

IZA DP No. 5503

**Minu, Startu and All That:
Pitfalls in Estimating the Sensitivity of a
Worker's Wage to Aggregate Unemployment**

Pedro Martins
Andy Snell
Jonathan Thomas

February 2011

Minu, Startu and All That: Pitfalls in Estimating the Sensitivity of a Worker's Wage to Aggregate Unemployment

Pedro Martins

*Queen Mary, University of London
and IZA*

Andy Snell

University of Edinburgh

Jonathan Thomas

University of Edinburgh

Discussion Paper No. 5503
February 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

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

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

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

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

ABSTRACT

Minu, Startu and All That: Pitfalls in Estimating the Sensitivity of a Worker's Wage to Aggregate Unemployment

In this paper we show that panel estimates of tenure specific sensitivity to the business cycle of wages is subject to serious pitfalls. Three canonical variates used in the literature – the minimum unemployment rate during a worker's time at the firm ($\min u$), the unemployment rate at the start of her tenure (Su) and the current unemployment rate interacted with a new hire dummy (δu) – can all be significant and “correctly” signed even when each worker in the firm receives the same wage, regardless of tenure (equal treatment). In matched data the problem can be resolved by the inclusion in the panel of firm-year interaction dummies. In unmatched data where this is not possible, we propose a solution for $\min u$ and Su based on Solon, Barsky and Parker's (1994) two step method. Our proposed solution method is however suboptimal because it removes a lot of potentially informative variation in average wages. Unfortunately δu cannot be identified in unmatched data because a differential wage response to unemployment of new hires and incumbents will appear under both equal treatment *and* unequal treatment.

JEL Classification: J50, J31, C18

Keywords: wage cyclicality, unemployment

Corresponding author:

Pedro S. Martins
School of Business and Management
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
E-mail: p.martins@qmul.ac.uk

1 Introduction and overview

There has been a recent upsurge in interest in the relationship between the tenure of a worker and the sensitivity of her wages to the business cycle. Despite a burgeoning empirical literature many issues in this area remain controversial. In particular arguments still persist about the extent to which the wages of new hires are more sensitive to current business cycle conditions than those of incumbents (See for example Baker, Gibbs and Holmstrom, 1994. who find they are different and Gertler and Tregari, 2009 who find they are not). More generally, others have investigated the general relationship between a worker's pay and the state of the business cycle during her tenure with the firm. Establishing reliable empirical stylised facts about these issues is crucial for macroeconomic theories of wage setting. A popular way to investigate how the sensitivity of wages to the business cycle varies with a worker's tenure is the inclusion of tenure related cyclical variates in standard Mincer wage equations. Three canonical examples of such variates are a) the minimum unemployment rate during a worker's tenure, " $\min u$ ", b) the unemployment rate of a worker at the start of his tenure, " Su ", and c) the current unemployment rate interacted with a new hire dummy, " δu ".¹ Henceforth we refer to variates such as these as unemployment-tenure interactions or UTI's for short. We argue in this paper that drawing inferences from the significance of UTI variates has serious pitfalls. In particular we show that they may be significant and "correctly" signed even when the wages of workers within a firm are equally sensitive to the business cycle *regardless of tenure*. Referring to the latter situation as equal treatment - our generic null hypothesis - we show analytically and numerically that under a number of plausible equal treatment models these three variates will be significant with a sign that would lead the investigator to find falsely in favour of a model based on unequal treatment contracts (forged via bilateral firm-worker bargaining) rather than equal treatment contracts (usually but not necessarily forged via collective firm-workforce bargaining). The problem - essentially one of endogenous tenure - arises because the average UTI for a firm embeds information on its current and past hiring decisions which, under equal treatment, may be correlated with that firm's wage level. A solution to the problem is to include firm-year interaction dummies to absorb firm specific wage components. If this is done, UTI variates will only be significant if the sensitivity of wages to the business cycle actually does vary with tenure - our generic alternative hypothesis. Unfortunately this cure is not always available because many panel datasets do not match workers to firms, the PSID being a classic case in point. We argue that without matched data it is impossible to identify asymmetric responses to unemployment of wages of new hires and incumbents - the case we are calling δu here. However for $\min u$ and Su Solon, Barsky and Parker's (1994, henceforth SBP) two step estimator may be adapted to control for the biases induced by the existence

¹ δu is not commonly used as a regressor directly in panel regressions although Gertler and Tregari (2009)'s estimates recently attracted some attention. It is more common to compare directly the properties of average wages of new hires with incumbents - something we do in the final section of this paper. Other examples of UTI's are the maximum unemployment rate since joining the firm, the maximum change in unemployment since joining the firm (see Macis, 2009) and the product of a tenure measure and unemployment (see Arozamena and Centen (2006)). Extensions of the analysis to these and other UTI's should be obvious.

of common firm specific wage components. Using the panel dimension to control for worker characteristics, SBP (and subsequently Shin and Shin, 2003 and Devereux and Hart, 2007) extract composition bias.free estimates of mean wages for different worker tenures in each time period. These data are then used to form a new time/tenure panel to investigate the business cycle sensitivity of wages across different tenures. ² We propose adding extra regressors to SBP's second stage regression to annihilate the biases in UTI estimates caused by equal treatment. However our proposed method is clearly inferior to adding firm-year interaction dummies to the original panel. Not only does it remove much of the cross tenure variation in wages but to work effectively it also requires normalised covariances between firm hiring and firm wages to be constant across the business cycle. We close the paper with a small empirical illustration from the PSID. In the application negative estimates of UTI effects from the panel dimension change sign and become insignificant when we apply the modified SBP method.

We emphasise at this point what this paper does *not* say. We do not argue that the tenure related cyclical effects found so far in the literature (in particular, the significantly negative coefficients found on $\min u$, Su and δu) are *necessarily* spurious. Instead the paper makes the important methodological point that UTI's *may* be spuriously significant and "correctly" signed. Furthermore it is quite likely that a large economy will be characterised by bargaining practices that vary from sector to sector. Some sectors could be characterised by equal treatment contracting whereas others could be characterised by unequal treatment (see for example Kilponen and Santavirta, 2010, who find variations in the importance of different contract mechanisms across different sectors of the Finnish economy). If this is the case, our results would also indicate that estimates of tenure specific cyclical effects may be biased rather than simply spurious. Whether or not this bias is upwards (towards zero) or downwards will depend on the nature of firm level bargaining in the sectors that are subject to equal treatment. For example in this paper we identify a number of equal treatment models that generate spurious negative coefficients on $\min u$. Even if these types of models are only relevant in a portion of the economy the coefficient on $\min u$ will still be downward biased. This would lead the investigator to an exaggerated view of the quantitative importance in the economy as a whole of the contracting environment that $\min u$ was designed to test for. Whilst it makes sense to focus this paper on equal treatment models capable of generating the negative signs on UTI's that we see in the empirical literature, we acknowledge that negative signs are not generic - other models will generate positive coefficients on UTI's. In such cases an exactly converse argument could be invoked - namely that the quantitative importance of the relevant unequal treatment contracting mechanism could be underestimated. Whatever the case, it is essential in these empirical exercises to correct the biases to get an accurate take on the quantitative importance (or not) of the relevant unequal treatment bargaining mechanism that is being tested by the particular UTI.

²The method - asymptotically equivalent to panel estimation when weighted least squares is used in the second step - was originally advanced to circumvent the large biases in standard errors that arise when the RHS (macro) variables have variation that is only a tiny fraction of the dependent variable (wages) - see Moulton(1990).

Second, it is well known that any variate that is correlated with a worker's tenure such as $\min u^3$ will also be potentially correlated with wages if human capital accrues through job experience. As pointed out by other authors, if tenure related human capital is not adequately controlled for, variates such as $\min u$ could be significant in Mincer equations even in the absence of tenure specific business cycle effects. The modus operandi of this effect, however, is completely different to ours and to emphasise this point we show UTI's will be significant even in the absence of tenure related human capital. Having said this, tenure is nearly always included in Mincer equations and its inclusion will affect the biases on UTI's. To assess this, we examine the impact of adding tenure measures to Mincer equations in numerical simulations at the end of the paper.

Thirdly and in a similar vein to the human capital argument, Hagedorn and Manovskii (2010) argue that $\min u$ and Su are significant because they proxy for unobserved match quality in a market clearing model with on the job search. They propose new proxies for match quality and argue that including these in a wage equation drives out the significance of $\min u$ and Su . Once again the modus operandi of their effect is completely different to ours and our results obtain in a world without unobservable match quality. Furthermore, in 4.6 below we argue that two of Hagedorn and Manovskii's newly proposed match quality proxies may themselves be spuriously significant in models of equal treatment even where workers always have *identical* match productivity and labour markets do not necessarily clear. In our paper, the potential spurious significance of tenure related cyclical variates is generated by the cross sectional correlation of firm wages with firm hiring decisions rather than via differences in human capital or match quality across workers.

The paper is organised as follows. Section 2 overviews the literature - theoretical and empirical - of wage setting in relation to the business cycle. Emphasis here is on the distinction between models that are founded on unequal treatment versus those founded on equal treatment. The former are necessarily founded on firm-worker bilateral bargaining whilst the latter are often - but not always - founded on firm-workforce collective bargaining. In section 3 we expose the main point made by the paper via a simple illustrative model. In section 4 we derive the properties of (pooled) panel regression estimates of UTI's under a generic alternative hypothesis of equal treatment within the firm. In this section we sharpen the main findings by assuming that wages and employment depend only on firm specific idiosyncratic shocks and hence display no aggregate business cycle. To avoid singularity of some of the regressions we assume that aggregate labour supply and hence the aggregate unemployment rate are variable. Despite the absence of a business cycle in aggregate wages and employment, estimated UTI coefficients are asymptotically nonzero and often take the expected negative sign. Also in this section we offer a digression which suggests that Hagedorn and Manovskii's match quality proxies may themselves be spuriously significant under equal treatment models where match quality is completely absent and where markets do not necessarily clear. In Section 5 we run simulations to quantify the estimated spurious effects in the more realistic setting of both aggregate and idiosyncratic shocks. We find that under

³It is easy to show (we do so below) that variates like $\min u$ can be re-written as a linear combination of tenure dummies.

several plausible parameter scenarios in two equal treatment models, the UTI estimates have a similar order of magnitude to those found in the empirical literature. Section 6 discusses SBP's method and its extensions used by Devereux and Hart(2007) .We show that these methods do not eliminate the problem. However if scaled cross sectional wage-employment covariances are acyclical, time t averages of composition-free wages can be used to obtain estimates of UTI's that are zero under the null of equal treatment and consistent under the alternative. A small empirical application to the PSID in this section shows that applying our method reverses the initially "correct" signs of initial panel based UTI coefficient estimates.

2 Models of wage formation and the business cycle

Much of the current theoretical macro literature on wage formation focuses on models where individual workers bargain with a firm bilaterally and independently of existing contracts that exist within that firm. Classic and vintage examples of these bilateral contracting models are the implicit contract models of Beaudry and Dinardo (1991 - henceforth BDN) and a host of search theoretic models that grew (and are growing) out of Mortensen and Pissarides' (1994) seminal paper (e.g. Cahuc, Postel-Vinay and Robin, 2006)⁴ In these models wages at time t are affected by the state of the economy (or more specifically the level of firm labour productivity) at the time of entry into the firm and may also depend on the state of the economy subsequent to that date. Hence the current level of an individual's wages is determined by the state of the business cycle - usually measured as the aggregate unemployment rate - at the start of and during his tenure.

There is, however, another class of contracting models where, for a given level of human capital be it firm or worker specific, each worker within the firm is paid the same wage. In these "equal treatment" models the wage may vary over the business cycle, but crucially is independent of a worker's tenure (again, modulo human capital). These models imply equal treatment in the sense that no matter how bad(good) current economic conditions are, new workers are not offered lower (higher) wages than incumbents. Classic and vintage examples are the efficiency wage models of Shapiro and Stiglitz(1984) and its variants and insider-outsider models such as that of Blanchard and Summers(1986). More recent examples are search theoretic models with a) staggered contracting (Gertler and Trigari, 2009), b) wage norms (Hall, 2005), c) bargaining over the marginal surplus under diminishing returns to labour (Elsby, 2010) and d) market clearing but with idiosyncratic unobserved match quality (Hagedorn and Manovskii, 2010). Finally the contracting models of Snell and Thomas(2010)

⁴In most of these models, constant returns to scale implies that the economy contains "jobs" not "firms". To take the models to the data where firms obviously do exist requires us to think of each firm as housing a number of jobs each with a wage determined by the bilateral bargain struck between the worker and the firm at the time of the job's inception. With firms so defined, the model predicts wage dispersion within firms even across workers of identical human capital.. Under equal treatment, however, wage dispersion within the firm can only occur via differences in human capital something that we abstract from in this paper.

and Martins, Snell and Thomas(2005,2010) build in equal treatment within the firm at the outset.

Whilst many of the bilateral contract models are assessed via their abilities to reproduce the salient moments of the relevant macro data (such as employment,wage and vacancy variability over the business cycle) companion empirical work tests the theory at hand by examining the significance of some tenure specific cyclical variable, typically a UTI.such as $\min u$. Significance of these variates when they are included in standard panel wage (Mincer) equations is construed as being supportive of *both* unequal treatment *and* of the particular type of bargaining that the variate was designed to capture. For example in one version of BDN's bilateral contract model, wages of new hires are synchronised with the state of the cycle at the time of joining the firm but because workers are mobile, wages must rise as the labour markets tighten in order to retain the worker. By contrast when the labour market slackens, the insurance implicit in the contract prevents workers' wages from falling. In an extension to their basic model (where they add an alternative to formal employment that displays aggregate diminishing returns), BDN show that the minimum unemployment rate since the worker joined the firm or " $\min u$ " for short is a sufficient statistic for his wages. The significance of $\min u$ in their empirics therefore, is taken as evidence against equal treatment and in favour of the specific form of bilateral bargaining embodied in their model. Another variant of the BDN model assumes worker commitment via costly labour mobility. In this world it is unemployment at the start of tenure that determines the worker's wage so that Su and not $\min u$ is the relevant variate. They also test a spot market model whereby u itself (the current unemployment rate) is the only relevant variable.Using data from the CPS and PSID they find $\min u$ dominates both Su and u . Subsequent empirical papers by Mcdonald and Worswick(1999) and Grant(2003) have found similar results with $\min u$ being by far the most robustly significant and correctly (negatively) signed of the three.

In a similar vein adherents of the Mortensen and Pissarides (henceforth MP) modelling approach measure the extent to which wages of new hires and incumbents differentially respond to current economic conditions. Adding u and δu (unemployment and an unemployment-new hire dummy interaction term) to a wage equation would help establish the extent of (if any) the differential response of new hire versus incumbent wages to current economic conditions. Finding such a differential would provide support for the bilateral contracting in the model. It would also aid the calibration of the model by quantifying the sensitivity of the bargained wage to current economic conditions (the worker's outside option). Gertler and Trigari (2009).extend the Mortensen and Pissarides model to allow for staggered contracts but they assume that devising new contracts for new hires incurs costs so that all wages within the firm are adjusted together - in short they assume equal treatment. In their companion empirical work they add u and ∂u .to a standard Mincer equation and find that after controlling for spell fixed effects ∂u is insignificant. They conclude that the wages of new hires have the same exposure to the business cycle as do those of incumbents.

Further examples of papers that include UTI's in Mincer equations include:- Montuenga, Garcia and Fernandez(2006), who add $\min u$ to an otherwise standard wage curve for a group

of EU countries, Schmieder and von Wachter (2010), who extend BDN's analysis to test for equality of $\min u$ coefficients between two consecutive work spells, Hartog, Opstal and Teulings(1997), who use UTI's to analyse inter industry wage differentials, Bertrand(2004) and Kilponen and Santavirta (2010) who use UTI's to assess the effects on wages of import competition, Arozamena and Centeno(2006) who interact unemployment with a tenure measure to allow for cyclicalities to vary with tenure, Vilhubert(1999) who uses UTI's to assess wage flexibility in Germany and Bell, Nickell and Quintini(2000) who add UTI's to an otherwise standard wage curve. Authors using SBP' method to estimate the importance of UTI's include Shin and Shin (2003) and Devereux and Hart(2007).

3 A simple illustrative model

In this section we fix ideas and intuition for our main results by analysing a simple equal treatment model. In keeping with the analytical results in the first half of this paper we work with one cross section at time t and abstract from the business cycle by assuming that all shocks to wages and employment are firm specific and idiosyncratic. Whilst this implies that average firm wages and employment are constant it leaves unspecified the time series properties of labour supply and hence of aggregate unemployment.⁵

In practice it would be foolish to try and identify the effect of the business cycle on wages using a single cross section and it would be impossible to do so when no business cycle is present. But *attempting* to do so illustrates our main point:- even in a world with no business cycle and where there is equal treatment we may still get significant UTI estimates. The modus operandi of the effect we identify in this paper is that the significance of the coefficient on the UTI arises from cross sectional (more specifically cross-firm) wage variation rather than from its time series correlation with current and past levels of unemployment. Later in the paper we show - again in the absence of a business cycle - that the results on the signs of biases from a single cross section extend to those obtained from a full panel.

We build this example around two key stylised facts of labour markets, namely that, controlling for firm and worker characteristics, larger firms pay more and have higher labour retention rates (lower turnover).⁶ Explicitly we have "low" and "high" firms. Low (high) firms have low(high) wages and low(high) retention rates. Whilst firm size is irrelevant to the model, it would be natural to think of the high wage firms as being large and low firms being

⁵Note that the orthogonal complement of $\partial_{ijt}^0 u_t$ namely $(1 - \partial_{ijt}^0)u_t$ or the unemployment rate (u_t) itself should also be included in order to be able to assess differential effects of the business cycle on new hires. It is easy to show that in the absence of business cycles - the initial scenario under which we operate - that omitting either term is innocuous. And of course in the single cross section we have here u_t is absorbed into the intercept.

⁶Lallemand et al (2003) estimate that in some EU countries a doubling of firm size - ceteris paribus - raises wages by around 5% and in the Survey of Consumer Finances, firms with <100 employees have an average turnover rate greater than 40% whilst for those with more than 100 employees the average rate is around 20% (Even and Macpherson,1996).

small. We analytically determine the signs of the three UTI coefficients and calibrate the model to get numerical values for them.

Suppose that a firm either pays high wages ($w_{ijt} = w^h$) or low wages ($w_{ijt} = w^l$) and that proportion $p^h(p^l)$ of time t 's labour force work in high(low) wage firms. High firms are assumed to have retention rate s_h which exceeds the rate for low firms (s_l).

To derive a form for the UTI coefficient for $\min u$ ($\hat{\beta}_{\min u}$) we could simply treat the group of firms paying high wages (and having low labour turnover) as a single high wage "firm".and do likewise for the low wage sector. Employment in both sectors (L^h and L^l) is assumed to be constant and this makes the tenure structure very simple. For the two sectors (indexed by $i = h$ and $i = l$) the number of tenure k workers surviving at time t ($L_t^{i(k)}$) is just

$$\begin{aligned} L_t^{i(k)} &= s_i^k (L_{t-k}^i - s_i L_{t-k-1}^i) \\ &= s_i^k \{(1 - s_i) L^i\} \quad i = h, l \end{aligned}$$

Each worker of tenure k at time t will have the same $\min u$.value so the average $\min u$ in the high wage sector (m_t^h) and in the low wage sector (m_t^l) is

$$\begin{aligned} m_t^i &= \left\{ \sum_{k=0}^{\infty} s_i^k (L_{t-k}^i - s_i L_{t-k-1}^i) u_{t-k}^m \right\} / L_t^i \\ &= (1 - s_i) \sum_{k=0}^{\infty} s_i^k u_{t-k}^m \quad i = h, l \end{aligned} \tag{1}$$

Note that by replacing u_{t-k}^m with u_{t-k} we get an exact formula for the Su case and we denote this as $S_t^h(S_t^l)$ for high(low) wage firms.

We can rewrite the expression for m_t^i more informatively as

$$m_t^i = u_t - \sum_{k=1}^{\infty} s_i^k (u_{t-k+1}^m - u_{t-k}^m) \quad i = h, l$$

In this last expression, the term in braces is always weakly positive and because $s_l < s_h$ it follows that $m_t^h < m_t^l$

Using some tedious OLS arithmetic we can now show that the three coefficient estimates for our UTI's are

$$\widehat{\beta}_{\delta u} = \frac{p}{\{(1 - s^h) + (1 - s^l p)\}\{s^h + s^l p\}u_t} \{(s_h - s_l)(w^l - w^h)\}$$

$$\widehat{\beta}_{\min u} = \frac{p^h p^l (m_t^h - m_t^l)(w_t^h - w_t^l)}{p^h m_t^{h(2)} + p^l m_t^{l(2)} - (p^h m_t^h + p^l m_t^l)^2}$$

where $m_t^{i(2)} = (1 - s_i) \sum_{k=0}^{\infty} s_i^k (u_{t-k}^m)^2 \quad i = l, h$

$$\widehat{\beta}_{S u} = \frac{p^h p^l (S_t^h - S_t^l)(w_t^h - w_t^l)}{p^h S_t^{h(2)} + p^l S_t^{l(2)} - (p^h S_t^h + p^l S_t^l)^2}$$

where $S_t^i = (1 - s_i) \sum_{k=0}^{\infty} s_i^k u_{t-k} \quad i = l, h$

$$S_t^{i(2)} = (1 - s_i) \sum_{k=0}^{\infty} s_i^k u_{t-k}^2 \quad i = l, h$$

Because $s_l < s_h$ and $m_t^h < m_t^l$ both $\widehat{\beta}_{\delta u}$ and $\widehat{\beta}_{\min u}$ are negative for any sequence of realisations of aggregate unemployment $u_t, u_{t-1}, u_{t-2} \dots$. The sign of $(S_t^h - S_t^l)$ will however depend on the realisations for aggregate unemployment so the sign of $\widehat{\beta}_{S u}$ cannot be determined.

To get a feel for numerical values we might expect from a cross section estimation ⁷we conduct a simple and crude calibration exercise based on data from the US economy. Below are data from the US Census Bureau on private sector employment by firm size.

<i>FirmSize</i>	1 - 4	5 - 9	10 - 19	20 - 99	100 - 499	5000 - 9999
<i>Employees(m)</i>	5.8	6.9	8.5	20.6	16.8	6.4
	500 - 749	750 - 999	1000 - 1499	1500 - 2499	2500 - 4999	10000+
	3.5	2.3	3.4	4.4	6.0	30.5

Using this data to rank employees by the size of the firm that they work in then the median worker's firm size is about 300 employees. Following on from above, we could label those workers working in firms of less than 300 employees "l" type and those above as "h" type. In this case, $p^l = p^h = .5$. Assuming a wage size premium elasticity of 5% ⁸ then, if the above firm size distribution applied for regardless of worker skill and industry sector, the wage premium for the above model i.e. $w_t^h - w_t^l$ would be about 40%. We set s_l and s_h equal to .6 and .8 respectively. If our reference points are the average firm size for "l"

⁷It is easy to show that in our acyclical world estimates from the full panel are a (positive) weighted average of the cross sectional estimates.

⁸Studies by Lallemand, et.al (2003) for European economies and by Oi and Idson (1999) find elasticities in the range 0 to 10% with an average estimate near to 5%.

and "h" category workers respectively, .6 and .8 are roughly consistent with the 1988-91 NLSY data in Even and Macpherson,(1996). Finally to calibrate $S_t^i, S_t^{i(2)}, m_t^i$ and $m_t^{i(2)}$ we use the realisations of annual US unemployment since 1948.and set u_t in the δu formula to 5%. These calibrated values give estimates of $\widehat{\beta}_{\min u}, \widehat{\beta}_{\delta u}$ and $\widehat{\beta}_{Su}$ of $-9.98, -2.14,$ and $-.78$ respectively. The value for $\min u$ is higher than those found in the literature by BDN, and others where the average estimate is around -5.0 but the numbers for Su and δu are a similar order of magnitude to estimates found in empirical work. Whilst the model and its calibration represent a rather crude caricature of the salient stylised facts of US labour markets, the exercise does at least show that the effect we identify in this paper is potentially quantitatively important.

The above shows that even in the absence of a business cycle, the three UTI's will be significant in panels and may have negative sign. It is easy to show that adding worker fixed effects does not cure the problem. By contrast adding firm fixed effects will fix the problem. It will annihilate the problematic firm specific wage components and yield consistent estimates of the β 's.⁹ However in a more general stochastic model firm wages and employment would be subject to firm specific and aggregate shocks (as in sections 5 and 6 below). In that scenario firm fixed effects would no longer remove firm specific wage components. Instead we would need firm-year interaction dummies to remove these components.

4 Estimates of UTI effects under equal treatment

In this section we expose analytically the behaviour of estimates of our three UTI variates under equal treatment within the firm. We derive our results under a single fixed economy wide retention¹⁰ rate to allow us to obtain closed form solutions for estimates etc. The formulae are easily adapted to allow for m possible retention rates ($s_i, i = 1, 2..m$) by grouping the firms into sectors each of which corresponds to a fixed s value.We do this for the simple case of $m = 2$ i.e. an economy with high and low s sectors with the high s sector having a high mean wage and high mean employment and vice versa for the low s sector. In this analysis it is important that the designation is fixed over time and independent of the shocks.that impinge on firms. This is consistent with a world where shocks that affect employment and wages in high and low firms are temporary and small relative to the difference in mean wages and mean employment between high and low firms. Hence, despite suffering idiosyncratic shocks over time, large high wage/high retention rate firms do not become small, low wage/low retention rate firms and vice versa.¹¹

⁹Under the alternative where wages are linear in each respective UTI, adding firm fixed effects is innocuous.

¹⁰In this paper we prefer to deal with worker retention rates rather than labour turnover rates. The latter is of course one minus the former.

¹¹The literature on the wage size premium focuses on both differences in plant and firm size. Obviously over a long period of time, both plant size and firm size can grow in size. So our fixed assignation of firms into small and large size is consistent with the relatively small T assumption of the paper.

The main aim in this section is to establish conditions on the cross section covariance of firm/sector wages and firm/sector employment under which these estimates have a non zero and negative probability limit. The plan is to start by analysing a single cross section and then to establish results for the full panel afterwards. As noted above we sharpen and simplify our analytical results by abstracting from a business cycle - all shocks are idiosyncratic rather than aggregate. The effects of allowing for a business cycle in wages and employment are considered via numerical simulations in section 5.

We assume we have a complete sample of workers in n firms which constitute the economy. Of course few datasets will be anything like this comprehensive (although the QP from Portugal approaches this). In subsection 4.8, we discuss the effects of random sampling of only proportion p of the workforce in the economy, and show that although this complicates the analytical details it does not change the central results, as long as the number of firms being sampled is large. As stressed above this paper abstracts from human capital. Our equal treatment hypothesis is that workers within a firm receive the same wage, up to an (worker specific) idiosyncratic shock.¹² Finally we assume that the retention rate in each firm is exogenous and is sufficiently low to avoid the firm having to make layoffs. Allowing for layoffs would introduce nonlinearities which would seriously confound the analysis, but we do not believe it is central to our results.¹³

4.1 OLS estimates of β in a single cross section under equal treatment

In what follows we consider the regression of wages on (an intercept and) a single UTI - hence we deal with each of our three UTI's separately and in turn. The global aim is to derive results for full (pooled) panel estimation over time periods $t = 1, \dots, T$, firms $j = 1, \dots, n$ ¹⁴ and individuals $i = 1 \dots L_{jt}$ within those firms, but as noted above our no business cycle assumption allows us to deal with a single cross section for the current purpose. We therefore estimate for a single time period t

$$w_{ijt} = \alpha + \beta c_{ijt} + error_{ijt} \quad (2)$$

where w_{ijt} is log of wages of individual i in firm j at t and c_{ijt} and $error_{ijt}$ are that individual's UTI cyclical variable and error (both to be specified) respectively.

¹²This may be measurement error and were we to ease our assumptions to allow for "idiosyncratic" (i.e. uncorrelated across workers and uncorrelated with tenure and macro variates) human capital, it could be that also.

¹³We should note that the average rate of annual firm level labour force turnover in the US is high - about 30% - although we admit that not all of this will be due to worker quits. One way of defending our assumption of no layoffs is by saying that the results only apply in data where adverse shocks to the firm are not too severe.

¹⁴Here n is assumed to be fixed across time but this is merely a notational simplification. The analysis whereby n is time subscripted would merely require the additional assumption that $\min(n_1, n_2, \dots, n_T) \rightarrow \infty$ for the asymptotics to carry through.

We focus on three specific cases for c namely the aggregate unemployment rate times a new hire dummy, $c_{ijt} = \delta_{ijt}^0 u_t$ ("δu") the minimum aggregate unemployment rate seen by worker i at time t since he/she joined firm j , $c_{ijt} = \min u_{ijt}$, ("min u") and the aggregate unemployment rate at the start of worker i 's tenure at firm j , $c_{ijt} = Su_{ijt}$ ("Su"). It should become clear that the analysis could be extended quite easily to other UTI variates such as "max u" the maximum unemployment rate since a worker joined the firm (relevant where there is one-sided (worker) commitment). A significantly negative estimate of β is typically interpreted by the investigator as support for the existence of the relevant form of bilateral contracting.

Of course (2) is not a proper regression equation but is merely a statement of what the investigator is estimating. Suppose now that (2) is in fact a misspecification in the sense that w_{ijt} is not *directly* related to c_{ijt} . Instead wages are equal to a firm specific component plus worker specific shock i.e.

$$w_{ijt} = w_{jt} + v_{ijt} \quad (3)$$

$$E(v_{ijt}, c_{ijt}) = 0 \quad E(v_{ijt}, w_{ijt}) = 0 \quad (4)$$

This equation makes clear what we mean by equal treatment - differences in wages may exist but these differences must not be correlated with UTI's. In adopting (2) we have ignored education and worker tenure as regressors whilst in the literature they are typically included. Excluding the former is innocuous in the absence of human capital but excluding worker tenure is not - tenure is manifestly correlated with c_{ijt} and adding it to the regression will change the estimates of the UTI parameters. The effect of adding tenure to the regression in (2) is taken up in the numerical simulations in section 5 below.

The regression estimate β for a single cross section of L_t workers is

$$\hat{\beta} = \frac{1}{\text{svar}(c_{ijt})} (\text{scov}(w_{ijt}, c_{ijt})) \quad (5)$$

where $\text{scov}(\cdot)$ and $\text{svar}(\cdot)$ are sample covariance and variance respectively. Later, we extend the results to a panel where T the number of time periods is fixed and small relative to the number of firms n . With this in mind we now analyse the sign of $\hat{\beta}$ as n goes to infinity.

The denominator in (5) is always positive so we can focus exclusively on the sign of the numerator.

Proposition 1:- The numerator in (5) can be written as

$$\text{scov}(w_{ijt}, c_{ijt}) = \frac{1}{L_t} \left(\text{scov}^f(w_{jt}, c_{jt}) - \text{scov}^f(L_{jt}, w_{jt}) \cdot \frac{1}{L_t} \sum_{j=1}^n c_{jt} \right) + o_p(1) \quad (6)$$

where $scov^f$ denotes sample covariance across firms $j = 1, 2, \dots, n$ at time t . rather than across individuals and where, in the absence of aggregate shocks to firm employment, $p \lim \bar{L}_t = \bar{L}$ is constant over time.

Proof: - See appendix

Equation (6) is important. It shows us that under an alternative hypothesis of equal treatment where wages and employment are acyclical, $\hat{\beta}_{\min u}$, $\hat{\beta}_{Su}$ and $\hat{\beta}_{\delta u}$ will in general not be zero and will take values that depend on the cross firm covariance of wages (w_{jt}) with the sum of the UTI values of workers in the firm (c_{jt}). The latter will be a weighted average of current and past employment levels of the firm where the weights are identical across firms. For example in the case of δu and where the rate of labour turnover $1 - s$ is fixed across firms, c_{jt} will be just $(L_{jt} - sL_{jt-1})u_t$. Therefore, in models where the firm's wage policy (w_{jt}) depends on current and past labour force levels, the cross firm correlation of w_{jt} and c_{jt} will in general be nonzero even though δu is by assumption irrelevant to the wage policies of firms.

We develop further the above expressions for specific choices of c_{ijt} namely, $\min u$, Su and ∂u . We then discuss the signs of the probability limits of the respective regression coefficients ($\hat{\beta}_{\min u}$, $\hat{\beta}_{Su}$ and $\hat{\beta}_{\delta u}$). in an economy that has firms with identical mean wages, mean employment and retention rates. We then extend the results on sign to cases of heterogenous mean employment, mean wages and retention rates.

4.2 Minimum unemployment rate during tenure:- $\min u$

We start by developing expressions for $c_{jt} = \min u_{jt}$ (the "aggregate" $\min u$ within firm j).

The c_{ijt} variate for the $\min u$ case is a tenure dummy for worker i multiplied by the minimum unemployment rate associated with her length of tenure. The sum of within-firm tenure dummies for any entry date k is

$$\partial_{jt}^k = \sum_{i=1}^{L_t} \partial_{ijt}^k \equiv L_{jt-k}^t - L_{jt-k-1}^t \quad (7)$$

where $\partial_{ijt}^k = 1$ if worker i is of tenure k and $\partial_{ijt}^k = 0$ if not. The "aggregate" $\min u$ within a firm ($\min u_{jt}$) will be related to past hiring and the cohort composition of the current

labour force as follows

$$\min u_{jt} \left(= \sum_{i=1}^{L_{jt}} \min u_{ijt} \right) = \sum_{k=0}^{\infty} \partial_{jt}^k u_{t-k}^m = \sum_{k=0}^{\infty} (L_{jt-k}^t - L_{jt-k-1}^t) u_{t-k}^m \quad (8)$$

$$= \sum_{k=0}^{\infty} s^k (L_{jt-k} - sL_{jt-k-1}) u_{t-k}^m \quad (9)$$

Following the lead of the analysis in section 3 above we can collect terms differently to get a different and more useful form for this expression as

$$\min u_{jt} = L_{jt} u_t - \sum_{k=1}^{\infty} s^k L_{jt-k} (u_{t-k+1}^m - u_{t-k}^m) \quad (10)$$

Summing across firms and dividing by the number of workers L_t gives the time t average $\min u$ as

$$\frac{1}{L_t} \sum_{j=1}^n \min u_{jt} = u_t - \frac{1}{L_t} \sum_{k=1}^{\infty} \sum_{j=1}^n s^k L_{jt-k} (u_{t-k+1}^m - u_{t-k}^m) \quad (11)$$

$$= u_t - \sum_{k=1}^{\infty} s^k \frac{L_{t-k}}{L_t} (u_{t-k+1}^m - u_{t-k}^m) \quad (12)$$

Setting $c_{jt} = \min u_{jt}$ in (6) and then using (10) and (12) in (6) gives a value for $p \lim \widehat{\beta}$ for the $\min u$ case as

$$p \lim \widehat{\beta}_{\min u} \propto - \sum_{k=1}^{\infty} s^k (\gamma_k - \gamma_0) (u_{t-k+1}^m - u_{t-k}^m) \quad (13)$$

where $\gamma_k = p \lim \text{scov}^f(L_{jt-k}, w_{jt})$ which - in keeping with the assumption of acyclical firm employment and firm wages - is assumed to be time invariant and where we have used the fact that in the absence of aggregate shocks $p \lim \frac{L_{t-k}}{L_t} = 1$. Here and henceforth the symbol \propto means "positively proportional to".

It may be more convenient sometimes to deal with the log of employment rather than its level. Many models of the labour market do so. It would be useful then to derive an analogue to (13) for covariances of logs. If (log) wages and firm employment are normally

distributed ¹⁵ with time invariant unconditional means and variances then $cov(w_{jt}, L_{jt-k}) = c^+ cov(w_{jt}, l_{jt-k})$. where $c^+ > 0$ and is independent of k . Using this (13) becomes

$$p \lim \widehat{\beta}_{\min u} \propto - \sum_{k=1}^{\infty} s^k \{\gamma_k^* - \gamma_0^*\} (u_{t-k+1}^m - u_{t-k}^m) \quad (14)$$

where $\gamma_k^* = cov(l_{jt-k}, w_{jt})$. $k = 0, 1, 2$,

Now note that $u_{t-k+1}^m - u_{t-k}^m$ is always by definition non negative. We can see from (13) therefore that if γ_0 is always negative and if it is also larger than or equal in absolute value to γ_k (for $k = 1, 2, \dots$), then $\widehat{\beta}_{\min u}$ will be negative. If the γ_k ($k > 0$) are all weakly positive then all we need is that γ_0 be negative. By contrast if γ_0 and γ_k are both positive then $p \lim \widehat{\beta}_{\min u}$ is only guaranteed to be negative if $\gamma_k > \gamma_0$ for all $k > 0$, something that is unlikely to be true in practice or that is unlikely to be a theoretical property of a model. However, given that s is below unity then the lead term may well dominate the sum in (13) or (14). In that case we would require just $\gamma_1 > \gamma_0$. We could repeat these arguments for (14) and develop identical conditions for γ_k^* in place of γ_k to determine the sign of $p \lim \widehat{\beta}_{\min u}$

4.3 Unemployment rate at start of tenure:- Su

For Su we could repeat the analytical steps used for $\min u$ but replacing terms in $u_{t-k+1}^m - u_{t-k}^m$ in (13) and (14) with $u_{t-k+1} - u_{t-k}$. This gives the analogue form of (13) and (14) as

$$p \lim \widehat{\beta}_{Su} \propto - \sum_{k=1}^{\infty} s^k (\gamma_k - \gamma_0) (u_{t-k+1} - u_{t-k}) \quad (15)$$

$$p \lim \widehat{\beta}_{Su} \propto - \sum_{k=1}^{\infty} s^k \{\gamma_k^* - \gamma_0^*\} (u_{t-k+1} - u_{t-k}) \quad (16)$$

Whereas $u_{t-k+1}^m - u_{t-k}^m$ in (13) and (14) is always positive, the sign of $u_{t-k+1} - u_{t-k}$ cannot be determined so we cannot say anything definitive about the sign of $\widehat{\beta}_{Su}$. It is important to note however that, for any given realisation of the unemployment rate sequence, $p \lim \widehat{\beta}_{Su}$ is non zero¹⁶

¹⁵Of course employment has bounded support so technically speaking it can only be approximately normally distributed.

¹⁶It is possible that when we evaluate its unconditional mean, i.e. $\int_0^1 \dots \int_0^1 p \lim \widehat{\beta}_{Startu} f(u_1, \dots, u_t) du_1 du_2 \dots du_t$ that this quantity could be zero. But for any particular realisation of the unemployment sequence it will be non zero and of course it remains nonzero asymptotically as $n \rightarrow \infty$.

4.4 Unemployment rate sensitivity of new hires:- ∂u

Setting $c_{ijt} = \partial_{ijt}u_t$ in (6) above and using $c_{jt} = \sum_{i=1}^{L_{jt}} \partial_{ijt}u_t = u_t(L_{jt} - sL_{jt-1})$ in (6) gives the analogues to (13) and (14) as

$$p \lim \widehat{\beta}_{\delta u} \propto -(s\gamma_1 - \gamma_0)u_t \quad (17)$$

$$p \lim \widehat{\beta}_{\delta u} \propto -(s\gamma_1^* - \gamma_0^*)u_t \quad (18)$$

If γ_0 is negative and γ_1 is either relatively small in absolute value or is positive then $\widehat{\beta}_{\delta u}$ will have a negative probability limit. Once again these conditions apply to γ_0^* and γ_1^* . By contrast if γ_0 and γ_1 are both positive and $\gamma_1 > \gamma_0$ than $\widehat{\beta}_{\delta u}$ will have a negative probability limit.

We now examine the implications for the signs of the estimates if there is heterogeneity across firms in mean employment, wages and retention rates.

4.5 Heterogenous mean wages, mean firm employment and retention rates

Some of the theoretical equal treatment models considered in MSTa and here generate a negative cross firm covariance between wages and size and as noted above this is in conflict with the wage size premium. In addition and again as noted above, smaller, lower wage firms tend to have lower retention rates. Here we extend the formulae for $scov(w_{ijt}, c_{ijt})$ given in (6) above to allow for 2 sectors each containing n^i $i = h, l$ firms with, respectively, retention rates s_i , average firm wages \bar{w}_i^i and average firm employment \bar{L}_i^i .¹⁷We also assume that n^l and $n^h \rightarrow \infty$ with n . These two sectors are assumed to be separate subeconomies within a larger economy and under the assumption of no business cycle in wages and employment (no aggregate shocks) for large n , their mean firm wage and firm employment, \bar{w}^i and \bar{L}^i , will be fixed over time. To match the empirical regularities on s, w and L we assume that $s_l < s_h$. Again whilst we do not require $\bar{L}^l < \bar{L}^h$ we could envisage this to be the case in order to be consistent with a positive cross-sector size-wage premium.

In this scenario $scov(w_{ijt}, c_{ijt})$ becomes

¹⁷Note that the results we derive here readily extend to the $m > 2$ case - a proof of this is available on request.

$$scov(w_{ijt}, c_{ijt}) = p^l scov^l(w_{ijt}, c_{ijt}) + p^h scov^h(w_{ijt}, c_{ijt}) + p^l (\bar{w}^{li} - \bar{w}_t)(\bar{c}^l - \bar{c}_t) + p^h (\bar{w}^{hi} - \bar{w}_t)(\bar{c}^h - \bar{c}_t) \quad (19)$$

where $scov^i(\cdot)$ denotes a sample covariance measured over the subsample of workers in sector i and where p^i is the proportion of the labour force in sector r at time t .

As before we can talk unambiguously of small, low wage, low s firms, and large high wage high s firms and use annotations l and h respectively. Noting that $p \lim(L_{t-k}^i/L_t^i) = 1$ for $i = l, h$ when n is large and following the analysis in section 3 (and in particular (1)), the time t average UTI's \bar{c}_t^i for each of the three cases are

$$\begin{aligned} p \lim \bar{\delta} u_t^i &= (1 - s_i) u_t \quad i = h, l \\ p \lim \bar{m}_t^i &= (1 - s_i) \sum_{k=0}^{\infty} s_i^k u_{t-k}^m \quad i = h, l \\ p \lim \bar{S}_t^i &= (1 - s_i) \sum_{k=0}^{\infty} s_i^k u_{t-k} \quad i = h, l \end{aligned}$$

respectively. Again following section 3 above, we see that $p \lim \bar{\delta} u_t^l < p \lim \bar{\delta} u_t^h$ and that $p \lim \bar{m}_t^l < p \lim \bar{m}_t^h$ independently of \bar{L}^l, \bar{L}^h . Using this and the fact that $\bar{w}^l < \bar{w}^h$, the last two terms in (19) have a negative probability limit for the $\min u$ and δu cases. Hence a sufficient condition for $p \lim scov(w_{ijt}, c_{ijt})$ to be negative in the $\min u$ and δu cases is that $p \lim scov^i(w_{ijt}, c_{ijt})$ for $i = l, h$ also be negative. Therefore, if wages and employment in the high and low subeconomies are driven by the same economic model (albeit with different mean firm wages and employment) then we need only analyse the sign of $p \lim scov(w_{ijt}, c_{ijt})$ for that economic model to determine the sign of $\hat{\beta}_{\min u}$ and $\hat{\beta}_{\delta u}$. For Su however, these sufficient conditions do not apply:- even if we can determine the sign of $p \lim scov^i(w_{ijt}, c_{ijt})$ in (19) we cannot determine the sign of the second term.

In a companion paper, (Martins, Snell and Thomas, 2011 - henceforth MSTa) analyse a number of theoretical models of firm and sector wage/employment determination, including static and dynamic multisectoral competitive models and a model of firm (labour market) monopsony subject to dynamic labour adjustment costs. In each case, we show that when parameters lie in a region suggested by the relevant empirical studies, the conditions for negativity of $p \lim \hat{\beta}_{\delta u}$ and $p \lim \hat{\beta}_{\min u}$ given at the end of sections 4.2 and 4.4 are satisfied. Under some calibrations, some of the models exhibit a negative wage size premium - in contradiction to the evidence. However by bolting on the high and low subeconomy structure outlined in this subsection and by allowing different intercepts (the parameters that determine mean wages and employment) in the two subeconomies, the models can be engineered to exhibit any desired wage size premium. We apply the high/low subeconomy structure below to one of the two models we simulate in section 5.

4.6 A digression:- Hagedorn and Manovskii's q^{HM} and q^{EH} variates

Hagedorn and Manovskii(2010 - henceforth HM) describe a search environment where workers' wages in a job are equal to a common cyclical wage (as would be the case under a simple market clearing model without search) plus an idiosyncratic firm specific match component. The latter is the workers' unobserved firm specific human capital. HM argue that $\min u$ and Su are significant in Mincer regressions because they proxy for this unobserved match quality. They show that the expected number of job offers a worker receives during his working spell (a spell during which employment is continuous, in which the worker switches firms only in response to higher offers and which is terminated when he is laid off) helps explain the wage in that work spell. In the version of their model with exogenous separations they develop two variables that act as proxies for the human capital component of a worker's wage. Defining labour market tightness θ_t as the ratio of aggregate vacancies to the unemployment rate, these two variates are q^{HM} , the sum of θ_t 's during the current job spell and q^{EH} the sum of θ_t 's during the work spell up to the point the job started. HM argue that the significance of $\min u$ and Su does not necessarily support the respective rigid wage contracting models they were designed to test because these variates are also significant under HM's flex wage market clearing world. At first glance this appears to be similar to the point of our paper, namely to show that $\min u$ and Su may be significant and "correctly" signed under polar opposite conditions (i.e, equal treatment) to those which motivated the respective variate's construction. However this is misleading. Our effect arises not because the variates $\min u$ and Su proxy for the human capital elements of wages (as in HM) but because they are correlated with firm level wage and employment policies. Furthermore, because HM's variates are - like $\min u$ and Su - partly constructed from tenure dummies that are correlated with a firm's wage and hiring policy under an equal treatment model with homogenous workers, they too may be spuriously significant under such alternative models. If this was the case HM's variates would suffer the same fate as $\min u$,etc in that their significance does not necessarily support HM's flex wage spot market model.but instead could arise under one of the rigid wage equal treatment alternatives considered in this paper. To illustrate this idea, we take a closer look at q^{HM} and q^{EH} under an equal treatment model market clearing model with identical workers.

The equal treatment models we consider in this paper do not say anything about the vacancy rate θ - it does not feature in them at all. However, below we derive all of our analytical results for panels under the absence of a business cycle.in aggregate wages and employment. Following this line for θ_t we specify it as a constant θ (say). Again this will make our results stark by showing that apparently significant estimates of q^{HM} and q^{EH} obtain even when θ is constant. This will show clearly that it is endogenous tenure not unobserved human capital that is at work here.

To follow HM to the letter we should examine the joint behaviour of q^{HM} and q^{EH} in a wage regression. But to keep things simple and tractable we consider their probability limits as single regressors separately. In MSTa we show under our assumption of constant θ and using tedious derivation that

$$p \lim \widehat{\beta}_{q^{HM}} \propto \frac{\theta}{p \lim \bar{L}_t} \sum_{k=0}^{\infty} s^k \left(k + \frac{1}{1-s} \right) (\gamma_k - s\gamma_{k+1}) - \frac{\theta}{p \lim \bar{L}_t} \gamma_0 \frac{1+s}{1-s}$$

where we assume $p \lim \gamma_k^f = \gamma_k$ independent of t .

If the data are generated by an equal treatment model whereby a firm's (or a sector's) wage covaries with its labour force, then $\widehat{\beta}_{q^{HM}}$ will be nonzero even in the absence of a business cycle and without human capital. If we further suppose that the covariance between a firm's (sector's) wages and its current labour force is negative (i.e. $\gamma_0 < 0$) whilst covariances between its lagged labour force and its wages are zero (i.e. $\gamma_k = 0$ $k = 1, 2, 3..$). then $p \lim \widehat{\beta}_{q^{HM}}$ is positively proportional to $\frac{-s}{1-s} \gamma_0$ and is hence positive. In MSTa we present some models which have covariances with this property. By contrast if γ_0 and γ_1 are positive with $\gamma_k = 0$ for $k > 1$ and where $\gamma_1 > \frac{\gamma_0}{1-s}$ then $p \lim \widehat{\beta}_{q^{HM}}$ is positive. Again MSTa present dynamic models capable of generating covariances with this property. (although we would never argue the property was generic).

For q^{EH} things are more tricky. This requires data on the length of the current job spell when a worker joined the firm and this is not a variable that enters the models in our paper or MSTa. Furthermore and unlike labour market tightness, abstracting from the business cycle does not help much. Even without cyclical variation in wages and employment, there could be systematic time variation in the average length of measured job spells across time. Denoting $\overline{\tau_{tjk}}$ as the average job spell length on joining firm j of workers of tenure k , the formula for $\widehat{\beta}_{q^{EH}}$ is a simple adaptation of the formula for Su (15) and is given as

$$\widehat{\beta}_{q^{EH}} \propto - \sum_{k=1}^{\infty} s^k (\gamma_k - \gamma_0) (\overline{\tau_{tjk-1}} - \overline{\tau_{tjk}})$$

Without knowing the sign of $(\overline{\tau_{tjk-1}} - \overline{\tau_{tjk}})$ we cannot determine the sign of $\widehat{\beta}_{q^{EH}}$. but it will in general be non zero.

In sum it is possible that HM's q^{HM} and q^{EH} be significant and "correctly" (positively) signed even though the true world is radically different to the one they specify namely, a world without human capital or a business cycle in either wages, employment or unemployment.

4.7 Pooled estimation on a full panel dataset

We now show how the above results for $\widehat{\beta}$ carry over from a single cross section (single time period) to a full panel. We take the absence of a business cycle in firm employment and

wages to imply that for a worker i in firm j

$$w_{ijt} = f(\widetilde{\xi}_{jt}) \quad \text{and} \quad L_{jt} = f(\widetilde{\xi}_{jt}) \quad (20)$$

where $\widetilde{\xi}_{jt}$ is a vector of firm specific idiosyncratic shocks with time invariant pdf's.

Proposition 2:-

$$if \quad p \lim\{scov(w_{ijt}, c_{ijt})\} < 0 \quad t = 1, ..T \quad then \quad (21)$$

$$p \lim\{scov^p(w_{ijt}, c_{ijt})\} < 0 \quad (22)$$

where $scov^p(.)$ denotes a sample covariance derived from a panel and $scov(.)$ denotes one taken from a single cross section.

Proof:- See appendix

In the absence of aggregate shocks to firm wages and employment then, if $\widehat{\beta}$ has a negative probability limit in the cross section it also has a negative limit in the entire panel. Therefore if the sufficient conditions on $\gamma_k(\gamma_k^*)$ for (asymptotic) negativity of $\widehat{\beta}$ discussed in subsections 4.2,4.3 and 4.4 above hold in both the high and low sectors, this is all we need consider. We now turn to analyse the effects of random sampling on our results.

4.8 The effects of using a random sample

Until now we have assumed that we have access to a complete dataset of all the workers in an economy with a large number of firms. But investigators typically only have access to a random subsample of a particular population (a remarkable exception is the QPdataset in the case of Portugal). The effects of random sampling add technicalities but provided that the number of firms being sampled remains large the probability limits of the estimates are unchanged.

Suppose we have a random sample consisting of a proportion ρ_{jt} of firm j 's workforce at time t where ρ_{jt} equals a constant $\rho > 0$ plus an independently distributed finite variance shock ε_{jt} . so that

$$L_{jt-k}^\rho = \rho L_{jt-k} + \varepsilon_{jt-k} L_{jt-k} \quad \text{and} \quad (23)$$

$$L_t^\rho = \rho L_t + \sum_{j=1}^n \varepsilon_{jt} L_{jt} \quad (24)$$

where superscript ρ denotes a quantity from a random sample so that $L_{jt-k}^\rho \{L_t^\rho\}$ are the

number of workers sampled ex post from firm j at time $t - k\{t\}$.¹⁸.

Proposition 3:- The asymptotic quantities computed in this paper for the entire population of workers in n firms are unchanged if we have instead a random sample with properties given in(23) and (24)

Proof:- See appendix

5 Some numerical simulations

Here we analyse the values of $\hat{\beta}$ from calibrated versions of two models. The main purpose is to be indicative rather than exhaustive. We wish to show that under reasonable parameter values, equal treatment models are capable of generating UTI effects of a similar order of magnitude as those found in the empirical literature.

The first equal treatment model we use is one of labour contracting under exogenous real wage rigidity due to Martins, Snell and Thomas (2010) which we call MST. The second is a standard textbook model of dynamic labour demand in a multisectoral competitive framework which we call DCM. Although competitive, the DCM can be interpreted equivalently as a model of dynamic (labour) monopsony because as MSTa show, the stochastic structures for each sector's wages and employment in DCM is identical to that of each firm in the monopsony model.

The MST and DCM models do not have firms but sectors (although if we interpret the DCM model as one of monopsony, the sectors would be considered as being firms). Sectors are presumed to be segmented labour markets. Whilst it is not clear how many such labour markets exist in any economy, their number will be an order of magnitude lower than that of firms. In the face of this uncertainty we run simulations for numbers of sectors $ns = 9, 21$ and 51. The frequency is assumed to be annual with the number of years, T , in the panel set to 5, 10 and 20 - typical spans for many US panel data studies.

The MST Model.

Martins, Snell and Thomas (2010) give an equal treatment model where real wages are assumed to be downwardly rigid. In that, the maximum amount per period they can fall is exogenously bounded. If we specify that the maximum rate of real wage decline is the

¹⁸We note four things. First, we should more formally write $L_{jt-k}^r = int(\rho_{jt-k} L_{jt-k})$ where *int* denotes integer truncation but doing so changes nothing so we suppress this for brevity. Second, our assumptions on ε do not rule out $\rho_{jt} = 0$ for some firm j - the crucial assumption is that its mean ρ is strictly positive and constant and remains so as $n \rightarrow \infty$. Third and in the proof of the proposition, w_{jt} is written without a ρ superscript because it pertains to firm j and is not changed by random sampling. Finally allowing ρ_{jt} to be stochastic means that the sample is not stratified with respect to firms but obviously the stratified case - where the variance of errors goes to zero, is encompassed here.

inflation rate then this model becomes one of absolute nominal wage rigidity. The two equations describing wage-price dynamics are a wage equals MPL condition and a wage adjustment rule. We assume these are given by respectively

$$MPL_{jt} = w_{jt} = K + \xi_{jt} - \alpha l_{jt} \quad (25)$$

$$w_{jt} = \max\{\xi_{jt}, \mu w_{jt-1}\} \quad (26)$$

where $\mu \leq 1$. Following a large negative productivity shock, wages fall only slowly to the new market clearing level at a rate determined by μ which is assumed to be exogenous. We choose this model for our simulations because it has so few parameters and because as MST show, it fits the postwar US data on unemployment and wages quite well. We split the artificial economy into high and low sub-economies as described above. For the MST model we adopt a more general firm productivity process than before, one that includes both idiosyncratic (firm or sector specific) and aggregate shocks, namely

$$\xi_{jt} = \phi t + \varepsilon_{jt} + \tau_{et} \quad (27)$$

$$\tau_{et} = \tau_{et-1} + \epsilon_t \quad (28)$$

where ε_{jt} and ϵ_t are *iid* normally distributed firm specific and aggregate (log) productivity shocks with variances σ_ε^2 and σ_ϵ^2 respectively and where ξ_{jt} is the log of the total factor productivity (TFP) of sector j at time t . Given this productivity process, MST will generate genuine business cycles in wages and unemployment.

Unfortunately there is no data on sectoral TFP for the MST model to help us calibrate values for σ_ε^2 . However the Bureau for Labour Statistics does produce TFP estimates for 20 or so manufacturing sectors. The postwar standard deviations of TFP growth in these sectors lie between 2 and 5% - substantially higher than that for aggregate TFP as one might expect.given that the sectors will in part be driven by idiosyncratic elements. We therefore run two sets of simulations with $\sigma_\varepsilon = .02$ and $\sigma_\varepsilon = .05$.respectively. This should give us an idea of how $E(\hat{\beta})$ changes with idiosyncratic TFP volatility. When ns is large, idiosyncratic shocks will wash out and the standard deviation of aggregate productivity growth will be σ_ϵ In postwar annual US data, this quantity is roughly .017. By setting σ_ε .to .015 we get a standard deviation of aggregate productivity growth slightly below .017 for large ns and slightly above for small ns . The parameter α , the inverse of the sector wage elasticity of labour demand is set to 1.4, roughly in line with results from studies of labour demand using postwar US data. The extent to which real wages can fall within any year (μ) we set to .97. If inflation stands at 3% per annum - close to the postwar US average - than this setting implies downward nominal wage rigidity (for a recent model of nominal wage rigidity see Elsby, 2010). The trend term ϕ is set to .01 implying 1% per year growth in real wages. Finally we have two separate scenarios:- the first has a single economy with homogenous sectoral means and retention rates and the second allocates the sectors into high and low

sub-economies as per section 3. In the latter exercise, sectors in the high sub-economy have twice the mean employment and 5% higher wages than the low sub economy, and the retention rates for each sub-economy were .8 and .6 respectively. We kept the total size of each subeconomy equal by allowing the low sub-economy to have twice as many sectors as the high. In each simulation the wage-size premium is about 5% in keeping with elasticities estimated in the empirical literature for firms - the implicit assumption being therefore that firms in high sectors are twice the size of those in low sectors. By having results for the two subeconomies and for a single (single s) economy we are able to assess the effect heterogenous mean wages and retention rates have on the parameter estimates.

The DCM Model

The DCM model is intrinsically a "deviations from trend" model and wages here are in levels not logs. Firms in each sector j determine employment to maximise discounted profits subject to a quadratic cost of new hires, and to sectoral labour supply and to productivity shocks. The reduced form equations for sectoral employment and wages are

$$L_{jt} = \lambda L_{jt-1} + \frac{\lambda}{cs} \{\xi_{jt} + u_{jt}\} \quad (29)$$

$$W_{jt} = \lambda W_{jt-1} + \frac{\lambda - 2cs}{cs} u_{jt} + 2\lambda u_{jt-1} + \frac{\lambda}{cs} \xi_{jt} \quad (30)$$

where c is a parameter determining hiring costs, s is the (common) sectoral labour retention rate, $\lambda = \frac{cs}{1+c+\delta cs(s-\lambda)}$ with δ being the discount rate, ξ_t is an aggregate (common) productivity shock and where u_{jt} are idiosyncratic shocks to sector j 's labour supply (see MSTa).

The standard deviation of ξ_t is $\sigma_\xi = .025$ and the standard deviation of u_{jt} is set at a level that makes the standard deviation of aggregate (detrended) employment equal to 2% - roughly in line with postwar US data. The model is linear and aggregates so λ is the AR(1) coefficient in aggregate wages and employment. We set it to .6 again in line with postwar aggregate employment and wage data. The parameter c is set to 2 which implies equilibrium labour turnover costs are around 8% of the wage bill - roughly in line with estimates given by Mincer(1989). The retention rate s is set to .7, roughly in line with the average for US firms. Finally the discount rate δ is set to .98, a value typically used in macroeconomic analyses undertaken at the annual frequency. Under these parameter values the model naturally exhibits a positive wage-size premium. (see MSTa) and so we do not adopt the high-low subeconomy structure for this model.

The Simulation Results

We derive average values for $\hat{\beta}_{\min u}$, $\hat{\beta}_{Su}$, and $\hat{\beta}_{\partial u}$ using 1000 simulations for each model and parameter set. We add one further estimate $\hat{\beta}_{\delta \bar{u}}$ which uses the de-meaned aggregate unemployment rate $u_t - \frac{\sum_{t=1}^T u_t}{T}$ to construct δu rather than the unemployment level itself. We do this because we believe it is a more satisfactory way of modeling the impact of unemployment on wages in that allows for an arbitrary scale of u_t . In keeping with the

empirical literature we include a linear tenure term in all regressions and for the $\widehat{\beta}_{\delta u}$ case we add the aggregate unemployment rate as an extra regressor.

Results for the MST model under a single economy (Table 1) and two sub-economies (Table 2) and for the DCM model (Table 3) are given below

Table 1
Estimates of $E(\widehat{\beta})$ for the MST Model without high/low subeconomies

	$\sigma_\varepsilon = .02$				$\sigma_\varepsilon = .05$			
	$\widehat{\beta}_{\min u}$	$\widehat{\beta}_{\delta u}$	$\widehat{\beta}_{\delta \bar{u}}$	$\widehat{\beta}_{Su}$	$\widehat{\beta}_{\min u}$	$\widehat{\beta}_{\delta u}$	$\widehat{\beta}_{\delta \bar{u}}$	$\widehat{\beta}_{Su}$
$ns = 9, T = 5$	-2.05	-1.82	-2.50	-.43	-.57	-.29	-.71	-.13
$ns = 21, T = 5$	-2.42	-2.07	-2.95	-.51	-.64	-.23	-.61	-.10
$ns = 51, T = 5$	-3.44	-2.88	-4.46	-.66	-.65	-.22	-.70	-.13
$ns = 9, T = 9$	-2.12	-1.88	-2.36	-.49	-.63	-.29	-.63	-.17
$ns = 21, T = 9$	-3.06	-2.54	-3.48	-.70	-.63	-.22	-.60	-.14
$ns = 51, T = 9$	-3.28	-2.39	-3.39	-.85	-.86	-.22	-.77	-.21
$ns = 9, T = 20$	-1.98	-1.57	-1.95	-.77	-.50	-.20	-.37	-.15
$ns = 21, T = 20$	-2.77	-2.13	-2.82	-.90	-.75	-.20	-.50	-.22
$ns = 51, T = 20$	-3.30	-2.32	-3.15	-1.09				

Table 2
Estimates of $E(\widehat{\beta})$ for the MST Model with high/low sub-economies

	$\sigma_\varepsilon = .02$				$\sigma_\varepsilon = .05$			
	$\widehat{\beta}_{\min u}$	$\widehat{\beta}_{\delta u}$	$\widehat{\beta}_{\delta \bar{u}}$	$\widehat{\beta}_{Su}$	$\widehat{\beta}_{\min u}$	$\widehat{\beta}_{\delta u}$	$\widehat{\beta}_{\delta \bar{u}}$	$\widehat{\beta}_{Su}$
$ns = 10, T = 5$	-2.11	-2.16	-.83	-.45	-1.06	-.79	-.87	-.12
$ns = 20, T = 5$	-3.61	-3.46	-2.11	-.59	-.99	-.73	-.88	-.12
$ns = 50, T = 5$	-4.79	-4.69	-3.02	-.61	-.96	-.75	-.74	-.11
$ns = 10, T = 10$	-3.35	-3.10	-2.30	-.69	-.79	-.55	-.51	-.14
$ns = 20, T = 10$	-4.16	-3.80	-3.10	-.75	-1.03	-.65	-.70	-.19
$ns = 50, T = 10$	-5.12	-4.74	-3.62	-.79	-1.33	-.77	-.89	-.24
$ns = 10, T = 20$	-3.20	-2.92	-2.27	-.80	-.88	-.55	-.45	-.16
$ns = 20, T = 20$	-4.09	-3.52	-2.92	-.99	-1.18	-.62	-.58	-.22
$ns = 50, T = 20$	-5.44	-4.64	-3.95	-1.34	-1.41	-.77	-.72	-.26

Table 3
Estimates of $E(\hat{\beta})$ for the DMC Model

DMC	$\hat{\beta}_{\min u}$	$\hat{\beta}_{\delta u}$	$\hat{\beta}_{\delta \tilde{u}}$	$\hat{\beta}_{Su}$
$\sigma_u = .02$	$\sigma_u = .01$			
$ns = 9, T = 5$	-.997	-.008	-.128	-.302
$ns = 21, T = 5$	-.963	-.007	-.100	-.286
$ns = 51, T = 5$	-.978	-.007	-.112	-.296
$ns = 9, T = 10$	-.913	-.007	-.145	-.318
$ns = 21, T = 10$	-.947	-.009	-.151	-.330
$ns = 51, T = 10$	-.991	-.008	-.153	-.328
$ns = 9, T = 20$	-.923	-.012	-.184	-.387
$ns = 21, T = 20$	-.966	-.012	-.180	-.390
$ns = 51, T = 20$	-.988	-.012	-.187	-.412

We see that all estimates have a negative sign. In terms of magnitude the $\min u$ and Su estimates in MST for low idiosyncratic variance are similar to those obtained by BDN. The estimates of δu in this scenario are higher in absolute value than empirical counterparts. MST estimates from the high σ_ε case are in line with empirical counterparts for Su and δu but those for $\min u$ are a bit low compared with the values found in the empirical literature. DCM seems to produce estimates for Su that are broadly in line with empirical work but the results for $\min u$ and δu are lower than that typically found. By comparing the numbers in Tables 1 and 2 we can see that the high/low subeconomy structure appears to increase the estimates in absolute value and considerably so in many cases. Finally we report that although all regressions included a linear tenure term, the addition of this term had minimal impact on the estimates.

6 Extending SBP' method to handle unmatched datasets

The problem of bias we have identified in this paper has its root in firm (or sector) specific components of the wage that are related to firm (or sector) hiring levels. As noted above we could remove these by adding firm-year interaction dummies to the panel regression. Under an equal treatment model as laid out in (3) this would reduce the regressand w_{ijt} to idiosyncratic noise whereas under the hypothesis that wages are linear in c_{ijt} the addition of such dummy terms is innocuous. But what if the dataset does not match workers to firms? Large matched panel datasets abound in Europe but in the US they are virtually non-existent. In this section we tentatively offer a solution to the bias which is implementable in unmatched datasets. The solution is considerably inferior to the addition of firm-year interaction terms and will only work if cross firm wage-employment covariances are constant over time.

SBP point out that in a panel data set, macro variates like unemployment, have extremely limited variation. For example, adding the aggregate unemployment rate to a Mincer equa-

tion in the PSID involves dealing with a regressor that takes on <50 different values to explain wages which take on around a million different values. As Moulton(1990) shows, this is likely to impart huge bias to standard errors because of error clustering. SBP's solution was to use the panel dimension to control for worker characteristics and extract from the panel "composition bias free" estimates of mean wages at each time t via the addition to the Mincer equation of time dummies. Coefficients on these dummies - common time effects in wages - would then be regressed on unemployment and other macro series of interest in a time series regression. In an extension of this idea to $minu$ and Su , Devereux and Hart(2007) add tenure-year interaction dummies to extract composition bias free estimates of average wages within each tenure-year cell of the panel data. $Minu$ and Su only vary between tenure-year cells and are constant for workers within these cells. So again the idea is to condense the data to guarantee that the "x-variable" varies between each data point. Finally Shin and Shin(2003) extract time means of respectively, stayers' and movers' wages to estimate differential effects of unemployment on new hires and incumbent wages via separate time series regressions. We show below that these aggregation methods do not remove the bias we have identified in this paper but they do point to a possible way forward to remove it.

Again and without loss of generality, we abstract from worker characteristics so we can focus on raw mean wages. In this section we operate under the hypothesis of equal treatment as written in equations (3) and (5) above. We maintain the high/low firm structure above with high/low mean wages, high/low mean employment and high/low retention rates and where each firm is subject to aggregate and idiosyncratic shocks. Under this scenario the tenure k time t average (log) wage is

$$\bar{w}_t^k = \frac{\sum_{j=1}^{n_h} s_h^k (L_{jt-k}^h - s_h L_{jt-k-1}^h) w_{jt}^h + \sum_{j=1}^{n_l} s_l^k (L_{jt-k}^l - s_l L_{jt-k-1}^l) w_{jt}^l}{\sum_{j=1}^{n_h} s_h^k (L_{jt-k}^h - s_h L_{jt-k-1}^h) + \sum_{j=1}^{n_l} s_l^k (L_{jt-k}^l - s_l L_{jt-k-1}^l)} \quad (31)$$

where L_{jt-k}^i $i = h, l$ is employment in a firm j at time $t - k$ that is located in the i sector and where n^i $i = h, l$ is the number of firms in sector i ..assumed fixed over time. Defining the proportion of firms in the high/low sectors as a fixed constant $p^i = \frac{n^i}{n}$, dividing the top and bottom of (31) by n and taking probability limits as the n^i both go to infinity gives $p \lim \bar{w}_t^k (= \mu_{k,t})$ as

$$\mu_{k,t}(= p \lim \bar{w}_t^k) = \frac{A}{C} + \frac{B}{C} \quad (32)$$

$$\begin{aligned} \text{where } A &= p^h s_h^k (\gamma_{k,t}^h - s_h \gamma_{k+1,t}^h) + p^l s_l^k (\gamma_{k,t}^l - s_l \gamma_{k+1,t}^l) \\ B &= p^h \mu_t^h s_h^k (\bar{L}_{t-k}^h - s_h \bar{L}_{t-k-1}^h) + p^l \mu_t^l s_l^k (\bar{L}_{t-k}^l - s_l \bar{L}_{t-k-1}^l) \\ C &= p^h s_h^k (\bar{L}_{t-k}^h - s_h \bar{L}_{t-k-1}^h) + p^l s_l^k (\bar{L}_{t-k}^l - s_l \bar{L}_{t-k-1}^l) \end{aligned}$$

where $\mu_t^i = \frac{\sum_{j=1}^n w_{jt}^i}{n^i}$ and \bar{L}_{t-k}^h are the (*unweighted*) average firm wage at t and firm employment(size) at $t - k$ in the i sector.respectively and where $\gamma_{k,t}^i$ is the probability limit of the sample covariance of L_{jt-k} and w_{jt} . We make two further simplifying assumptions. First we assume that $\bar{L}_{t-k}^l = \rho \bar{L}_{t-k}^h$ ($\rho < 1$) i.e. that employment in the high and low sectors has common cyclicalty.¹⁹ Second we assume that the normalised covariances $\gamma_{k,t}^{i*} = \frac{\gamma_{k,t}^i}{\bar{L}_{t-k}^i \mu_t^i}$ are constant over time and henceforth drop the t subscript. Under these assumptions (32) takes the form

$$\mu_{k,t} = \mu_t^h + \frac{a_k - a_{k+1} \Delta_{kt} - (b_k - b_{k+1} \Delta_{kt}) w p_t}{c_k - c_{k+1} \Delta_{kt}} \quad (33)$$

$$\text{where } a_k = p^h s_h^k \gamma_k^{h*} + \rho p^l s_l^k \gamma_k^{l*} \quad c_k = p^h s_h^k + \rho p^l s_l^k \quad \text{and } b_k = \rho p^l s_l^k$$

where $\Delta_{kt} = \bar{L}_{t-k-1}^h / \bar{L}_{t-k}^h$. Under a constant or slow moving labour supply, Δ_{kt} is approximately one plus the change in the aggregate unemployment rate at time $t - k$.and as it only enters tenure k 's cell mean it is the same as the change in the "start unemployment rate". Equation (33) shows that under equal treatment, average wages in the tenure-year cells $\mu_{k,t}$ will, in general, vary with tenure and time. In fact even if wages were equal across firms. ($\gamma_k^{i*} = 0$ and $\mu_t^i = \mu$).as long as employment was cyclical, cell mean wages would still display cyclical variation over time and tenure. The SBP method²⁰ uses \bar{w}_t^k to estimate $\mu_{k,t}^h$. We then regress \bar{w}_t^k (which form a balanced panel dataset) on the relevant cell value of the UTI, c_{kt} say. In the case of δu , k takes the value 0 for new hires and 1 for all other tenures (incumbents) and there are two regressors; δu_{kt} and u_t . We consider the consequences of using the SBP method for each of our three UTI's in turn.

a) *Su* : – Equation (32) shows that under equal treatment $\mu_{k,t}^h$ will be related to the

¹⁹This assumption would hold true if each firm's employment was linear in idiosyncratic shocks and in aggregate shocks.with the latter entering with coefficient $\alpha(\rho\alpha)$ in high(low) firms. We should note that Moscarini and Postel-Vinay, (2008) find that "high" firms (large firms with high average wages) have more cyclical employment than do "low" firms. We make the assumption of common cyclicalty to simplify matters but it should be clear from the discussion that greater cyclicalty of high firms would make our results even more pronounced.

²⁰More properly its extension in Devereux and Hart (2007).

change in start unemployment. Hence Su will be significant both under bilateral contracting of the Su variety *and* when there is equal treatment.

b) $\min u$: – Again problems arise here because of potential comovement of $\min u$ with \bar{w}_t^k over t and k . under equal treatment. To give a specific and simple example we could return temporarily to the base scenario of this paper and suppose that aggregate shocks are absent so that mean wages *and* mean firm employment at time t are constant²¹. Economic models of wage determination lead us to expect that the $|\gamma_{k,t}^h|$ will decline with k albeit not necessarily monotonically²². Equally we know that $\min u$ will decline with k although again not necessarily monotonically. If the $\gamma_{k,t}^h$ are predominantly negative we would expect a spuriously negative coefficient in the regression of \bar{w}_t^k on the t, k . cell $\min u$.

c) δu : – The SBP method has been used several times in the empirical literature to estimate the differential response of new hire wages to unemployment so we now flesh out more explicit results for this case. The mean wage of incumbents (μ_{It}) is

$$\begin{aligned} \mu_{It} &= p \lim \frac{\sum_{j=1}^{n_h} s_h L_{jt-1}^h w_{jt}^h + \sum_{j=1}^{n_l} s_l L_{jt-1}^l w_{jt}^l}{\sum_{j=1}^{n_h} s_h L_{jt-1}^h + \sum_{j=1}^{n_l} s_l L_{jt-1}^l} = \frac{s_h \gamma_{1,t}^h + s_l \gamma_{1,t}^l}{(s_h + \rho^* s_l) \bar{L}_{t-1}^h} + \frac{s_h \mu_t^h + s_l \rho^* \mu_t^l}{s_h + \rho^* s_l} \\ &= \mu_t^h + \frac{s_h \gamma_1^{h*} \mu_t^h + s_l \gamma_1^{l*} \mu_t^l}{s_h + \rho^* s_l} + \frac{s_l \rho^*}{s_h + \rho^* s_l} w p_t \end{aligned} \quad (34)$$

where $\rho^* = \rho p^l / p^h$ is the ratio of the number of workers in low firms to those in high firms in the economy as a whole. Adapting (32) with $k = 0$ to get the corresponding case for new hires gives

$$\mu_{0t} = \mu_t^h + \frac{a_0 - a_1 \Delta_{kt} - (b_0 - b_1 \Delta_{0t}) w p_t}{c_0 - c_1 \Delta_{0t}} \quad (35)$$

Equations (34) and (35) show that the mean incumbent and new hire wages are both weighted averages of the μ_t^l 's but the former has fixed weights whereas the latter has weights that vary with Δ_{0t} . (approximately one plus the change in the current unemployment rate). An interesting special case is where wages in firms are acyclical - constant to make this an extreme case - but where aggregate employment is cyclical. Linearising the second term in (35) around $\Delta_{0t} = 1$ we can rewrite (35) as

²¹As before we would require aggregate labour supply to vary over time in order to avoid a constant aggregate unemployment rate).

²²In dynamic models, the γ_k^l 's will be non zero because current and lagged idiosyncratic shocks affect firm wages and firm employment. These models usually embed stationarity guaranteeing that the γ^l 's > 0 with k .

$$\mu_{0t} \approx \text{const} \tan t - \frac{\rho^*(s_h - s_l)wp}{(1 - s_h + \rho^*(1 - s_l))^2}(\Delta_{0t} - 1) = \alpha + \beta\Delta_{t-1} \quad (36)$$

where $\beta < 0$. As noted above $\Delta_{0t} - 1$ is approximately the change in the aggregate unemployment rate. Unlike μ_{It} therefore, μ_{0t} would appear to be procyclical and regressing $\{\mu_{It}, \mu_{0t}\}$ on the aggregate unemployment rate and a new hire dummy times the unemployment rate (δU) would yield a zero coefficient on the former but a spuriously negative coefficient on the latter.²³

As a final note and in contrast to the above, if we again assume common cyclicity of employment in the high and low sectors, we can show that wages averaged over *all* workers at time t (μ_t) do not display spurious cyclicity under a null of equal treatment. Using $\bar{L}_{t-k}^l = \rho \bar{L}_{t-k}^h$, $k = 0, 1, 2..$ and following familiar arithmetic manipulations it is easy to show that μ_t is given by

$$\mu_t = \frac{1}{1 + \rho^*} (1 + \gamma_0^{*h}) \mu_t^h + \frac{\rho^*}{1 + \rho^*} (1 + \gamma_0^{*l}) \mu_t^l \quad (37)$$

where $\frac{1}{1 + \rho^*}$ and where $\frac{\rho^*}{1 + \rho^*}$ are proportions of the workforce in low and high firms respectively. Hence, whilst (32) to (36) show how mean wages at time t for tenure k will in general display spurious cyclicity and spurious tenure effects (37) shows that - under our simplifying assumptions,- average wages across all workers (tenures) at time t will not. If T was large the investigator could regress composition bias free estimates of μ_t on \bar{c}_{ijt} . (to capture the alternative hypothesis) and on presumed determinants of μ_t^i such as u and trend (to capture the null). The significance (and "correct" sign) of \bar{c}_{ijt} would favour the alternative hypothesis of the UTI in question. However very often T is too small to get reliable estimates this way and in any event, ignoring cross tenure variation in wages will severely reduce power under the alternative.

6.1 Adapting the SBP method:- An empirical illustration

SBP obtain composition bias free estimates of mean wages for each relevant tenure category. For $\min u$ and Su this means using the panel dimension to control for worker characteristics and averaging the residual wages in each tenure-time cell to obtain estimates of the μ_t^k . These are then regressed on the relevant cell value of $c_{ijt} \cdot (c_{kt})$ (see for example Devereux and Hart, 2007). However as argued above, equation (32) shows how this may lead to spurious results. To eliminate this possibility, we suggest adding extra regressors to absorb the terms in (32) i.e. those terms that would appear if our equal treatment model held true. Taking the simplifying assumptions of the previous section on board again here (constant normalised

²³This is a relatively simple result to derive and proof is available on request.

covariances and equal cyclical of employment in high and low firms) it is easy to show that we can linearise (32) to get

$$\mu_{k,t} \simeq a_k + b_k \mu_t + c_k \Delta u_{t-k} \quad (38)$$

where μ_t is a weighted average of μ_t^h and μ_t^l . If we further assume that μ_t is driven by a deterministic trend (t) and by aggregate unemployment (u_t) we could regress (estimates of) $\mu_{k,t}$ on $t, u_t, \Delta u_{t-k}, c_{kt}$ and on tenure dummies allowing coefficients on all but the last two to differ across tenures. Another way of viewing this procedure is to see it as a set of k regression equations, one for each tenure subject to the cross equation restriction of a single common coefficient on c_{kt} . We call this the modified SBP method (MSBP). To apply it we first of all need to use the panel dimension to factor out worker composition effects from $\mu_{k,t}$.

Our empirical model may be summarised as

$$w_{ijt} = \nabla' \mathbf{x}_{ijt} + \alpha \tau_{ijt} + \beta c_{ijt} + w_{jt} + v_{ijt} \quad (39)$$

where $cov(w_{jt}, x_{ijt}) = 0$

$$H_0 : \beta = 0 \quad H_1 \beta < 0$$

where \mathbf{x}_{ijt} is a $ax1$ vector of worker characteristics such as educational attainment (it may also include worker fixed effect dummies), τ_{ijt} is worker tenure and v_{ijt} is an idiosyncratic error term independent of all the RHS variables. As before w_{jt} is an unobserved firm j specific component of wages that will in general contain aggregate variates such as unemployment and a time trend as well as idiosyncratic components such as firm specific productivity shocks. The way the hypotheses are set up allows firm specific wage components w_{jt} to exist under H_1 . As noted in the introduction to this paper it is quite likely that several contracting mechanisms simultaneously co-exist in a large economy in different sectors. Alternatively wages within a sector or firm may have a firm specific component and a differential tenure related business cycle component. Equations (34), 35 and (36) above suggest that it may be impossible to reject the existence of firm specific wage components in unmatched datasets.

Under the assumptions in (39) we can obtain consistent estimates of ∇ under both null and alternative by executing the OLS regression

$$w_{ijt} = \nabla' \mathbf{x}_{ijt} + \sum_k \sum_t \mu_{k,t} \partial_{ijt}^k + e_{ijt} \quad (40)$$

where ∂_{ijt}^k is unity if the worker is of tenure k at time t . and zero otherwise. The tk estimates of $\mu_{k,t}$ ($\hat{\mu}_{k,t}$) provide us with composition-free wage means for each k, t cell to be used in the second stage regression.

To illustrate this procedure and to get a handle on what difference it may make in US panel data we collected an unbalanced panel dataset from the PSID. for the years 1976 to 1993. This is a period which nests the years selected by BDN (1976-84) and which displays much time series volatility (an oil price shock and two major recessions) We collected information on workers' real log wages (real 1983 \$ using the CPI deflator), occupation (7 categories), education (7 categories), State of residence, age, tenure (in years) and race (white, Hispanic, and other). Despite differences between our data collection and that of BDN ²⁴ our panel estimates for $\min u$, Su and u for the subsample in 1976-84 (the BDN years) are close to that obtained by BDN as rows 1. to 6. in Table 4 show. Extending the data to 1993 and more than doubling the number of observations makes little qualitative difference as lines 7. to 9. show although the estimates are somewhat smaller in absolute value here. Adding year effects - there is a negative trend in aggregate wages during this period - does not change the sign or nominal ²⁵ significance of the estimates. Finally all coefficients on characteristics were correctly signed and had reasonable orders of magnitude.

When it came to implementing the MSBP method we encountered some problems At large tenures, some tenure/year (k, t) cells were empty and some others contained too few observations to give reliable estimates of wage means. To avoid null or sparsely populated cells we computed cell means for 9 tenure categories - tenures 0 to 8 and a final category consisting of all tenures in excess of 8 years. Table 5 gives the SBP and MSBP estimates for the 1976-93 sample. Lines 1 and 2 show the results for the regression of $\hat{\mu}_{k,t}$ on trend, tenure and $\min u$ and on trend, tenure and Su respectively. $\min u$ and Su have "correct" sign but only the former is significant. This is in keeping with results in the literature where $\min u$ has been consistently found to be negative and significant in a variety of datasets and specifications whereas success with Su has been mixed.

Using (38) as a guide we add extra regressors to purge the regression error of terms whose presence is induced by the existence of equal treatment wage components. Explicitly we add u_t and $\delta_{kt}u_t$ ($k = 0, 1, \dots, 8$), t and $\delta_{kt}t$ ($k = 0, 1, \dots, 8$), $\delta_{kt}\Delta u_{t-k}$ ($k = 0, 1, \dots, 9$) and δ_{kt} ($k = 0, 1, \dots, 9$). where δ_{kt} is a dummy variable indicating tenure k . Lines 3 and 4 show that adding these terms reduces $\min u$ to being wholly insignificant and both Su and $\min u$ now take the wrong sign. The two Wald tests (available on request) on the 8 $\delta_{kt}u_t$ terms and on the 8 δ_{kt} terms respectively were wholly insignificant. However these terms turned out to be highly collinear and a test for joint significance of all 16 of them had a p-value below 1%.

The shortage of degrees of freedom inhibit applying this method rigorously to the BDN years (37 regressors but only 81 observations) but for completeness' sake we report the results for this subsample anyway in lines 5 to 8. Asymptotic inference is unreliable here but the

²⁴Tenure was taken directly from answers to the question relating to "present employer". By contrast BDN employ the algorithm of Altonji and Shakotko(1987) to modify the raw tenure data. However they argue it made little difference to their results. We also note that average tenure from our data for the relevant subsample is within 5% of BDN's. BDN also have 13 industrial sectors, marital status and union membership. They also add worker fixed effects but their results show that these have little qualitative impact on their results.

²⁵The word nominal is used because of the Moulton problem.

results do seem to be qualitatively similar to those in the larger sample.

Before closing we note two more things. First, the $\delta_{kt}\Delta u_{t-k}$ terms were significant (χ^2_9 values of 22.4 and 21.9 in the $\min u$ and Su regressions respectively) and this is quite interesting. Whilst there may be stories to support existence for tenure varying trends and intercepts (for example a complex rewards to tenure scheme), the existence of tenure varying responses to the change in initial unemployment (Δu_{t-k}) is hard to rationalise using economic arguments. Second, and by contrast, there is an obvious caveat to this procedure. Tenure related terms added to the regression will soak up a lot of the cross tenure variation in $\min u$ and Su . In short the method undermines the power of tests under the alternative that $\min u$ and Su do actually determine wages. This brings us back to the point made earlier in the paper that the first best solution to the problem is to purge w_{ijt} of any firm specific wage components via the addition of firm-year interaction dummies to the original panel.

Table 4
Panel estimates of $\min u, Su$ and δu from the PSID

	<i>u</i>	<i>Su</i>	<i>Minu</i>
BDN 1976-84 (N=19958)			
1.	-.020 (.002)		
2.		-.030 (.002)	
3.			-.045 (.003)
MST 1976-84 (N=19749)			
4.	-.023 (.002)		
5.		-.025 (.002)	
6.			-.054 (.003)
1976-93 Panel (N=46057)			
7.	-.010 (.011)		
8.		-.017 (.001)	
9.			-.033 (.016)

Table 5Estimates²⁶ from the PSID using BSP and MBSP.

	<i>Su</i>	<i>Minu</i>	τ	<i>t</i>
SBP 1976-1993 (N=162)				
1.		-.015 (.006)	.015 (.002)	-.015 (.001)
2.	-.003 (.059)		.017 (.001)	-.015 (.001)
MSBP 1976-1993 (N=162)				
3.		.001 (.013)		
4.	-.005 (.007)			
N=162				
SBP 1976-1984 (N=81)				
5.		-.046 (.008)	.005 (.002)	-.004 (.003)
6.	-.025 (.008)		.012 (.002)	-.008 (.004)
MSBP 1976-1984 (N=81)				
7.		.007 (.027)		
8.	.002 (.012)			

References

- [1] Altonji, J. (1986) "Intertemporal Substitution in Labor Supply: Evidence from Micro Data". *Journal of Political Economy*, Vol. 94 pp176-215.
- [2] Altonji, J and R.Shakotko (1987). "Do wages rise with job seniority?". *Review of Economic Studies*, Vol 54. pp437-59.
- [3] Baker G, Gibbs, M. and B. Holmstrom (1994). "The Wage Policy of a Firm," *The Quarterly Journal of Economics*, MIT Press, vol. 109(4), pp. 921-55.
- [4] Solon, G, Barsky, R and J Parker(1994). "Measuring the Cyclicity of Real Wages: How Important Is Composition Bias?," *The Quarterly Journal of Economics*, MIT Press, vol. 109(1), pages 1-25, February.
- [5] Beaudry, P. and J. Dinardo (1991). "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data," *Journal of Political Economy*, vol. 99 (4), pp. 665-88, August.
- [6] Blanchard, O. and L. Summers (1986). "Hysteresis and the European Unemployment Problem", *Macroeconomics Annual*, 1.
- [7] Cahuc, P, Postel-Vinay, F. and J-M. Robin (2006). "Wage Bargaining with On-the-job Search: Theory and Evidence", *Econometrica*, Vol 74 No.2, pp323-64.
- [8] Carneiro, A. Guimares, P. and P. Portugal (2009) "Real Wages and the Business Cycle: Accounting for Worker and Firm Heterogeneity," *IZA Discussion Papers* 4174.
- [9] Carruth, A. and A. Oswald (1987). "On Union Preferences and Labour Market Models: Insiders and Outsiders", *Economic Journal*, Vol. 97 pp 431-45, June
- [10] Devereux, P. and R.A. Hart (2006). "Real Wage Cyclicity of Job Stayers, Within-Company Job Movers, and Between-Company Job Movers," *Industrial and Labor Relations Review*, Vol 60 No. 1 pp105-119
- [11] Elsby, M. (2009). "Evaluating the Economic Significance of Downward Nominal Wage Rigidity", *Journal of Monetary Economics*, vol. 56, no. 2
- [12] Estevao, M. and B. Wilson (1998). "Nominal Wage Rigidity and Real Wage Cyclicity", *Finance and Economics Discussion Series* 1998-21, Board of Governors of the Federal Reserve System (U.S.). <http://www.federalreserve.gov/pubs/feds/1998/199821/199821pap.pdf>
- [13] Even, W.E. and David A. Macpherson (1996). "Employer Size and Labor Turnover: The Role of Pensions", *Industrial and Labour Relations Review*, vol. 49, No. 4, pp. 707-728

- [14] Montuenga, V. García, I and M. Fernández(2006). "Regional Wage Flexibility: the Wage Curve in Five EU Countries", in *The European Labour Market*, AIEL Series in Labour Economics, 2006, 245-265, DOI: 10.1007/3-7908-1680-9_12
- [15] Grant, Darren (2003). "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from National Longitudinal Surveys", *Industrial and Labor Relations Review*, 56, 393–408.
- [16] Gertler and Tregari (2009). "Unemployment Fluctuations with Staggered Nash Wage Bargaining", *Journal of Political Economy*, vol 117 no. 1.
- [17] Hagedorn, M. and I.Manovskii (2010), "Spot Wages over the Business Cycle?", Working Paper, October.http://www.econ.upenn.edu/~manovskii/papers/Spot_Wages.pdf
- [18] Hall, R.E. (2005). "Employment Fluctuations with Equilibrium Wage Stickiness", *American Economic Review*, vol. 95 no.1
- [19] Kilponen, J and Santavirta(2010). "New Evidence on Implicit Contracts from Linked Employer-Employee Data", *Bank of Finland Research Discussion Paper No. 12/2010*.
- [20] Kudlyak, M (2009). "The Cyclicity of the User Cost of Labor with Search and Matching." *Federal Reserve Bank of Richmond Working Paper 09-12 (September)*.
- [21] Lallemand, T., Plasman, R. and F.Rycx (2003) "The establishment-size wage premium: evidence from European countries". *Empirica*, Vol 35 no.4
- [22] Lindbeck, A. and D.Snowder (2001). "Insiders Versus Outsiders" *Journal of Economic Perspectives*, American Economic Association, vol. 15(1), pp 165-188, Winter.
- [23] Lockwood, B and A.Manning (1989). "Dynamic wage-employment bargaining with employment adjustment costs", *Economic Journal*, Vol 99 pp. 1143-1158.
- [24] McDonald, J and C.Worswick(1999). "Wages, Implicit Contracts, and the Business Cycle: Evidence from Canadian Micro Data", *Journal of Political Economy*, Vol. 107 No.4, pp 884-892.
- [25] Martins, Snell and Thomas (2005). "Wage Dynamics, Cohort Effects, and Limited Commitment Models", *Journal of the European Economic Association*, Vol 3 pp353-359
- [26] Martins, Snell and Thomas (2010). "Real and Nominal Wage Rigidity in a Model of Equal-Treatment Contracting", forthcoming in the *Scandinavian Journal of Economics* (Earlier version: IZA DP 4346.)
- [27] Martins, Snell and Thomas (2011). "Minus, Startu and all that: Pitfalls in Estimating the Sensitivity of a Worker's Wage to Aggregate Unemployment". *University of Edinburgh School of Economics WP199*. (http://www.econ.ed.ac.uk/papers//Martins_Snell_Thomas_sign%20of%20minus%20january.pdf)

- [28] Mincer, J. (1989). "Job Training: Costs, Returns and Profiles". NBER WP No. W3208
- [29] Montuenga, Garcia and Fernandez(2006), "Regional Wage Flexibility: the Wage Curve in Five EU Countries" in "The European Labour Market" AIEL Series in Labour Economics, 2006, 245-265,
- [30] Mortensen and Pissarides (1994)."Job Creation and Job Destruction in the Theory of Unemployment", Review of Economic Studies,vol 61, pp. 397-415.
- [31] Moscarini, G. and F. Postel-Vinay, 2008, "The Timing of Labor Market Expansions: New Facts and a New Hypothesis", NBER Macroeconomics Annual, 23, 1-51.
- [32] Moulton, Brent (1990). "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", The Review of Economics and Statistics, Vol. 72(2), pp. 334-338.
- [33] Oi, W.Y. and L.T. Idson (1999), "Firm Size and Wages", "The Handbook of Labour Economics", Vol 3, Part 2 pp2165-2214 (Elsevier)
- [34] Rossi-Hansberg, E. and M.Wright, (2007) "Establishment Size Dynamics in the Aggregate Economy", American Economic Review, vol 97, No.5
- [35] Schmieder, J and T von Wachter (2010) "Does Wage Persistence Matter for Employment Fluctuations? Evidence from Displaced Workers" American Economic Journal: Applied Economics, Volume 2, Number 3, July 2010 , pp. 1-21(21)
- [36] Shapiro and J. Stiglitz(1984)."Equilibrium Unemployment as a Worker Discipline Device", American Economic Review, vol.74 no. 3
- [37] Shin, D. and K.Shin (2003). "Why Are The Wages of Job Stayers Procyclical?," ISER Discussion Paper 0573, Institute of Social and Economic Research, Osaka University.(<http://www.iser.osaka-u.ac.jp/library/dp/2003/DP0573.pdf>)
- [38] Snell, A. and J.Thomas (2010) "Labor Contracts, Equal Treatment, and Wage-Unemployment Dynamics", forthcoming in the American Economics Journal: Macroeconomics.

Proof of Proposition 1:-

Note that in a single complete cross section the number of observations is the labour force at time t , L_t . Using this, the numerator in (5) is

$$scov(w_{ijt}, c_{ijt}) = \frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} w_{ijt} c_{ijt} - \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} w_{ijt} \right) \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} c_{ijt} \right) \quad (\text{A1})$$

We can substitute (3) into the RHS of (A1) to get

$$\begin{aligned} scov(w_{ijt}, c_{ijt}) &= \frac{1}{L_t} \sum_{j=1}^n w_{jt} \sum_{i=1}^{L_{jt}} c_{ijt} - \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} w_{jt} \right) \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} c_{ijt} \right) \\ &+ \frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} v_{ijt} c_{ijt} - \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} v_{ijt} \right) \left(\frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} c_{ijt} \right) \\ &= \frac{1}{L_t} \sum_{j=1}^n w_{jt} c_{jt} - \frac{1}{L_t} \sum_{j=1}^n L_{jt} w_{jt} \frac{1}{L_t} \sum_{j=1}^n c_{jt} + o_p(1) \end{aligned} \quad (\text{A2})$$

$$(\text{A3})$$

where $c_{jt} = \sum_{i=1}^{L_{jt}} c_{ijt}$ and where the $o_p(1)$ terms derive from the fact that the v 's are idiosyncratic and that L_t goes to ∞ with n .

$$scov(w_{ijt}, c_{ijt}) = \frac{1}{L_t} \left(\frac{1}{n} \sum_{j=1}^n w_{jt} c_{jt} - \frac{1}{n} \sum_{j=1}^n L_{jt} w_{jt} \frac{1}{L_t} \sum_{j=1}^n c_{jt} \right) + o_p(1) \quad (\text{A4})$$

$$= \frac{1}{L_t} \left\{ \frac{1}{n} \sum_{j=1}^n w_{jt} c_{jt} - \frac{1}{n} \sum_{j=1}^n w_{jt} \frac{1}{n} \sum_{j=1}^n c_{jt} - \frac{1}{n} \sum_{j=1}^n (L_{jt} - \bar{L}_t) w_{jt} \frac{1}{L_t} \sum_{j=1}^n c_{jt} \right\} + o_p(1) \quad (\text{A5})$$

$$= \frac{1}{L_t} \left(scov^f(w_{jt}, c_{jt}) - scov^f(L_{jt}, w_{jt}) \cdot \frac{1}{L_t} \sum_{j=1}^n c_{jt} \right) + o_p(1)$$

which establishes (6) in the text.

Proof of Proposition 2

It follows from this that the time t averages of wages (\bar{w}_t) are

$$\bar{w}_t = \frac{1}{L_t} \sum_{j=1}^n \sum_{i=1}^{L_{jt}} w_{ijt} = \frac{1}{L_t} \left(\frac{1}{n} \sum_{j=1}^n L_{jt} w_{jt} \right)$$

Again allowing n the number of firms to go to infinity gives us the probability limit.²⁷

$$p \lim \bar{w}_t = \frac{\gamma_0}{L} + \mu_w$$

where $\mu_w = p \lim \left(\frac{1}{n} \sum_{j=1}^n w_{jt} \right)$. and as before $\gamma_0 = p \lim \text{scov}^f(w_{jt}, L_{jt})$.

In general then wages would vary from firm to firm as would employment. In aggregate however and in a large economy, employment and average wages at time t - whether measured across firms or across a sample of individuals working at those firms - are constant over time.
28

We focus on $\hat{\beta} = \frac{1}{\text{svar}^p(c_{ijt})} (\text{scov}^p(w_{ijt}, c_{ijt}))$ with $i = 1 \dots L_{jt}, j = 1 \dots n$ and $t = 1, \dots, T$. The superscript p denotes a sample covariance from full panel. As before we are only interested in the sign of $\hat{\beta}$ so we can focus on the $p \lim$ of the numerator alone.

We can always write a sample covariance over T time periods as a weighted average of the within time covariances plus "across time" covariances i.e.

$$\text{scov}^p(w_{ijt}, c_{ijt}) \equiv \sum_{t=1}^T p_t \text{scov}(w_{ijt}, c_{ijt}) + \sum_{t=1}^T p_t (\bar{w}_t - \bar{w})(\bar{c}_t - \bar{c}) \quad (\text{A7})$$

where $\bar{w} = \frac{1}{N} \sum_{t=1}^T \sum_{j=1}^n \sum_{i=1}^{L_{jt}} w_{ijt}$. is the average wage in the entire panel, $p_t = \frac{L_t}{N}$ ²⁹ is the proportion of panel observations (N) occurring at time t . Our assumption for wages implies that

²⁷Note that we assume the number of firms is constant across time. This is purely to save notation. It would not change anything if we allowed the number of firms to vary over time and instead based a probability limit on $n_{\min} = \min(n_1, n_2, \dots, n_T) \rightarrow \infty$. Similarly allowing firm *composition* to change across time would merely increase notation: - All workers are identical and firms only differ in that each has its own wage driven by an idiosyncratic shock(s).

²⁸As we have already noted we require some movement in aggregate labour supply over time as a device to generate some variation in $\min u$ and Su over individual workers in the panel.

²⁹Note that $p \lim p_t = \frac{1}{T}$ so panel sample covariances are the simple unweighted average of their cross sectional counterparts.

$$p \lim \bar{w}_t = p \lim \bar{w} \quad t = 1, \dots, T$$

Under these assumptions the second term in (A7) vanishes asymptotically and (21) and (22) in the text directly follow.

Proof of Proposition 3

We show the result for $\hat{\beta}_{\min u}^\rho$. Adaptation of the analysis below to δu , Su and to sample means computed in Section 7 is obvious and straightforward and is available on request.

The numerator of $\hat{\beta}_{\min u}^\rho$ for a random sample from a single cross section at time t can be found via a simple adaptation of (13) namely

$$\text{Numerator}(\hat{\beta}_{\min u}^\rho) = -\frac{1}{L_t^\rho} \sum_{k=1}^{\infty} s^k (\gamma_k^{f\rho} - \frac{L_{t-k}^\rho}{L_t^\rho} \gamma_0^{f\rho}) (u_{t-k+1}^m - u_{t-k}^m) \quad (\text{A8})$$

The assumptions in (23) and (24) imply that

$$\begin{aligned} p \lim(\bar{L}_t^\rho) &= \rho \bar{L}_t \\ p \lim \gamma_k^{f\rho} &= p \lim \{ \text{scov}^f(\rho L_{jt-k}, w_{jt}) \} + p \lim \{ \text{scov}^f(\varepsilon_{jt-k} L_{jt-k}, w_{jt}) \} = \rho \gamma_k \\ p \lim \frac{L_{t-k}^\rho}{L_t^\rho} &= p \lim \frac{L_{t-k}^\rho/n}{L_t^\rho/n} = \frac{\bar{L}_{t-k}}{\bar{L}_t} \end{aligned}$$

Using these three probability limits in (A8) we see that the numerator in $\hat{\beta}_{\min u}^\rho$ is asymptotically unchanged by random sampling.

For the denominator we have

$$\begin{aligned} \text{Denominator}(\hat{\beta}_{\min u}^\rho) &= \frac{1}{L_t^\rho/n} \left\{ \sum_{j=1}^n \sum_{i=1}^{L_{jt}^\rho} \min u_{ijt}^2/n \right\} - \left\{ \frac{1}{L_t^\rho/n} \right\}^2 \left\{ \sum_{j=1}^n \sum_{i=1}^{L_{jt}^\rho} \min u_{ijt}/n \right\}^2 \\ &= \frac{1}{L_t^\rho/n} \{A^\rho/n\} - \left\{ \frac{1}{L_t^\rho/n} \right\}^2 \{B^\rho/n\}^2 \end{aligned} \quad (\text{A9})$$

We can expand the terms A^ρ and B^ρ by adapting (11) in the text to get

$$A^\rho = \sum_{j=1}^n \left(L_{jt}^\rho u_t^2 - \sum_{k=1}^{\infty} s^k L_{jt-k}^\rho (u_{t-k+1}^m - u_{t-k}^m)^2 \right)$$

$$B^\rho = \sum_{j=1}^n L_{jt}^\rho u_t^2 - \sum_{k=1}^{\infty} s^k L_{jt-k}^\rho (u_{t-k+1}^m - u_{t-k}^m)$$

This shows that both terms in the denominator (A^ρ and B^ρ) are weighted sums of "head counts" of workers of different tenures surviving within firm j . They are therefore linear in $\sum_{j=1}^n L_{jt-k}^\rho \cdot k = 0, 1, \dots$. Note also that setting $\rho = 1$ in the above expressions gives us the corresponding formulae for the full sample. Using (23) and (24) and taking probability limits gives

$$p \lim A^\rho/n = \rho \cdot p \lim \sum_{j=1}^n \left(L_{jt} u_t^2 - \sum_{k=1}^{\infty} s^k L_{jt-k} (u_{t-k+1}^m - u_{t-k}^m)^2 \right) /n = \rho \cdot p \lim A^1/n \quad (\text{A10})$$

$$p \lim B^\rho/n = \rho \cdot p \lim \left(\sum_{j=1}^n L_{jt} u_t^2 - \sum_{k=1}^{\infty} s^k L_{jt-k} (u_{t-k+1}^m - u_{t-k}^m) \right) /n = \rho \cdot p \lim B^1/n \quad (\text{A11})$$

Taking probability limits of (A9), using $p \lim(\frac{1}{L_t^\rho/n}) = \frac{1}{\rho} p \lim(\frac{1}{L_t/n})$, therein and using (A10) and (A11) gives a new form for (A9) as

$$\begin{aligned} p \lim \{ \text{Denominator}(\widehat{\beta}_{\min u}^\rho) \} &= p \lim \left\{ \frac{1}{L_t^\rho/n} \right\} \cdot p \lim \{ A^\rho/n \} - \left(p \lim \left\{ \frac{1}{L_t^\rho/n} \right\} \cdot p \lim \{ B^\rho/n \} \right)^2 \\ &= p \lim \left(\frac{1}{L_t/n} \right) \cdot p \lim \{ A^1/n \} - \left(p \lim \left\{ \frac{1}{L_t/n} \right\} \cdot p \lim \{ B^1/n \} \right)^2 \\ &= p \lim \{ \text{Denominator}(\widehat{\beta}_{\min u}^1) \} \end{aligned}$$

where again we have used $\widehat{\beta}_{\min u}^1$ to denote the estimate based on the full sample.

This establishes the Proposition.