over the hours of trained workers can enhance contract efficiency. If firms would prefer even longer hours in order to enhance the firm's surplus, we show how an overtime premium schedule could optimally be brought into play. It is in the parties' joint interests to guarantee overtime work and premium pay to relatively senior trained workers. We show that the optimal pay configuration is to remunerate the basic hours of trained workers at a rate below marginal product and their overtime pay at above marginal product.

References

- 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.
- Brechling, F. "The relationship between output and employment in British manufacturing industries", *Review of Economic Studies* 32, (1965), 187–216.
- 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.
- 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. "Firm-specific human capital as a shared investment", (1981) *American Economic Review* 71, 475-82.
- Hashimoto, M and B T Yu, "Specific capital, employment contracts and wage rigidity", *Bell Journal of Economics*, (1980), 11, 536-49.
- Hutchens, R M. "Seniority, Wages and Productivity: A Turbulent Decade", (1989) *Journal of Economic Perspectives* 3, 49-64.
- Income Data Services. *Overtime*, (1997) Study 617, London: Income Data Services Ltd.
- Kahn, S and K Lang, "Constraints on the choice of hours", *Journal of Human Resources* 27, (1992), 661 - 678.
- Kinoshita T, 'Working hours and hedonic wages in the market equilibrium', (1987). *Journal of Political Economy* 95, 1262-1277.
- Lazear, E P. "Agency, earnings profiles, productivity, and hours restrictions", *American Economic Review* 71 (1981): 606-20.
- Lewis, H.G. "Employer interests in employee hours of work", (1969), mimeo, University of Chicago.
- Malcomson, J M. "Individual employment contracts", *Handbook of Labor Economics*, volume 3B (eds. Orley Ashenfelter and David Card), (1999), Elsevier Science.

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.

Appendix: An extended Carmichael model with overtime pay

We present modelling developments that lie behind the discussion in sections 2, 3 and 4. Essentially, we extend the contract model of Carmichael (1983) to incorporate working hours and overtime pay.

(i) Underlying framework

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 $VMP_1 = w_a - C/h_1$, where h_1 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 $E(VMP_2) = w_a + M$, where VMP_2 and M are both assumed to be hour-invariant for simplicity. We relax this restriction in section (v).

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 $VMP_2 = w_a + M + \eta$. 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

$$Q = Q(\theta^*) = \int_{-\infty}^{\theta^*} q(\theta) d\theta \quad (A1)$$

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

$$F = F(\eta^*) = \int_{-\infty}^{\eta^*} f(\eta) d\eta \quad (A2)$$

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_1 and h_2 weekly hours in periods 1 and 2 respectively, with the corresponding disutilities represented by $d(h_1)$ and $d(h_2)$. For simplicity, we assume that weekly hours in period 1 are fixed to $h_1 = h_a$, where h_a is alternative employment working hours. But the number of weekly hours in period 2 is a choice variable h_2 .

The parties' joint wealth consists of the returns arising from three mutually exclusive and exhaustive events, weighted by the probability of their occurrence. 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 ($w_a \cdot h_a - C - d(h_a)$). If the worker is fired

or voluntarily quits, the second period surplus to the worker is given by the market value $w_a.h_a - d(h_a)$; 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:

$$\begin{aligned}
 W = & F.[w_a.h_a - C - d(h_a) + w_a.h_a - d(h_a)] \quad (\text{the worker is fired}) \\
 & + (1-F).Q.[w_a.h_a - C - d(h_a) + w_a.h_a - d(h_a)] \quad (\text{the worker quits}) \\
 & + (1-F).(1-Q).\{w_a.h_a - C - d(h_a) + h_2[w_a + M + E(\eta|\eta > \eta^*) + E(\theta|\theta > \theta^*)] - d(h_2)\} \\
 & \quad (\text{the worker stays}). \quad (A3)
 \end{aligned}$$

(ii) The first-best solution without the problem of asymmetric information

Similar to Carmichael (1983), the first-best solution without the problem of asymmetric information may be derived by simply choosing a (θ^*, η^*, h_2) triplet to maximize the joint wealth W in (3). This gives the following first-best solution (the details are available upon request).

A worker quits if the job satisfaction θ is too low:

$$\theta < \theta^* = - [w_a (h_2 - h_a)/h_2 + M + E(\eta|\eta > \eta^*) + d(h_a)/h_2 - d(h_2)/h_2] \quad (A4)$$

The firm fires a worker if the realised hourly productivity in period 2 is too low:

$$\eta < \eta^* = - [w_a (h_2 - h_a)/h_2 + M + E(\theta|\theta > \theta^*) + d(h_a)/h_2 - d(h_2)/h_2] \quad (A5)$$

and the optimal working hours h_2 in period 2 are determined by

$$d'(h_2) = w_a + M + E(\eta|\eta > \eta^*) + E(\theta|\theta > \theta^*) \quad (A6)$$

These conditions imply that the first-best solution can only be achieved if the party wishing to separate is made to internalise the entire expected losses from the separation.

(iii) Asymmetric information and the second-best solution

Now suppose that information concerning job satisfaction and productivity cannot be exchanged *ex post*. Then the firm and its workers will determine the separation rules of fire and quit unilaterally. The joint wealth W in (3) is maximised subject to the constraints of the two separation rules. That is, the wage contract w_2 is offered to ensure that the firm will fire the workers whenever productivity is too low

$$\eta < \eta^* = -(w_a + M - w_2). \quad (A7)$$

Equivalently, workers will quit whenever job satisfaction is too low; i.e.

$$\theta < \theta^* = -w_2 + [w_a \cdot h_a - d(h_a) + d(h_2)]/h_2. \quad (A8)$$

This indicates that there are only two - instead of three in Section (ii) - choice variables (w_2 , h_2) to maximize the joint wealth. Therefore, the solution under asymmetric information is a second-best solution. (Full details of the solution are available upon request).

However, given w_2 , we find that the probability of a marginal worker quitting (Q) is negatively related to working hours in period 2 (h_2), that is,

$$\frac{\partial Q}{\partial h_2} = \left(\frac{\partial Q}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial h_2} \right) = q(\theta^*) \frac{h_2 d'(h_2) - d(h_2)}{h_2^2} < 0 \quad (A9)$$

if the disutility is not too large; that is if the elasticity of disutility with respect to hours e_h is less than unity, or

$$e_h = \frac{h_2 d'(h_2)}{d(h_2)} < 1.$$

We also find that the probability of firing (F) is independent of working hours in period 2 (h_2), i.e.

$$\frac{\partial F}{\partial h_2} = \left(\frac{\partial F}{\partial \eta^*} \right) \left(\frac{\partial \eta^*}{\partial h_2} \right) = 0$$

This implies that longer hours would induce workers to stay as the wage income rises with the hours, but longer hours do not affect the probability of fire. Hence, longer hours may improve the efficiency of the second-best solution if the elasticity of disutility with respect to hours e_h is less than unity.

(iv) Overtime pay and contractual efficiency

Suppose that the optimized period 2 working hours from Section (iii) above are given as h_o . In this section, we ask the following question. If the firm pays a pre-announced fixed-number of trained and longer tenured workers (N_p) additional overtime hours at a premium rate $k.w_2$ (with $k>1$) to induce them to work extra α hours on top of h_o basic hours, will this further improve contract efficiency? Under this arrangement, a junior trained worker waits in a queue until his turn arrives to work the guaranteed overtime. The firm's marginal hourly cost is $k.w_2$ while the marginal hourly replacement cost is w_2 . As in Carmichael (1983), efficiency may be gained because the cost of retaining a marginal worker differs from the pay of the marginal worker.

Similar to the design of Carmichael (1983, p.254), in our overtime premium scheme, workers are promoted (or tenured) sometime in the middle of their second

period. Therefore, at the beginning of the period 2, the *ex ante* expected joint wealth W_{ot} between the firm and a worker becomes:

$$\begin{aligned}
 W_{ot} = & F.[w_a.h_a - C - d(h_a) + w_a.h_a - d(h_a)] \quad (\text{the worker is fired}) \\
 & + (1-F).Q.[w_a.h_a - C - d(h_a) + w_a.h_a - d(h_a)] \quad (\text{the worker quits}) \\
 & + (1-F).(1-Q). \{ w_a.h_a - C - d(h_a) \\
 & + (h_o + \alpha\rho)[w_a + M + E(\eta|\eta > \eta^*) + E(\theta|\theta > \theta^*)] - (1-\rho)d(h_o) - \rho d(h_o + \alpha) \} \\
 & (\text{the worker stays}). \tag{A10}
 \end{aligned}$$

where $\rho = N_p/N_2$ is the pre-announced fraction of longer-serving workers who work overtime over the total number of workers in period 2, α is the number of overtime hours.

If $\alpha = 0$ and hence $\rho = 0$, then joint wealth W_{ot} collapses to that in Section (iii). However, if $\alpha > 0$ for long-serving workers, would that increase W_{ot} and hence improve the efficiency of the contract? The answer is yes. This is due to the fact that the expected wage income, net of the expected disutility of hours, of a marginal worker in period 2 is given by:

$$y_2 = w_2.h_o + (1 - \rho)[w_2.h_o - d(h_o)] + \rho[w_2.h_o + k.w_2.\alpha - d(h_o + \alpha)] \tag{A11}$$

A worker would quit if

$$\theta < \theta^* = \frac{w_a.h_a - y_2}{h_o + \rho\alpha}. \tag{A12}$$

Hence, there is a negative relationship between the probability of quit (Q) and overtime hours (α), that is

$$\begin{aligned}
 \frac{\partial Q}{\partial \alpha} &= \left(\frac{\partial Q}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial \alpha} \right) \\
 &= -q(\theta^*)\rho[w_a.h_a + w_2.h_o(k-1) + (1-\rho)d(h_o) + \rho d(h_o + \alpha) - (h_o + \rho\alpha)d'(h_o + \alpha)] / (h_o + \rho\alpha)^2 < 0.
 \end{aligned}$$

(A13)

There is also a negative relationship between the probability of quit (Q) and proportion of overtime workers (ρ), that is

$$\begin{aligned} \frac{\partial Q}{\partial \rho} &= \left(\frac{\partial Q}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial \rho} \right) \\ &= -q(\theta^*) [w_a h_a + w_2 h_o (k-1) \alpha + (h_o + \alpha) d(h_o) - h_o d(h_o + \alpha)] / (h_o + \rho \alpha)^2 < 0. \end{aligned} \quad (A14)$$

Both inequalities of (A13) and (A14) would hold if the elasticity of disutility with respect to hours e_h is less than unity and overtime hours are not too long (the full details are available upon request).

The probability of quit (Q) is also negatively related to the overtime premium (k), that is

$$\frac{\partial Q}{\partial k} = \left(\frac{\partial Q}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial k} \right) = -q(\theta^*) \frac{\rho w_2 \alpha}{(h_o + \rho \alpha)^2}. \quad (A15)$$

Taking inequalities (A13) to (A15) together, imply that a package consisting of an overtime premium ($k > 1$) offered to a group of longer-serving workers ($\rho > 0$) for overtime hours ($\alpha > 0$) would induce a marginal worker to stay.

On the other hand, the firing decision (A7) is unaffected by the overtime package since $\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial \rho} = \frac{\partial F}{\partial k} = 0$. This is due to the fact that there are a fixed number of ‘with overtime’ posts and automatic promotion from the ‘without overtime’ pool when there is a vacancy. Hence, a layoff can only save the firm $w_2 h_o$ under this system (this mechanism is similar to that in Carmichael (1983)).

Furthermore, substituting the fire-quit constraints (A7) and (A12) into the joint wealth W_{ot} , we have:

$$\frac{\partial W_{ot}}{\partial \alpha} = \left(\frac{\partial W_{ot}}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial \alpha} \right) > 0 \quad (\text{A16})$$

$$\frac{\partial W_{ot}}{\partial \rho} = \left(\frac{\partial W_{ot}}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial \rho} \right) > 0 \quad (\text{A17})$$

$$\frac{\partial W_{ot}}{\partial k} = \left(\frac{\partial W_{ot}}{\partial \theta^*} \right) \left(\frac{\partial \theta^*}{\partial k} \right) > 0 \quad (\text{A18})$$

where

$$\frac{\partial W_{ot}}{\partial \theta^*} = - (1-F)q(\theta^*) \{ d(h_a) + h_o E(VMP_2) + \rho \alpha [E(VMP_2) - k w_2] \} < 0 \quad (\text{A19})$$

$$E(VMP_2) = w_a + M + E(\eta | \eta > \eta^*).$$

The inequality (A19) will hold if the overtime premium k is not too high.

Inequalities (A16) to (A18) together reveal that a scheme of overtime premium ($k > 1$) offered to a group of longer-serving workers ($\rho > 0$) for overtime hours ($\alpha > 0$) could indeed improve the contract efficiency.

Next we derive the automatic compensating rule for the overtime pay scheme. As long as firing is avoided, we have:

$$\eta \geq \eta^* = -(w_a + M - w_2).$$

This implies

$$VMP_2 = w_a + M + \eta \geq w_2. \quad (\text{A20})$$

Furthermore,

$$\frac{\partial F}{\partial w_2} = \left(\frac{\partial F}{\partial \eta^*} \right) \left(\frac{\partial \eta^*}{\partial w_2} \right) = f(\eta^*) > 0 \quad (\text{A21})$$

which implies that a low wage rate w_2 would reduce the probability of fire. Therefore, it is more likely that the outcome $VMP_2 > w_2$ will be realised, which is the second inequality of the automatic compensating rule (9) in the main text.

The rational of the first inequality in (9) is due to the fact that both $\frac{\partial Q}{\partial k} < 0$ (A15) and $\frac{\partial W_{ot}}{\partial k} > 0$ (A18). Hence high overtime premium pay $k.w_2$, even above VMP_2 , would reduce the probability of quit and improve the efficiency of the contract.

Finally, substituting the fire-quit constraints (A7) and (A12) into the joint wealth W_{ot} , we have:

$$\frac{\partial W_{ot}}{\partial w_2} = \left(\frac{\partial W_{ot}}{\partial \eta^*} \right) \left(\frac{\partial \eta^*}{\partial w_2} \right) < 0 \quad (A22)$$

where

$$\frac{\partial W_{ot}}{\partial \eta^*} = - (1-Q)f(\eta^*) \{ (h_o + \alpha\rho)[w_2 + E(\theta|\theta > \theta^*)] - (1-\rho)d(h_o) - \rho d(h_o + \alpha) \} < 0$$

$$\frac{\partial \eta^*}{\partial w_2} = 1 > 0.$$

From (A18) and (A22), we have

$$\frac{\partial w_2}{\partial k} = \frac{\frac{\partial W_{ot}}{\partial k}}{\frac{\partial W_{ot}}{\partial w_2}} < 0. \quad (A23)$$

This proves the inequality (10) in the main text.

(v) The case of hours-variant VMP

Here we relax the assumption that the value of marginal product VMP is constant and invariant with respect to working hours. We show that this is an innocuous assumption that would not affect our conclusion qualitatively.

Suppose that the *hourly* value of marginal product for a marginal worker in period 2 is $VMP_2 = VMP(h_2)$, where h_2 is working hours in period 2, and $VMP'(h_2) < 0$ due to such influences as fatigue and boredom. The *total* value of marginal product for a marginal worker in period 2 is given as $TVMP(h_2) = \int_0^{h_2} VMP(h)dh$, with $TVMP'(h_2) = VMP(h_2) > 0$, and $TVMP''(h_2) = VMP'(h_2) < 0$.

In our theoretical model we have $VMP_2 = w_a + M + \eta$. For simplicity, we maintain the assumption that both the value of the alternative wage w_a and the random shock η are hours-invariant. Hence, an hours-variant VMP implies that the *hourly* specific human capital M after training for a marginal worker in period 2 is also hours-variant - i.e. $M(h_2)$ with $M'(h_2) < 0$. The *total* specific human capital TM after training for a marginal worker is given by

$$TM(h_2) = \int_0^{h_2} M(h)dh, \text{ with } TM'(h_2) = M(h_2) > 0, \text{ and } TM''(h_2) = M'(h_2) < 0.$$

The firm's firing rule (A7) then changes to:

$$\eta < \eta^* = - \left[w_a + \frac{1}{h_2} \int_0^{h_2} M(h)dh - w_2 \right]. \quad (A24)$$

One of the important implications of hours-variant VMP is that the firing rule is no longer independent of hours. In other words, working hours now have an influence on the firm's firing decision.

However, the relationship between the probability of fire (F) and working hours in period 2 is negative:

$$\frac{\partial F}{\partial h_2} = \frac{\partial F}{\partial \eta^*} \frac{\partial \eta^*}{\partial h_2} = \frac{f(\eta^*)}{h_2^2} \left[\int_0^{h_2} M(h) dh - h_2 M(h_2) \right] < 0 \quad (\text{A25})$$

if the effect of fatigue is not too large, i.e., the elasticity of total specific human capital (TM) with respect to hours $e_{TM,h}$ is greater than unity:

$$e_{TM,h} = \frac{h_2 TM'(h_2)}{TM(h_2)} = \frac{h_2 M(h_2)}{\int_0^{h_2} M(h) dh} > 1.$$

This implies that the hours-variant VMP simply introduces an additional constraint on working hours. Hours should not be excessively long so as to reduce VMP substantially due to the fatigue. Subject to this constraint $e_{TM,h} > 1$, longer working hours would reduce the probability of firing and therefore could still improve the contract efficiency.

On the other hand, hours-variant VMP also affects the inequality (A19). If the expected VMP_2 is lowered due to fatigue, then it will place an additional constraint on the size of the overtime premium k . If this constraint is fulfilled, then an overtime premium could still improve the contract efficiency.

To conclude, an hours-variant VMP imposes additional constraints on the length of working hours and the magnitude of the overtime premium. If these constraints are satisfied, hours-variant VMP would not change the main results of our model.

a

Figure 1 Second-period hourly pay profile

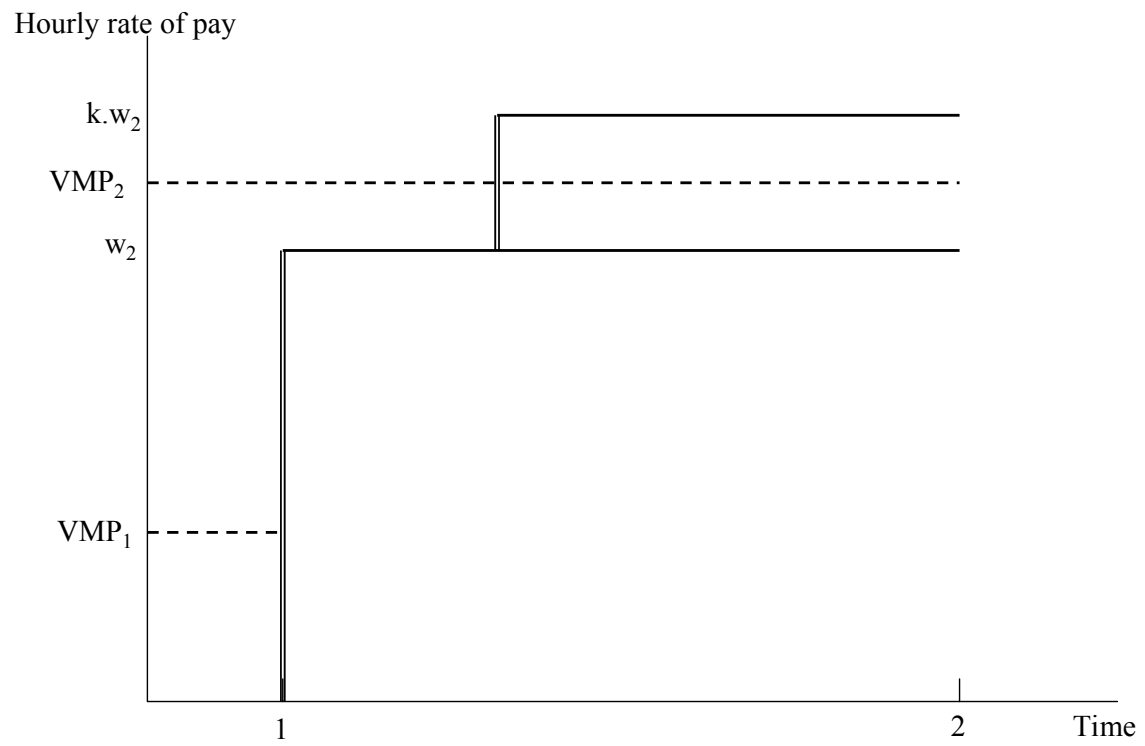


Table 1 Proportions of British employees who regularly work overtime or hours in excess of normal working hours, by largest occupational group within the establishment (paid and unpaid overtime hours)

	Weighted % (based on 2295 establishments)
1. All (100%)	7.8
2. Almost all (80-99%)	7.9
3. Most (60-79%)	10.4
4. Around half (40-59%)	15.2
5. Some (20-39%)	20.4
6. Just a few (1-19%)	22.5
7. None (0%)	13.7
8. Other*	2.0

Source: Workplace Employee Relations Survey (WERS) (Management Survey), 2003.

* Including refusal to give information and don't know.

Table 2 Probability of working paid overtime: male workers (BHPS 1991-2005)

Explanatory Variables	Probit	Probit (with individual fixed effects)
TENURE	0.007* (0.003)	0.007 (0.005)
(TENURE) ² /100	0.003 (0.01)	0.009 (0.022)
EXPERIENCE	-0.009* (0.002)	-0.009* (0.005)
(EXPERIENCE) ² /100	-0.076 (0.004)	-0.026* (0.009)
BELONG TO UNION	0.261* (0.015)	0.308* (0.030)
COHABITING	0.042* (0.018)	0.035 (0.035)
AGE OF YOUNGEST CHILD	-0.003* (0.002)	-0.007* (0.003)
Constant	-1.228* (0.039)	-1.813* (0.078)
Other Controls [§]	Yes	Yes
Sample size	37,678	37,678

Notes: Bracketed figures are standard errors and * denotes 5% significance.

§ Other controls are education dummies (covering six levels of education from university degree-level to legal minimum years of schooling) and year dummies

Table 3 Probability of Job Separation and Working Paid Overtime: male workers (BHPS 1991-2005)

Explanatory Variables	All workers	All workers excluding managers, professionals, and associate professionals
WORKED PAID OVERTIME IN PREVIOUS YEAR	-0.016 (0.022)	-0.067* (0.003)
EXPERIENCE	-0.028* (0.004)	-0.028* (0.004)
(EXPERIENCE) ² /100	0.008 (0.007)	0.008 (0.009)
BELONG TO UNION	-0.305* (0.024)	-0.415* (0.031)
COHABITING	0.047 (0.027)	0.072* (0.035)
AGE OF YOUNGEST CHILD	-0.003 (0.003)	-0.005 (0.003)
Other Controls [§]	Yes	Yes
Sample size	28,386	17,790

Notes: Bracketed figures are standard errors and * denotes 5% significance.

§ Other controls are education dummies (covering six levels of education from university degree-level to legal minimum years of schooling) and year dummies.