Limited Commitment Models of the Labour Market

Jonathan P. Thomas (University of Edinburgh)

Tim Worrall (Keele University)

Version: May 2007

Abstract

We present an overview of models of self-enforcing labour contracts in which risk-sharing is the dominant motive for contractual solutions. A base model is developed which is sufficiently general to encompass the two-agent problem central to most of the literature. We consider two-sided limited commitment and look at its implications for macroeconomics; we then consider what empirical support exists for the model. Subsequently we look at the one-sided limited commitment problem for which there exists a considerable amount of empirical testing.

Keywords: Labour contracts, self-enforcing contracts, business cycle, unemployment.

JEL Codes: E32, J41.

[0]We gratefully acknowledge the financial support of the Economic and Social Research Council (Research Grant:RES-000-23-0865).

1 Introduction

In this paper we consider long-term labour contracts under limited commitment. Firms and workers are allowed to sign, or implicitly agree to, contingent contracts but also to renege on these contracts when it is to their advantage. That is to say, there are no courts to enforce contracts and low mobility or "lock-in" costs. We first develop a general framework for analysing contracts in this class of repeated interactions. The logic of these contracts follows that of repeated games, in that a party called upon to sacrifice current utility to maintain the insurance is prepared to do so in anticipation of receiving reciprocal benefits in the future. However in general first-best risk sharing can not be achieved, and it is what happens in the second best contracts which is of particular interest. What then follows is a selective overview of the existing literature where we concentrate on models where risk-sharing is the primary motivation for having any sort of contract.

We start by developing a basic two agent (worker/firm) model in which either agent can quit the relationship at any time, potentially at a cost, although we allow for the cost to be zero. The agents agree initially to a contingent sequence of wages (and potentially a termination rule) which satisfies certain incentive constraints. The outside environment can be summarised in the evolution of the respective outside options for the two agents. The basic characterisation of second-best contracts can then be applied to specific models, and we do this this to summarise the existing theoretical work in the area. In the development of the model we do not use the dynamic programming framework that is usually used for this environment, but instead show that the model can be solved by using local variational arguments, thus avoiding the need to establish a number of technical properties of value functions.

Although the basic characterisations of the second-best contracts have been known for some time, there has recently been an upsurge of interest in applications of this type of model to macroeconomics, and of testing of the model particularly in the one-sided limited commitment case where workers are mobile but firms can commit. We attempt to summarise the main findings which are generally very supportive of the one-sided model.

2 A general model of limited commitment

2.1 A baseline model

The model is as follows. There is an infinite horizon, $t = 1, 2, 3 \ldots \infty$. Workers are risk averse with per period twice differentiable utility function $u(c)$, $u' > 0, u'' < 0$, where $c \geq 0$ is the income/consumption of the single good received within the period; crucially, it is assumed that they cannot make capital market transactions, so the only possibility for consumption smoothing across states of nature or over time arises if the firm provides insurance. There is no disutility of work, but hours are fixed so that workers are either employed or unemployed (although we relax this assumption below). The firm is assumed to be risk neutral. We consider a single match between one worker and one firm, and for the moment we do not need to fill in the details of the outside environment. There is perfect information within the match. We suppose that output at time t within this match is $z(s_t) \geq 0$, where s_t is the current state of nature.[1] The state of nature s_t follows a time-homogeneous Markov process, with finite state space S, and initial distribution p over S, and from state s state $r \in S$ is reachable next period with transition probability: $\pi_{sr} \geq 0$. Let $h_t := (s_1, s_2, \ldots, s_t)$ be the history at t. Workers and firms discount the future with common discount factor $\beta \in (0, 1)$.

[1] We do not identify the state of nature directly with productivity (z) as it may be that other firms face different productivity shocks, and so the outside options will not depend wholly on the match productivity. In fact, even if there is a common productivity shock, because optimal contracts do not necessarily express current wages as a function of only the current state, the outside option may not be a function of the current state only (although in most models the payoff from a new contract, and hence the outside option, will only depend on the current state).

At the start of date 1, after the initial state s_1 is observed,[2] the firm offers the worker a contract $(w_t(h_t))_{t=1}^{T} = ((w_1(s_1), w_2(s_1, s_2), w_3(s_1, s_2, s_3), \ldots))$, where $w_t(h_t) \geq 0$ is the wage at t after history h_t, and $T > 1$ is the (random) date at which the contract is terminated.[3] The within period timing is as follows. At the start of each period, both agents observe the current state of nature, s_t. At this point either party can quit and take their outside option. Otherwise, they trade at the agreed terms, in which case the value of output $z(s_t)$ is realised, and the firm then makes a wage payment according to the contract. The value (discounted utility) of the outside option for the worker is denoted by $\chi(s)$ in state s, and for the firm it is always zero.[4]

Let $V_t(h_t)$ denote the continuation utility from t onwards from the contract (assuming it does not terminate at t):

$$V_t(h_t) := u(w_t(h_t)) + E\left[\sum_{t'=t+1}^{T-1} \beta^{t'-t} u(w_{t'}(h_{t'})) + \beta^{T-t}\chi(s_T) \mid h_t\right], \tag{1}$$

where E denotes expectation. Likewise the firm's continuation profit is

$$\Pi_t(h_t) := z(s_t) - w_t(h_t) + E\left[\sum_{t'=t+1}^{T-1} \beta^{t'-t}\left(z(s_{t'}) - w_{t'}(h_{t'})\right) \mid h_t\right]. \tag{2}$$

The contract is said to be *self-enforcing* if the following hold for all dates t, $T-1 \geq$

[2]If matches also start at later dates, the characterisation developed below, which depends only on the state prevailing at the time the contract starts, is the same.

[3]So that at $t = T$, after observing the current state s_t, the partnership dissolves and both agents get their outside options. T is a random variable (a stopping time) so that the length of the contract will in general depend on the history of shocks. At this level of generality, termination must be allowed for as there may be no continuation values that satisfy participation constraints.

[4]This reflects that fact that in much of the existing literature it is assumed that competition among firms drives profits to zero from new matches. This is inessential and nothing of what follows depends on it. If other inputs such as capital were included in the firm's profits, then the participation constraint for the firm would require that it covers capital costs. This would make the firm's outside option state dependent like the workers's if say, the interest rate varied with the state. See e.g. Calmès (2007) for a model including a fixed capital component where the firm's outside option is state dependent.

$t \geq 1$, and for all positive probability h_t (with initial state s_1):

$$V_t(h_t) \geq \chi(s_t) - C_w, \tag{3}$$

$$\Pi_t(h_t) \geq -C_f, \tag{4}$$

where C_f and C_w are respective directly incurred quitting/mobility costs for the firm and worker.[5] Inequality (3) is the worker's *participation constraint* (henceforth PC) that says that at any point in the future the contract must offer at least what a worker can get by quitting, net of quitting costs, while (4) is the corresponding constraint for the firm.

We are interested in *constrained efficient contracts*, that is to say contracts which are self enforcing and are not Pareto dominated by any other self-enforcing contracts. Efficient contracts are thus solutions to the following problem:

$$\max_{(w_t(h_t))_{t=1}^{T}} \Pi_1(h_1) \qquad \text{Problem A}$$

subject to (3), (4), and

$$V_1(h_1) \geq \bar{V}_1. \tag{5}$$

The term $\bar{V}_1$ measures how much utility the worker gets from the relationship, and as this is varied across feasible values (i.e. values for which self-enforcing contract exist), all efficient contracts are traced out.[6]

Lemma 1 *In an optimal contract in which the firm's (worker's) participation constraint is slack at $t+1$, wages cannot fall (rise) between t and $t+1$.*

Proof. Suppose we are at h_t, and *suppose that the firms's participation constraint at $t+1$ in some state s is not binding*. Consider, starting from the optimal contract,

[5] We assume that these are also incurred if the contract is terminated by agreement (i.e. at $t = T$), so they are costs which cannot be avoided on match break-up. Obviously it would be equivalent to factor these directly into outside options. MENTION HIRING FIRING COSTS HERE???

[6] The issue of existence of solutions to this problem for feasible $\bar{V}_1$ is standard in this environment.

reshuffling wages between t, and $t+1$ in state s, to *backload* them. Increase the wage at $t+1$ after state s by a small amount Δ, and cut the wage at t by x so as to leave the worker indifferent; do not change the contract otherwise:

$$\pi_{s_t s}\beta u'\left(w_{t+1}(h_t, s)\right)\Delta - u'\left(w_t(h_t)\right)x \simeq 0.$$

This backloading satisfies all worker participation constraints since the worker's utility rises at $t+1$, and so even if her constraint were binding, it will not be violated; at t her constraint holds as her utility is unchanged, and likewise it is unchanged earlier since utility is held constant over the two periods. The change in profits (viewed from h_t) is

$$-\pi_{s_t s}\beta\Delta + x \simeq -\pi_{s_t s}\beta\Delta + \frac{\pi_{s_t s}\beta u'\left(w_{t+1}(h_t, s)\right)\Delta}{u'\left(w_t(h_t)\right)},$$

which is positive for Δ small enough if

$$\frac{u'\left(w_{t+1}(h_t, s)\right)}{u'\left(w_t(h_t)\right)} > 1. \tag{6}$$

If (6) holds (so that wages are falling), then the backloading would raise profits at t, so the firm's PC would hold at t, and at $t+1$ by assumption the firm's PC is slack, so a small change to the wage will not violate it. Thus all constraints are satisfied by this change, and profits have increased, contrary to the optimality of the original contract. So (6) cannot hold: marginal utility growth cannot be positive, or equivalently, *wages cannot fall*. By a symmetric argument if the worker's participation constraint is slack at $t+1$, then wages cannot rise between t and $t+1$.[7]

■

Next, we need to characterise more precisely what happens to the wage when one of the participation constraints binds. First, let $\left(w_t(\bar{V}_1; s)\right)_{t=1}^{\infty}$ be an optimal

[7]Suppose that the strict opposite of (6) holds, $u'\left(w_{t+1}(h_t, s)\right)/u'\left(w_t(h_t)\right) < 1$. The reverse argument can be used: frontloading would be profitable but it might violate the worker's $t+1$ PC, since wages fall at this point. In other words, if the worker's PC constraint is slack at $t+1$, wages cannot rise, but we cannot rule out a rise when this constraint is binding.

contract in Problem A starting from state $s_1 = s$. This must deliver precisely $\bar{V}_1$ to the worker, otherwise we can cut the period 1 wage without violating the worker's constraint, thus increasing profits.[8] We define $\underline{w}_s := w_1(\chi(s) - C_w; s)$, i.e. , the period 1 wage specified by an optimal contract starting in state s which delivers exactly the worker's outside option, $\bar{V}_1 = \chi(s) - C_w$. It must be unique by a simple convexity argument (see below). A key observation is the following: it must be optimal *at any date* t in state s to set $w_t = \underline{w}_s$ whenever $V_t(h_t) = \chi(s) - C_w$. This follows from the fact that the future distribution over states depends only on s, and that the continuation contract must itself be optimal (otherwise replacing the continuation contract by a lower cost one which delivered the same continuation utility would reduce the initial costs but satisfy all participation constraints). *Thus,* $\underline{w}_s$ *is the wage in state* s *at any* t *if the participation constraint is binding.* Similarly define $\overline{w}_s$ to be the period 1 wage specified by an optimal contract starting in state s which delivers profits of exactly $-C_f$.

It can then be established that if an optimal contract offers a higher utility, then it must offer a higher first period wage:

Lemma 2 *If* $V' > V$, *then* $w_1(V'; s) > w_1(V; s)$.

Proof. Assume otherwise, so that $w_1(V'; s) \leq w_1(V; s)$. Suppose at some point in the future on some path h_t that $w_t(V'; s) > w_t(V; s)$ for the first time. This h_t must exist as the V' contract offers higher utility. This implies that wage growth between $t-1$ and t is greater in the V' case, which from Lemma 1 can only be true if one or both of the following occur: (i) the worker's participation constraint binds at t for the V' contract; (ii) the firm's constraint binds in the V contract. In case (i) in the V' contract, wages are weakly lower than in the V contract until

[8]Provided $w_1 > 0$; otherwise, since it is assumed that the outside option dominates zero consumption for ever, it is easily shown that there must be a point in the future at which $w_t > 0$ and the worker's constraint is not binding, so wages can be cut at this point instead.

minimum continuation utility is obtained (so the V contract cannot offer less from this point); thus discounted utility cannot be greater in the V' contract, contrary to assumption. In case (ii) in the V' contract, wages are weakly lower than in the V contract until maximum continuation utility is obtained in the V contract, again contrary to assumption. ■

Proposition 3 *An optimal contract evolves according to the following updating rule. In state $s \in S$ either (a) the contract (always) terminates, or (b) there is associated a minimum and a maximum wage, $\underline{w}_s$ and $\overline{w}_s$ respectively ($\underline{w}_s \leq \overline{w}_s$), such that in an optimal contract if at date $t+1$ state s_{t+1} occurs then w_{t+1} is updated from w_t by*

$$w_{t+1} = \begin{cases} \overline{w}_{s_{t+1}} & \text{if } w_t > \overline{w}_{s_{t+1}}, \\ w_t & \text{if } w_t \in \left[\underline{w}_{s_{t+1}}, \overline{w}_{s_{t+1}}\right], \\ \underline{w}_{s_{t+1}} & \text{if } w_t < \underline{w}_{s_{t+1}}. \end{cases}$$

Proof. If there exist self-enforcing continuation utilities from s (i.e. if a self-enforcing contract exists) then by definition an efficient contract should continue as each player gets at least their outside option, and cannot be worse off. Otherwise termination must occur. Thus we w.l.o.g. assume termination does not occur at s for the remainder of the proof. We start by showing that $\underline{w}_s$ is unique. Suppose otherwise: then there are two distinct contracts that deliver $\chi(s) - C_w$ to a worker, both of which satisfy participation constraints and yield the same costs. Take a strict convex combination of these two contracts (i.e., a convex combination of wages at each h_t). From (1) and the concavity of $u(\cdot)$ it is clear this increases a worker's utility, and satisfies the participation constraint at each point. Costs are linear in wages, and hence are unchanged. Thus a small reduction in the initial wage (in state s) will still satisfy participation, and will lead to lower costs, a contradiction. So $\underline{w}_s$ is unique. Likewise $\overline{w}_s$ is unique. Moreover, by Lemma 2 $\underline{w}_s \leq \overline{w}_s$ since a contract that delivers the firm $-C_f$ (i.e., corresponding to $\overline{w}_s$) must deliver the worker at least $\chi(s) - C_w$ (i.e., corresponding to $\underline{w}_s$), otherwise the worker's PC would be violated. Next, suppose that $\underline{w}_{s_{t+1}} < \overline{w}_{s_{t+1}}$ and $w_t \in \left(\underline{w}_{s_{t+1}}, \overline{w}_{s_{t+1}}\right)$. If the worker's participation constraint at $t+1$ in state s_{t+1} binds, $w_{t+1} = \underline{w}_{s_{t+1}}$, i.e., wages fall (as $w_t > \underline{w}_{s_{t+1}}$), but then the firm's constraint is slack ($w_{t+1} \neq \overline{w}_{s_{t+1}}$), so this contradicts Lemma 1 which asserts that wages do not fall. Thus the worker's constraint does not hold, and we know from Lemma 1 that wages cannot rise. Likewise as $w_t < \overline{w}_{s_{t+1}}$ the worker's constraint cannot bind, and wages cannot fall. Thus for $w_t \in \left(\underline{w}_{s_{t+1}}, \overline{w}_{s_{t+1}}\right)$, wages remain constant. Conversely, if $w_t \leq \underline{w}_{s_{t+1}}$, then if the worker's constraint does not hold ($V_1 > \chi(s_{t+1}) - C_w$), by Lemma 1 wages cannot rise, so $w_{t+1} \leq \underline{w}_{s_{t+1}}$. However, $V_1 > \chi(s_{t+1}) - C_w$ would imply by Lemma 2 (comparing with the contract that delivers $\chi(s_{t+1}) - C_w$) that $w_1 > \underline{w}_{s_{t+1}}$, a contradiction. So the constraint binds and $w_1 = \underline{w}_{s_{t+1}}$. A symmetrical argument establishes that $w_{t+1} = \overline{w}_{s_{t+1}}$ if $w_t > \overline{w}_{s_{t+1}}$. ■

Thus wages evolve in a simple fashion: they remain constant unless this takes the wage outside the interval of efficient wages $[\underline{w}_s, \overline{w}_s]$ for the current state, in which case the wage changes by the minimum amount needed to bring it into this interval. The only thing remaining to be determined is the initial wage, $w_1(s_1)$. This will be determined by $\bar{V}_1$ in Problem A, and this can in turn be thought of as depending on the bargaining strengths of the two parties or the initial outside options of the two parties. By varying the initial wage all possible splits of the joint surplus will be traced out.[9]

The state-dependent wage intervals $[\underline{w}_s, \overline{w}_s]$ will in general depend on all the parameters of the model including the worker's preferences and the stochastic process for productivity. However, the outside option of the worker χ_s and the quitting and mobility costs C_w and C_f, will also play a crucial role in the determination of these interval endpoints. We will specify the more specific assumptions in various models as we encounter them. In the first paper to analyse a problem of this

[9]See Malcomson (1999) which shows ???

type, Thomas & Worrall (1988), it is assumed that C_w and C_f are zero, and if a worker reneges, thereafter he or she can find work only at the spot market wage, where because of competition among firms, the wage equals current productivity $z(s_t)$ (which is assumed to be a common shock across all firms). Similarly, if a firm reneges it is assumed it can hire at the spot market rate. This may be motivated as follows. Suppose there are, in addition to infinitely lived workers and firms, at each date m workers and n firms, $n > m$, who live for only one period. Since there is no enforcement mechanism and no mobility costs, the one-period-lived agents trade at the spot market wage. The infinitely lived agents are competitive and thus treat these spot market wages as given. this is then in line with reputation models of repeated games, and corresponds to the most severe credible punishment. It requires that when an agent reneges he is observed by everyone else, and once he has reneged he has proved himself unreliable and no one will sign a contract with him again. Likewise for a firm which has reneged in the past. The implication is that a worker who reneges will receive a consumption stream equal to productivity at each date, and so $\chi(s)$ equals the discounted expected utility generated by this stream.

2.2 Introducing variable hours

The baseline model presented above is important in understanding the behaviour of wages as the insurance motive partially disassociates wages from productivity. It is commonly observed in many countries that labour market fluctuations are characterised by large procyclical variations in hours, but far smaller variations in wages. It has been suggested that the insurance provided in wage contracts can help explain this (Rosen (1985), Azariadis (1975)). Abowd & Card (1987) and Boldrin & Horvath (1995) have tested the the implicit contract model of full insurance against the spot market alternative and have found some weak support for the contracting hypothesis over the alternative.

In order to address the behaviour of both wages and hours in the limited commitment model this subsection shows how the baseline model presented above can be extended to allow for joint determination of wages and hours within the contract. In this case a contract will specify not only a profile for wages $(w_t(h_t))_{t=1}^{T}$ but also a profile for hours worked $(H_t(h_t))_{t=1}^{T}$. It is assumed that the worker has per-period twice differentiable strictly concave utility function $u(c, H)$ where work is disliked, so $u_H < 0$. It will further be assumed that leisure is a normal good so that the Engel curve for hours worked is downward sloping. In the previous model we were implicitly assuming that hours were fixed, say at 1 unit and $u(c) = u(c, 1)$. As before it is assumed that workers cannot engage in capital market transactions so that consumption is equal to earnings, $c(h_t) = w(h_t)H(h_t)$. The continuation utilities are defined analogously to equations (1) and (2) but with the per-period payoffs of the worker and the firm are replaced by $u(c_t(h_t), H(h_t))$ and $z(s_t)\,H(h_t) - c_t(h_t)$ respectively. The self-enforcing constraints are then still given by equations (3) and (4) and constrained efficient contracts can be found by solving

$$\max_{(c_t(h_t), H_t(h_t))_{t=1}^{T}} \Pi_1(h_1) \qquad \text{Problem A}'$$

subject to (3), (4), and (5). Again if matches start at a later date the characterisation is the exactly same as it depends only on the state in which the match is initiated.

The first thing to note about the solution to Problem A′ is that hours will be chosen efficiently so that for every history[10]

$$-\frac{u_H(c_t(h_t), H_t(h_t))}{u_c(c_t(h_t), H_t(h_t))} = z(s_t). \tag{7}$$

To see this consider a pure intratemporal reallocation of consumption and hours that leaves profits unchanged. That is consider a change in consumption of Δc and a change in hours δH such that $\Delta c = z\Delta H$. The net effect on utility is approximately $u_c(c, H)\Delta c + u_H(c, H)\Delta H = (u_c(c, H)z + u_H(c, H))\,\Delta H$. Thus if $-u_H/u_c < z$ a

[10]This was first pointed out in Beaudry & DiNardo (1995).

small decrease in hours, $\Delta H < 0$ would raise utility and if $-u_H/u_c > z$ a small increase in hours would raise utility. Hence at the optimum (7) must hold. The reason why this condition holds is that the self-enforcing are concerned only with the intertemporal allocation and thus do not interfere with the efficient intratemporal allocation of hours.[11]

It is further possible to find the updating rule analogous to Proposition 3. To do this we define the marginal utility of consumption

$$\lambda_t(h_t) = u_c(c_t(h_t), H_t(h_t)).$$

Associated with each $\lambda_t(h_t)$ is an interval $\left[\underline{\lambda}_{s_{t+1}}, \overline{\lambda}_{s_{t+1}}\right]$ and the updating rule for λ is given by

$$\lambda_{t+1} = \begin{cases} \overline{\lambda}_{s_{t+1}} & \text{if } \lambda_t > \overline{\lambda}_{s_{t+1}}, \\ \lambda_t & \text{if } \lambda_t \in \left[\underline{\lambda}_{s_{t+1}}, \overline{\lambda}_{s_{t+1}}\right], \\ \underline{\lambda}_{s_{t+1}} & \text{if } \lambda_t < \underline{\lambda}_{s_{t+1}}. \end{cases} \tag{8}$$

Here $\overline{\lambda}_{s_t}$ is the value of λ which delivers the exactly the worker's outside option and $\underline{\lambda}_{s_t}$ is the value that delivers the firm's outside option. The initial value of λ will be determined by the bargaining strength or initial outside options of the parties as reflected by $\bar{V}_1$ in equation (5).[12] It is easy to see that if hours are fixed then $\overline{\lambda} = u'(\underline{w})$ and $\underline{\lambda} = u'(\overline{w})$.

The contractual solution for the path of wages and hours therefore satisfies the two equations

$$u_c(c_t(h_t), H_t(h_t)) = \lambda_t(h_t) \tag{9}$$

$$-u_H(c_t(h_t), H_t(h_t)) = \lambda_t(h_t) z(s_t) \tag{10}$$

where $c_t(h_t) = w_t(h_t)H_t(h_t)$ and $\lambda_t(h_t)$ follows the updating rule given by equation (8). The solutions to the two equations (9) and (10) are the Frisch-type demand

[11] If there were also a moral hazard or adverse selection problem then (7) would not hold and in general there would be an interaction between the intratemporal and intertemporal allocation problems.

[12] Here λ is the inverse of the multiplier in Problem A$'$ so a lower value of λ corresponds to a greater bargaining strength for the worker.

functions $c(\lambda, z)$ and $H(\lambda, z)$.[13] It is easy to check that provided leisure is a normal good, hours are increasing and the wage rate is decreasing in λ.[14] Equally hours are increasing and the wage rate is decreasing in productivity z.[15] The intuition is that a decrease in λ holding z fixed, and hence holding the marginal rate of substitution constant, is a pure positive income effect and therefore because leisure is normal leads to a decrease in hours worked. Equally, an increase in productivity holding the marginal utility of consumption fixed leads to a substitution effect and therefore an increase in hours worked.

The implications of the model have been considered and tested by Beaudry & DiNardo (1995). Consider first the case of complete insurance so that λ is fixed and determined by the initial bargaining position at the time the contract is initiated. This may vary from worker to worker. Thus workers who enter the contract with a better bargaining position will in any given state (and hence productivity z) have higher wages and lower hours. Looking at a cross section of workers therefore it is to be expected that hours are negatively related to wages. This is to be contrasted with the standard intertemporal model of labour supply. In that model equations (9) and (10) apply with $z = w$ and with λ determined by an Euler equation of the form $\lambda_t = (1 + r_t)\beta \mathrm{E}[\lambda_{t+1}]$. Since the standard intertemporal model of labour supply allows the worker to self-insure through borrowing and lending, earnings need not equal consumption and therefore it follows directly from equation (10) that the function $H(\lambda, w)$ is increasing in w provided only that the marginal disutility of work is increasing.[16] This is the intertemporal substitution effect that as wages

[13] The Frisch demand functions are derived by keeping the marginal utility of wealth constant and where the marginal rate of substitution equals the real wage.

[14] This is easy to see if utility is separable in earnings and hours worked: with z fixed an increase in λ increase the marginal disutility of labour and hence the hours worked. Equally an increase in λ increases the marginal utility of consumption so that consumption or earnings is decreased. Since hours are increased it follows that the wage rate falls.

[15] The effect of an increase in z on earnings is ambiguous and depends on whether the marginal utility of consumption increases or decreases with hours worked (i.e. on the sign of u_{cH}): if utility is separable earnings are independent of z for a fixed λ.

[16] The function $H(\lambda, w)$ is the Frisch labour supply function with the inverse marginal utility

rise more hours work are supplied so that wages and hours should be positively associated holding the marginal utility of wealth fixed.

Now consider the case where the participation constraints are binding in some states. Depending on the history of states any individual worker may have any $\lambda_t(h_t) \in \left[\underline{\lambda}_{s_t}, \overline{\lambda}_{s_t}\right]$ for a given state and productivity $z(s_t)$. This has three important effects. First although different workers initially employed at different dates may have different λs, as soon as both workers are constrained in a particular state (or the firm is constrained for both workers), their λs will be equalised and therefore they will have the same wages and hours in subsequent periods. Thus the cross-sectional variation in wages and hours across employees should be lower with increasing tenure. Second, for any worker who is constrained following an increase in productivity, there will be a decrease in λ and two offsetting effects: the hours worked will increase because of the increase in productivity but the decrease in λ will offset this and tend to reduce hours worked. Similarly the wage rate will rise because of the decrease in λ but fall because of the increase in productivity. Thus the model will predict an ambiguous or weak effect of changes in productivity on hours and wage rates. Thirdly, for workers with different starting points the change in λ experienced by different workers will be different. Therefore the consequent growth rates in wages and hours will vary across workers of different tenure.

2.3 Endoginising the workers' outside option

As explained above in Thomas & Worrall (1988) the worker's outside option $\chi(s)$ is determined by what a worker would get on the spot market for evermore, which depends only on the exogenous productivity process. In order to justify this assumption it is necessary to assume that all firms can observe when a worker reneges

of wealth held constant. Of course λ will not in general be constant over time and therefore the long-run elasticity of wages on hours will depend on the evolution of λ.

on a contract, and punish the worker by not offering her anything other than a spot contract. This requires that firms can perfectly observe a worker's past history. An alternative assumption is that firms treat all new workers in the same way, irrespective of whether or not they have reneged on a previous contract. According to this view, when a worker quits a firm, she can look for a new job offering as much insurance as in the contract from which she just quit. If however, workers and firms could move costlessly to other contracts then no non-spot contracts could be sustained.[17] Therefore it will be necessary to assume either that the firms can commit to contracts or that there are other frictions such as search costs in the labour market. We deal with each of these in turn.

2.3.1 Search frictions

In this section we discuss two papers (Sigouin (2004), Rudanko (2006)) which embed the above model into a matching framework to analyse the association of certain variables with aggregate productivity. Both argue that the two-sided limited commitment model performs better than full commitment models and other versions such as spot contracts, one-sided limited commitment or continuous bargaining. Sigouin (2004) allows hours, but not employment, to vary, while Rudanko (2006) allows employment and vacancies to vary. However in both of these matching models there is also the possibility of an unemployment spell before a new contract is found, so $\chi(s)$ is less than the utility from a new contract.

Sigouin (2004) develops the model with variable hours by allowing the outside option $\chi(s)$ to be determined by contracts offered by other firms, rather than on a spot market as in the Thomas-Worrall model. He assumes however, that if a worker

[17] This assertion assumes that the surplus split from a new contract is always the same. Otherwise quitters could be punished effectively by starting a new contract so that the other agent gets all the surplus from the relationship. For example, in the Thomas-Worrall model this would imply that punishments are as severe as consignment to trading on the spot marker, so the same set of contracts are self-enforcing.

quits from one firm he or she faces a probability of not being matched with a new firm (even though if matching does occur, it happens without a delay) and being unemployed.[18] This is sufficient to drive a potential wedge between what a worker can get by remaining in the contract and what is available by quitting, and allows for some insurance to be sustained. Then $\chi(s)$ is determined by what a worker would get by quitting and waiting for a job; because of competition among firms a new job yields the worker the maximum surplus from a self-enforcing contract; however the worker may be unlucky and suffer unemployment, so this is also factored into $\chi(s)$.

Each worker has a total time endowment which is normalised to one, and can supply up to this amount to a single firm at any date. The productivity per hour worked is $z(s_t)$ at time t, which is common to all firms. However there is also a match specific shock, which can reduce productivity to zero (where it remains). If this happens, the match is dissolved. A worker has separable preferences at t given by

$$\mathrm{E}_t \sum_{j=t}^{\infty} \beta^{j-t} \left[\ln(c_j) + B(1-\eta)^{-1}(1-H_j)^{1-\eta} \right],$$

where c_j is consumption and H_j is hours supplied at time j. With separable preferences the updating rule of (8) implies that each state $s \in S$ is associated minimum and maximum *earnings*, $\underline{W}_s$ and $\overline{W}_s$ ($\underline{W}_s \leq \overline{W}_s$), such that earnings are kept constant if possible and otherwise move by the smallest amount to $\underline{W}_s$ or $\overline{W}_s$. In addition earnings and hours satisfy equation (7) that the marginal rate of substitution equals the marginal product. With the separable specification of preference

[18] There is no cost to posting a vacancy, but only a fixed fraction of the unemployed are able to make a match, or rather, to 'see' wage offers (i.e. they are nor directly matched, but are able to enter into a contract, whereas the unlucky ones cannot). This implies that the unemployment rate does not vary. Essentially he posits a matching function where the matching or "seeing" probability does not depend on the number of vacancies but only on the number seeking work. Moreover, although each entrepreneur can only match with a single worker, there are more entrepreneurs than workers so that competition between entrepreneurs for the fraction of unemployed workers who can see offers drives profits down to zero.

this gives

$$(w_{t+1}H_{t+1})\,B\,(1-H_{t+1})^{-\eta}=z\,(s_t)\,.$$

Notice that under a full commitment contract with these preferences a risk-neutral firm will stabilise total earnings while hours will vary procyclically with productivity (according to the intertemporal elasticity of labour supply described above). This leads to the (counterfactual, given the very weak empirical correlation) conclusion that the wage rate is perfectly negatively correlated with hours supplied. On the other hand under a spot labour contract, where the wage is always equal to productivity $z\,(s_t)$, these preferences have the property that income and substitution effects of a wage change cancel out (assuming that all income is labour income and there are no taxes). In this case hours do not vary at all with the wage or productivity (this contradicts the positive correlation between hours and productivity typically found in the data).

The situation will however, be slightly different when the enforcement constraints may bind. For relatively small changes in productivity (and assuming that earnings are not already up against the constraint that tightens) such that $w_tH_t\in\left[\underline{W}_{s_{t+1}},\overline{W}_{s_{t+1}}\right]$, so neither constraint is binding with strictly positive shadow value, the rule says that earnings stay constant, so there is no income effect, and hours change with productivity according to the intertemporal elasticity of supply. On the other hand, if the change is large enough that a constraint binds, then earnings change and there will be an income effect which reduces to an extent the change in hours. For example, a large increase in productivity may imply only a small increase in hours if earnings rise substantially, so the wage will also rise.[19] In this case then there is a positive correlation. The overall effect may then be that the correlation is very weak, in accordance with the evidence.

[19]This depends on how $\left[\underline{W}_{s_{t+1}},\overline{W}_{s_{t+1}}\right]$ varies with z_{t+1} but Sigouin shows through numerical simulations that the intuition will be correct in many situations.

Rudanko (2006) also embeds the basic model in a model of search. She addresses issues recently raised by Shimer (2005) who argues that the Mortensen-Pissarides model cannot account for the magnitude of unemployment and vacancy fluctuations without assuming unrealistically high volatility in productivity. Hall (2005) argues that some form of wage rigidity may be sufficient to solve this puzzle. Rudanko looks at different versions of a contracting model in a directed search model of the labour market, following Moen (1997), rather than the random matching model typically used in this literature. The model has similarities with the Sigouin model in that match specific productivity is composed (as the product of) a common (economy wide) component and match component that is unity initially, but transits to an absorbing state of 0 with a fixed probability each period. As in Sigouin, when this occurs, the match dissolves and the worker looks for a new job. Likewise there are a large number of risk-neutral entrepreneurs operating under constant returns to scale. (Unlike Sigouin, however, hours are fixed.) The model is one of competitive search: At the start of each period, after observing the current aggregate productivity level, each entrepreneur posts an offer of a wage contract, but has to pay a cost k for keeping a vacancy open. Worker search can be directed to a particular wage contract σ. There is a matching function defined as follows: if there is a measure N_u unemployed agents searching for σ and measure N_v vacancies offering σ, the measure of matches taking place this period is given by a Cobb-Douglas matching function

$$m(N_u, N_v) = K N_u^{\alpha} N_v^{1-\alpha}$$

where $0 < K < 1$ and $0 < \alpha < 1$. Defining $\theta = N_v/N_u$ to be the vacancy unemployment ratio ("labour market tightness"), the probability that a worker finds a contract σ this period is $m(\theta) := m(N_u, N_v)/N_u$, and the corresponding probability for an entrepreneur is $q(\theta) := m(N_u, N_v)/N_v$. Thus the payoff to a worker from searching for σ is

$$\mu\left(\theta(z)\right) V_\sigma(z) + \left(1 - \mu\left(\theta(z)\right)\right) V_u(z)$$

where $V_\sigma(z)$ is the discounted worker utility from finding a job with contract σ, while $V_u(z)$ is the corresponding utility from being unemployed, where both are functions of the prevailing aggregate state z. $V_u(z)$ is the discounted utility from consuming the unemployment benefit today and searching again tomorrow. Likewise $V_\sigma(z)$ is just the expression given in the original model for contract utility with a stochastic termination added, at which point the worker gets $V_u(z')$ if z' is the current state as she is unemployed for a period and then has to seek a new job. The firm's profit per job will depend on the probability that a job is filled, $q(\theta)$, and equals $q(\theta(z)) F_\sigma(z) - k$ where $F_\sigma(z)$ is the discounted profit from σ, but this is only achieved if a match occurs, but the vacancy cost k must be incurred in any case. Because of competition among entrepreneurs, this profit is driven to zero in equilibrium. The self-enforcing constraints specify that a worker cannot gain by leaving the contract, which requires that continuation utility must not be below $V_u(z')$ (the worker is unemployed for at least a period), and again that the continuation profits of the entrepreneur are nonnegative. In addition, for equilibrium to obtain it must be the case that there is no other contract that could be offered which would offer greater profits, where the corresponding θ will equate the returns to workers from searching in either market.[20] As in Sigouin, the model endogenises the worker's outside option so that it depends on what she would get by starting a new contract, but again the risk of unemployment (here it will last at least one period) is a sufficient deterrent to allow non-spot contracts to be sustained.

[20]Rudanko shows that only a single contract is ever offered to new matches in equilibrium. Moreover, it is equivalent to a model with undirected search in which a weighted Nash product of surpluses (relative to $(V_u(z), 0)$) is maximised, with weights proportional to the exponents in the matching function, i.e., α and $1 - \alpha$.

2.3.2 One-sided limited commitment

We next consider the influential paper by Beaudry & DiNardo (1991)(hereafter BD91). They develop a model of labour contracting where a risk-neutral firm offers insurance to risk-averse employees, but there is no worker commitment and unlike the search models considered above a worker who quits can immediately start work elsewhere (perfect mobility). In terms of our model above, they assume that $C_f = \infty$ (firm commitment) and $C_w = 0$ with $\chi(s_t)$ given by the utility from starting a new job (perfect labour mobility). We derive their basic characterisation, which is a generalisation of Holmstrom (1983) who considered a two-period model. We then describe the other ingredients of their model which lead to empirically testable predictions, and finally we discuss the empirical evidence. Their work is particularly important for two reasons. First, they provide strong evidence in favour of the perfect mobility model. Secondly, the paper addresses how wages respond to unemployment levels over the business cycle. There is a voluminous literature that examines how real wages respond to contemporaneous movements in unemployment which generally has not found a very strong relationship, but the results in BD91 suggest that this literature may have been looking at the wrong business cycle variable. If one looks at the lowest unemployment rate since a worker started a job, this appears to show a much stronger effect.

Given that $C_f = \infty$, we can treat the value of $\overline{w}_s$ as being infinite. (Alternatively, we can just ignore the firm's constraint in all the above arguments, so Lemma 1 directly asserts that wages cannot fall, etc.). Thus the intervals for efficient wages become $[\underline{w}_s, \infty)$. The ratchet nature of wages follows from Proposition 3: $w_{t+1} = \underline{w}_{s_{t+1}}$ if $w_t < \underline{w}_{s_{t+1}}$, and otherwise $w_{t+1} = w_t$. To pin down the values for the $\underline{w}_s$, we need to specify the process for $\chi(s_t)$ and how the contractual surplus is split between worker and firm.

BD91 assume that there are a large number of identical firms and workers, with new workers entering each period to replenish the labour force, replacing workers who die.[21] It is assumed further that because firms operate under constant returns to scale, competition for workers drives profits for a new worker to zero, so any surplus goes to the worker. A worker who quits a firm can immediately seek employment with another firm. Moreover the only source of uncertainty is the common shock to productivity each period. What this implies is that $\chi(h_t)$, the utility of the worker's outside option, equals the utility from an optimal contract which generates zero profits.[22] Given the updating rule, it is then possible to calculate the initial wage of a contract starting in state s for which discounted expected wages and discounted expected productivity are equal. This must therefore be $\underline{w}_s$.

What is perhaps surprising at first glance is how it is possible to offer any insurance at all when the worker can quit and *restart*the contract at a different firm, without any penalty.[23] Normally in repeated game models of cooperation players are induced to take short-term sub-optimal actions (such as paying out on insurance) by the promise of long-term rewards relative to reneging on this, which yields termination. But here a worker who quits is able to immediately start a new contract with a different firm so that whenever productivity is such that the contract demands a sacrifice by the worker, the worker can quit. The resolution of this apparent paradox is that contracts demand up-front payments by workers in that initially the worker receives a wage below productivity, to be compensated later by the likelihood of wages above productivity.[24,25]

[21]BD91 also have firm death, but we shall abstract from this in the exposition that follows.

[22]BD91 express the worker's participation constraint equivalently as the fact that the contract must never offer strictly positive profits, looking forward from any point—if it did then the worker would be bid away.

[23]In fact this intuition is correct in the two-sided case where the firm could also terminate the relationship costlessly. In the Sigouin and Rudanko models discussed above, there is the possibility of unemployment if a worker quits, and this is sufficient to support non-trivial contracts.

[24]This issue has been explored by Krueger & Uhlig (2006) in a general risk-sharing context where both parties to the contract are risk averse.

[25]The feature that workers initially receive wages below productivity with a rising wage profile

To see how the contract works, we shall take a very simple example. Suppose that in period 1 productivity $z(s_1) = 10$. From period 2 onwards, productivity will be in an absorbing state, fixed at either 11 (good state) or 9 (bad state) with equal probability. Thus for $t \geq 2$, in the good state the wage in a new contract must be constant at 11,and likewise in the bad state it must be constant at 9 (in either case, this is the first-best zero-profit contract and hence cannot be improved upon). An optimal contract at date 1 must pay at least 11 from $t = 2$ if the good state occurs, otherwise the worker will leave and get 11 elsewhere (and the constant 11 contract is the best way of delivering the utility, so no nonconstant contract can be better). In fact the optimal contract will pay a wage $\tilde{w}$ between 9 and 10 in period 1, and will continue paying $\tilde{w}$ for ever if the bad state occurs, and will pay 11 for ever if the good state occurs. The firm makes sufficient profit $(10 - \tilde{w})$ in period 1 to offset the future loss $(\tilde{w} - 9)$ suffered each period in the bad state, so that overall it makes exactly zero profit (the precise value of $\tilde{w}$ will depend on the discount and death rates). To get some intuitive idea of why downward rigidity is optimal, notice that if the bad state occurs the participation constraint is not binding. Thus there can be no benefit from cutting wages—a constant wage offers perfect smoothing and does not violate participation. It is now possible to see how insurance can be provided (the alternative would be that the worker gets paid the productivity level each period), despite the quitting option: the worker effectively pays an up-front insurance premium $(10 - \tilde{w})$ in return for later insurance (i.e., a wage greater than productivity, if the bad state occurs). The firm will then move from breaking even to negative discounted profits if the bad state occurs. (If the firm could also quit then the contract breaks down as it will happily take the up-front premium but dissolve the relationship when called upon to pay out later.)

In order to get testable restrictions, it is necessary to link the productivity

is of course reminiscent of the agency models where rising wages provide incentives for effort, see e.g. Lazear (1981).

level in the theoretical model to an observable variable. Notice that the optimal wage contract depends *only* on the productivity process—a very convenient feature. Moreover the labour market must always clear, since at the point of hiring there are no restrictions on wages. However when productivity is high, the wage and expected utility for a new entrant is high. BD91 posit an alternative sector in which a worker could be employed which is subject to a (fixed) decreasing returns technology. Thus a new entrant to the labour market faces a choice between a period in the alternative sector and then getting a contract, versus getting a contract in the original sector right away (by construction of the equilibrium, once a worker has a contract, the alternative of moving to the alternative sector will offer the same as a new contract, and so is always weakly dominated due to the participation constraint). In equilibrium workers will be indifferent (there are always some workers employed in the alternative sector) so a high wage in the original sector must go along with a low level of employment in the alternative sector (due to decreasing returns, this raises the wage), or equivalently, high employment in the original sector. BD91 associate high employment in the original sector with low unemployment.[26] Putting this together, high wages go together with high unemployment levels.

BD91 conclude, then, that with no worker commitment (perfect mobility), where the worker is free to quit at any point, the wage follows a ratchet like process, rising whenever the labour market is tighter than hitherto (since the worker joined the firm), but staying constant otherwise; hence the current wage is determined by the tightest labour market during a worker's tenure. Tightness of the labour market is measured by how low the unemployment rate is.

[26] It is tempting to interpret the alternative sector as leisure or some sort of household production, although the decreasing returns to total labour input makes this interpretation difficult.

3 Empirical evidence

The data typically show that real wages are only weakly correlated with productivity or even mildly countercyclical. Hours on the other hand are found to be quite strongly positively correlated with productivity. To match this observed pattern in the data using standard real business cycle models requires a very high intertemporal elasticity of substitution for labour supply that is not supported by estimates from micro data. Recently Shimer (2005) has suggested that standard search models under-predict the volatilities and unemployment because of the flexibility of wage movements to productivity unless implausibly large shocks for productivity are assumed. This section therefore considers some of the available empirical evidence on whether these puzzles might be resolved within the limited commitment labour contracting model.

3.1 One-sided commitment and perfect worker mobility

We first consider the one-sided commitment case with perfect labour mobility of Beaudry & DiNardo (1991) As explained above, this model leads to the prediction that the current wage of an individual worker is determined by the tightest labour market during the worker's current job tenure. In testing, this perfect mobility model does better than two alternatives: a spot market model in which current unemployment determines wages, and a full commitment model in which unemployment at the time of hiring is the determining factor. In the spot market model, wages are determined solely by the value of a worker's current marginal product, in the full-commitment contracting model, wages are constant but the level is determined by the worker's outside opportunity at the point at which he/she joins the firm. Beaudry & DiNardo (1991) test these three models against each other on U.S. data (Panel Study of Income Dynamics (PSID)/CPS). Perhaps surprisingly, the latter model appears to perform much better than the other two, which we describe in more detail:.

Commitment: a binding contract is signed when the worker joins a firm. Because the worker is risk-averse, the risk-neutral firm acts as an insurance company, completely stabilising wages. (This results from our above general model by imposing $C_w = \infty$, so that $\underline{w}_s = -\infty$.) In equilibrium workers will be offered a fixed wage contract (where the wage will equal the expected discounted value of a worker's productivity so firms make zero profits). The wage will be fixed at a level corresponding to conditions *at the point the worker joins the firm*—it equals the best estimate of a worker's lifetime productivity, and under the assumed productivity process this will depend only on his productivity at this point, which is, as explained above, proxied by the unemployment rate, U_t, at that point.

Spot market contract: no long-term contract is possible, so this implies that $w_t = z(s_t)$. (If a firm offered say a fixed wage contract, then whenever the wage was less than $z(s_t)$ the worker could just walk away, and go to another firm, while if the wage was greater than $z(s_t)$ the firm could sack the worker.) Thus wages fluctuate with $z(s_t)$ which is proxied by U_t.

The general model can be expressed as follows: the natural log of the real wage for worker j at time $\tau + t$ for a worker who started the job at time τ satisfies:

$$\ln w_{j,\tau+t} = \alpha_1 X_{j,\tau+t} + \alpha_2 C(\tau, t) + \varepsilon_{j,\tau+t}$$

where $X_{j,\tau+t}$ is a vector of individual variables[27], α_1 is the vector of coefficients on these variables, $\varepsilon_{j,\tau+t}$ is an error term, and α_2 is the coefficient on the business cycle (i.e., unemployment) variable, with the 3 possibilities for the business cycle variable

[27]For individual characteristics, BD91 used experience, experience squared, how much schooling, job tenure, and dummies for industry, region, race, union status, marriage, and metropolitan area (SMSA).

$C(\tau, t)$ being:

$$C(\tau, t) = \begin{cases} U_{\tau+t} & \text{spot market model} \\ U_{\tau} & \text{fully binding contract} \\ \min\{U_{\tau+k}, k = 0, 1, \ldots, t\} & \text{non-binding on worker} \end{cases}$$

where the unemployment rate is denoted by U, with U_{τ} the rate prevailing at the start of the job and $U_{\tau+t}$ the rate at time $\tau + t$ where t denotes tenure with the employer.

In some specifications[28] in which all three variables are included, the coefficient on the minimum unemployment rate is the only correctly signed (i.e., negative) significant one (PSID, no fixed effects), and in all specifications it is much larger than the other coefficients, implying that a 1% drop in the minimum unemployment rate (e.g., from 4% to 3%) leads to an increase in current wages of between 3% and 8%.

The implications for our understanding of real wage cyclicality are considerable. Typically studies have looked at how wages respond to contemporaneous unemployment movements. For example, using the PSID for men over the period of 1968-69 to 86-87, Solon, Barsky & Parker (1994) found that a one percentage point reduction of the unemployment rate leads to a rise in the real wage rate of workers who stayed in their jobs by 1.2 percent (movers appear to be subject to greater procyclical wage movements). Similar estimates are found in Shin (1994) and Devereux (2001). BD91's results suggest that the response of wages to the minimum unemployment rate is substantially larger. On the other hand, as argued in Grant (2003), because the minimum unemployment rate does not actually vary as much as contemporaneous unemployment (consider a worker whose minimum value occurred early in a job spell), minimum unemployment may not explain very much of the variability of aggregate wages over the business cycle.

[28]See Table 2 of their paper.

Several recent empirical studies have largely confirmed the robustness of BD91's main empirical findings over different periods and using different datasets, that the minimum rate of unemployment since hiring is a statistically important determinant of the current wage of an individual (McDonald & Worswick 1999, Grant 2003, Shin & Shin 2003, Devereux & Hart 2005). Both Grant, and Devereux and Hart, however, find more of a role for the current unemployment rate than did BD91. Grant (2003) extends BD91's analysis (using six cohorts from the National Longitudinal Surveys) to cover the time period 1966 to 1998. He finds that the significance and importance of $\min u$ is broadly robust with respect to the addition of fixed time dummies (to rule out any effects coming through macroeconomic variables, and thus the coefficient on $\min u$ is estimated only through variation across individuals in each year), of tenure dummies (to capture nonlinear tenure effects), a tenure-unemployment interaction term (to capture tenure effects that vary over the business cycle), and using subsamples selected on the basis of age, and sex. As mentioned, however, current unemployment levels also have some explanatory power.

A somewhat different methodology was adopted by Shin & Shin (2003), using the PSID for the period 1974-91, which includes one business cycle more than BD91. They run the BD91 regressions over the whole period and get very similar results—but as Grant does, they also find more significant results for contemporaneous unemployment.[29,30] They also estimate a complementary econometric model,

[29]In comparing their estimates with those of Grant, it is interesting to note that the PSID sample has a higher average age than in the NLS except for the NLSY Older Men. Estimates in Table 2 of Grant (2003) show that the effect of the minimum unemployment rate on current wages dominates those of the other unemployment variables more in the NLSY "Older Men" cohort than in "Young Men" or "Women," and the "Older Men" results are closest to the PSID estimates. This suggests that the BD91 model may work better for older workers.

[30]Shin & Shin (2003) include a trend, which might matter as the period studied has a generally rising unemployment rate, so that the a job's minimum unemployment rate is negatively correlated with time elapsed since the date at which the minimum is attained; as wages are rising omitting this trend might overstate the effect of the minimum unemployment rate. However it makes little difference, as one would anticipate from Grant's analysis with time dummies. Likewise, to rule out nonlinear effects of tenure they find that the addition a squared tenure term does not matter to the worker fixed effect model (although it does to the no-fixed effects specification); again this confirms Grant's findings.

only using the current unemployment rate as a business cycle regressor, but look for asymmetric effects of tight labour markets. Thus they split a job history into periods of tightening and loosening labour markets, and subdivide the former category into two sub-categories, when unemployment is falling but above its minimum for the current job, and when it is below the minimum. Tenure is measured with considerable error in the PSID; thus a mismeasurement in tenure may lead to an incorrect value for $\min u$, used in BD91's estimation, whereas here it will lead to the respective periods when unemployment is falling but below or respectively above the historical minimum, to be measured incorrectly. It is argued that the former is more likely to be problematical. The results are that most of the wage adjustment occurs in periods when the unemployment rate falls below the historical minimum level observed since the start of the current job in accordance with the perfect mobility model (according to the model, wages should be constant in other periods): For the sample of male household heads, the estimated coefficient on the unemployment rate is -0.026 (i.e., a one percentage point reduction in the unemployment rate is associated with a 2.64 percent rise in real wages). The coefficient on unemployment when it is falling but not below the historical minimum is much smaller at -.0076, but not significant. For periods of contraction, the coefficient is still smaller and insignificant. So again there is a strong confirmation of the perfect mobility model. They also confirm the findings of other studies that the wages of job stayers are procyclical, but less so than those of movers.

3.2 Two-sided limited commitment

The most direct testing of the two-sided model is by Macis (2006), using Italian panel data on a sample of 1500 firms which includes detailed information on all workers with these firms.[31] He conducts a number of tests. One test, which extends

[31]It should be noted that it is difficult to distinguish the limited commitment hypothesis from that of efficient incomplete contracts to overcome hold-up when there are exogenous switching

the approach of Beaudry & DiNardo (1991), is to allow both the best and *the worst* realisations of outside opportunities (proxied by unemployment rates) since the start of the employment relationship. Controlling for current outside opportunities, and other observable characteristics of both workers and firms, current wages respond to both the lowest and the *highest* unemployment rates recorded since the start of the employment relationship. The fact that both the best and worst labor market conditions since hiring have a significant impact on current wages suggests that both the worker's and the firm's outside option constraints matter. It should however be noted that Grant (2003) also used the highest unemployment rate in a Beaudry-DiNardo style regression on U.S. data, and found less evidence for its significance, although Devereux & Hart (2005) did find it to be significant in U.K. data.[32]

A second implication of the model is that "cohort effects"—differences between wages for different entry cohorts within a firm—will tend not to persist. The wage intervals will be cohort independent, so that a large change in outside opportunities should eliminate any differences if all cohorts need to "renegotiate" (have binding self-enforcing constraints). Consistent with this, Macis finds that the correlation between the unemployment rate prevailing at the time of hiring and current wages declines with tenure. A further test is based on the following observation: if a worker's wage rose between $t-1$ and t, then according to the model he is constrained at t. This implies an *asymmetric* response to changes in outside options. Suppose first that unemployment rises between t and $t+1$ so that the outside option worsens. This should relax the constraint, and certainly a small change should not imply that the firm's constraint binds, so the wage will be unchanged; a larger rise in unemployment will however cause the firm's constraint to bind and the wage to fall. On the other hand, if unemployment falls, the improvement in the outside option

costs (Macleod & Malcomson (1993); see also Malcomson (1997)).

[32]Grant (2003) finds maximum unemployment to matter in a basic individual fixed effects specification, but not if year and tenure dummies are included (whereas the effect of minimum unemployment is largely robust to these additions; see below).

will further tighten the constraint, pushing up the wage, even if this change is small. This is what Macis finds in the data. However the prediction should also work in the opposite direction when wages fall between $t-1$ and t, so that the firm's constraint can be assumed to be binding, but in this case small increases in unemployment at $t+1$ (which should further tighten the firm's constraint) do not appear to reduce wages.

In Rudanko (2006) the model is calibrated to US data, and the volatilities of real wages and of the vacancy-unemployment ratio are analysed. Not surprisingly, if there is commitment in the wage contract then wages vary too little with productivity (only new matches are responsible for any variability), but even with two-sided limited commitment the model only approximates the empirical correlation between wages and productivity (and simultaneously the corresponding vacancy-unemployment ratio correlation) if the replacement ratio is around 80%, which is considerably higher than usually assumed (although Rudanko argues that this is not necessarily an unreasonable number).

3.3 Hours and wages

Beaudry & DiNardo (1995) use PSID 1976–1989 for male heads of household to estimate the relationship between hours and wages according to the equation

$$\Delta \ln H_{j,\tau+t} = \alpha_1 \Delta \ln w_{j,\tau+t} + \alpha_2 \Delta \ln z_{j,k,\tau+t} + \alpha_3 \Delta X_{j,\tau+t} + \epsilon_{j,\tau+t}.$$

Hours, $H_{j,\tau+t}$ measure annual hours at date $\tau+t$ of worker j hired in year τ, $X_{j,\tau+t}$ measures marital and union status and $\epsilon_{j,\tau+t}$ is the error term. The wage rate, $w_{j,\tau+t}$ are measured in two alternative ways, either as an annual average or as the reported "point in time" estimate from the survey information. The productivity term $z_{j,k,\tau+t}$ is decomposed into into industry specific terms (k denotes the industry), and a quadratic experience and tenure profile for each worker. The equation is estimate in log differences to account for worker specific productivity differences at

The results presented in Beaudry & DiNardo (1995) show a statistically significant negative relationship between hours and wages. A Basemann overidentification test however, shows that the instruments for productivity are not only affecting hours through their effect on wages. Therefore Beaudry & DiNardo restrict data either to non-union contracts or by excluding workers that have recently switched jobs and find that for these subsets the overidentification restrictions are rejected less frequently while the coefficient α_1 remains significantly negative. This is in line with what the contracting model would predict. It is however, important to note that this model is not testing against an alternative. Thus unless assumptions are made about the *long-run* intertemporal elasticity of substitution this cannot be taken as evidence against the spot market model. When the estimates for α_1 are combined with the results of Beaudry & DiNardo (1991) (see above) this suggests that a 2% reduction in unemployment would lead to a 6–8% increase in the wage rate and therefore a 1–2% reduction in hours worked absent changes in productivity. This gives quite plausible numbers for the change hours.

4 Closing Comments

We presented an overview of models of self-enforcing labour contracts in which risk-sharing is the dominant motive for contractual solutions. A basic two-agent (firm-worker) model was developed which is sufficiently general to encompass the problem considered in most of the literature. We then considered some implications for macroeconomics and what empirical support exists for the model. Subsequently we look at the one-sided limited commitment problem in detail and discussed the recent empirical work which is generally supportive of the model.

Here we have concentrated on the implications of the limited commitment model for explaining observed regularities of wages and hours with productivity and unemployment. Another issue is whether the model can help explain observed patterns in wages at the firm level where it is typically found that larger firms pay higher wages and fast growing firms pay lower wages. An approach along these lines combining contracts with credit constraints for firms can be found in Michelacci & Quadrini (2005).

References

Abowd, J. M. & Card, D. (1987), 'Intertemporal labor supply and long-term employment contracts', *American Economic Review* **77**(1), 50–68.

Azariadis, C. (1975), 'Implicit contracts and underemployment equilibria', *Journal of Political Economy* **83**(6), 1183–1202.

Beaudry, P. & DiNardo, J. (1991), 'The effect of implicit contracts on the movement of wages over the business-cycle - evidence from micro data', *Journal of Political Economy* **99**(4), 665–688.

Beaudry, P. & DiNardo, J. (1995), 'Is the behavior of hours worked consistent with implicit contract theory?', *Quarterly Journal of Economics* **110**(3), 743–768.

Boldrin, M. & Horvath, M. (1995), 'Labor contracts and business cycles', *Journal of Political Economy* **103**(5), 972–1004.

Calmès, C. (2007), 'Self-enforcing labour contracts and macroeconomic dynamics', *International Advances in Economic Research* .

Devereux, P. J. (2001), 'The cyclicality of real wages within employer-employee matches', *Industrial Labor Relations Review* **54**(4), 835–850.

Devereux, P. J. & Hart, R. A. (2005), The spot market matters: Evidence on implicit contracts from Britain, IZA Discussion Papers 1497, Institute for the Study of Labor (IZA).

Grant, D. (2003), 'The effect of implicit contracts on the movement of wages over the business cycle: Evidence from national longtitudinal surveys', *Industrial and Labor Relations Review* **56**(3), 393–408.

Hall, R. E. (2005), 'Employment fluctuations with equilibrium wage stickiness', *American Economic Review* **95**(1), 50–65.

Holmstrom, B. (1983), 'Equilibrium long-term labor contracts', *Quarterly Journal of Economics* **98**(3), 23–54.

Krueger, D. & Uhlig, H. (2006), 'Competitive risk sharing contracts with one-sided commitment', *Journal of Monetary Economics* **53**(7), 1661–1691.

Lazear, E. (1981), 'Agency, earnings profiles, productivity, and hours restrictions', *American Economic Review* **71**(4), 606–620.

Macis, M. (2006), Wage dynamics and insurance. Mimeo, University of Chicago.

Macleod, W. B. & Malcomson, J. M. (1993), 'Investments, hold-up, and the form of market contracts', **83**(4), 811–847.

Malcomson, J. M. (1997), 'Contracts, hold-up, and labor markets', *Journal of Economic Literature* **35**(4), 1916–1957.

Malcomson, J. M. (1999), Individual employment contracts, *in* O. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics', Vol. 3, Elsevier, chapter 35, pp. 2291–2372.

McDonald, J. T. & Worswick, C. (1999), 'Wages, implicit contracts, and the business cycle: Evidence from Canadian micro data', *Journal of Political Economy* **107**(4), 884–892.

Michelacci, C. & Quadrini, V. (2005), Financial markets and wages, NBER Working Papers 11050, National Bureau of Economic Research, Inc.

Moen, E. R. (1997), 'Competitive search equilibrium', *Journal of Political Economy* **105**(2), 385–411.

Rosen, S. (1985), 'Implicit contracts - a survey', *Journal of Economic Literature* **23**(3), 1144–1175.

Rudanko, L. (2006), Labor market dynamics under long-term wage contracting and incomplete markets. Mimeo, University of Chicago.

Shimer, R. (2005), 'The cyclical behavior of equilibrium unemployment and vacancies', *American Economic Review* **95**(1), 25–49.

Shin, D. (1994), 'Cyclicality of real wages among young men', *Economics Letters* **46**(2), 137–142.

Shin, D. & Shin, K. (2003), Why are the wages of job stayers procyclical?, Technical report. CIBC Working Paper No. 2003-7, The University of Western Ontario.

Sigouin, C. (2004), 'Self-enforcing employment contracts and business cycle fluctuations', *Journal of Monetary Economics* **51**(2), 339–373.

Solon, G., Barsky, R. & Parker, J. A. (1994), 'Measuring the cyclicality of real wages: how important is composition bias?', *Quarterly Journal of Economics* **109**(1), 1–26.

Thomas, J. P. & Worrall, T. (1988), 'Self-enforcing wage contracts', *Review of Economic Studies* **55**(4), 541–554.