Bias in Returns to Tenure When Firm Wages and Employment Comove: A Quantitative Assessment and Solution*

Andy Snell
University of Edinburgh, 31 Buccleuch Place, EH8 9JT, UK

Pedro Martins
Queen Mary College University of London,
Mile End Road, London, E1 4NS, UK and IZA

Heiko Stüber
Institute for Employment Research (IAB), Regensburger Str. 104,
90478 Nürnberg, Germany and
Friedrich-Alexander Universität Erlangen-Nürnberg (FAU)

Jonathan Thomas
University of Edinburgh, 31 Buccleuch Place, EH8 9JT, UK

November 2016

It is well known that, unless worker-firm match quality is controlled for, returns to firm tenure (RTT) estimated directly via Mincer equations will be biased. In this paper we show that even if match quality is properly accounted for there is a further pervasive source of bias, namely the co-movement of firm employment and firm wages. In a simple analytical model we show that positively covarying shocks (either aggregate or firm level) to a firm's employment and its wages cause downward bias in OLS regression estimates of RTT. We show that the long established procedures for dealing with "traditional" RTT bias do not circumvent the additional problem we have identified. We argue that firm-year fixed effects must be added to the standard Mincer equations in order to eliminate this bias. Estimates from two large panel datasets from Germany and Portugal show that the bias is empirically important. Compared with the average estimate obtained from 4 traditional methods using our correction raises RTT in Germany (Portugal) by about 2.5% (3.5%) of wages at 10 years of tenure — around 20% (45%) of the total RTT level itself. Finally we show that the results extend to tenure correlates used in macroeconomics such as the minimum unemployment rate since joining the firm.

*This work was partly supported by the Economic and Social Research Council [grant number ES/L009633/1] and the priority program 1764 of the German Research Foundation (DFG).

1 Introduction and Overview

There is a large empirical literature that attempts to identify and consistently estimate returns to firm tenure (RTT for short). The aim of this literature is to obtain the pure "causal effect of tenure on wages" (Altonji and Shakotko, 1987), the effect on the wage of one more year of tenure, holding constant years of experience and job match quality broadly interpreted. In turn this causal effect is implicitly or explicitly viewed as being a measure of the returns to firm specific human capital and/or to contractual mechanisms that reward tenure for incentive reasons. The traditional approach is to use coefficient estimates of wages on deterministic tenure in a Mincer regression to obtain a measure of RTT. This reduced form method is easy to implement and avoids having to make structural economic assumptions about worker entry and exit from the firm.

However, the existence of unobservable wage shocks that drive firm hiring and worker exit may complicate the interpretation of reduced form estimates; their existence will make tenure endogenous. Put another way, in the presence of such shocks, reduced form estimates cannot be interpreted as the causal effect of tenure on wages. Much of the past literature has focused on worker-firm match quality as the key unobservable confounding factor for RTT. In particular if we believe that better matches tend to last longer, tenure will be endogenous and failing to control for match quality will induce upward bias in reduced-form RTT estimates. Three canonical methods have been used to circumvent this problem: i) the two step estimator of Topel (1991), ii) the IV approach of Altonji and Shakotko (1987) and iii) the method of controlling for completed tenure of Abraham and Farber (1987).[1] More recently the emergence of very large panel datasets that record complete work histories of workers have allowed investigators to absorb unobserved match quality by adding firm-worker match fixed effects (see for example Battisti, 2012). The downside of doing this is that – as in Topel's (1991) method – the estimated tenure effect will include the effect of linear experience and this must be backed out using an auxiliary regression. The upside however is that it automatically controls for the impact of time invariant worker and firm heterogeneity; employing fixed effects for this purpose avoids the concern that RTT estimates may be sensitive to the investigator's selection of controls. A specification that controls for match quality using worker-firm interaction (match) fixed effects provides us with our fourth "traditional" method for eliminating the upward bias in RTT due to unobserved match quality.

In this paper we identify a further and potentially equally pervasive source of bias to RTT: the existence of a time-varying wage component that is common to all a firm's workers but that comoves with its employment. We argue that even in a world where match quality is irrelevant, the failure to account for these wage components will bias estimates of returns to tenure, and most likely in a downwards direction. The mechanism generating the bias is simple: suppose firms that have a relatively high (low) wage at time t also have high

[1] For a recent example of an application of the first two of these methods see Devereux et al. (2013).

(low) employment, high (low) hiring and low (high) average firm tenure at t.[2] This induces negative feedback from equal treatment wage shocks to tenure. In this paper we show that traditional estimators — ones designed to eliminate the effects of unobservable worker/firm match quality — are not immune to potentially sizeable biases arising from this effect.

Drivers of a firm's wage/employment comovements may include both aggregate (business cycle) shocks and firm-specific shocks. In both cases the shocks that are the root cause of the problem are assumed to impact all workers in the firm. We call these common components of wages *equal treatment shocks* following the relevant macro literature (see, e.g., Snell and Thomas, 2010, Gertler and Trigari, 2009, and Moscarini and Postel-Vinay, 2013, for macro models subject to within-firm equal treatment). Because equal treatment shocks are the same for each worker in a firm in a particular year, we propose that they be controlled for via the addition of firm-year interaction fixed effects[3] to panel wage regressions whilst at the same time also controlling for the more traditional match-quality problem.

In an empirical application we use two large samples drawn from matched panel datasets from Germany and Portugal to show that the four traditional methods produce RTT estimates that are substantially lower than that obtained using our proposed correction. If we take the average RTT estimate from the four traditional methods as a benchmark then adding firm-year fixed effects to wage equations (whilst controlling for worker-firm match quality) increases estimated RTT in Germany (Portugal) by about 2.5% (3.5%) of wages at 10 years of tenure. This amounts to about 20% (40%) of the bias-corrected RTT level itself. Although investigators may have been aware of this problem (see for example a discussion in Topel, 1991, on high wage/employment growth firms), to the best of our knowledge we are the first to quantify its importance and to propose a (simple) solution.

One interesting supplementary result from our estimation method is that the fitted firm-year fixed effects appear to follow a unit root; like unobserved match quality, the equal treatment shocks also appear to have a permanent impact on a worker's wages within a firm. Given that entry into and exit out of a firm are likely driven by permanent (rather than transient) wage shocks, this is consistent with our finding that equal treatment shocks cause bias in RTT estimates. It suggests that if one wishes to obtain the causal effects of tenure on wages, one must control for all permanent wage shocks whether they arise from equal treatment shocks or from match quality.

A further implication of our results is that using regressors that interact macroeconomic variables, such as unemployment, with deterministic tenure, will also result in biased inference. Canonical examples of such variables are Beaudry and DiNardo's (1991) minimum unemployment rate during a worker's tenure ("*minu*"), and a new hire dummy interacted with unemployment to measure the incremental cyclicality of new hire wages. The empirical

[2] There is also a steady state cross-sectional effect: high-paying firms tend to have low labour turnover, and hence longer tenure. However this type of time invariant cross-sectional effect is usually removed via the addition of firm fixed effects.

[3] Aggregate business cycle shocks can be controlled for by including time fixed effects; in fact we find that shocks below the aggregate level account for almost all of the bias.

importance of these variates found in the literature adds a further twist because their omission from Mincer equations will be yet another source of bias to RTT estimates. Another way of saying this is that if wage growth within the firm contains both the effects of human capital and implicit contracts, then to consistently estimate these separate effects requires inclusion of the relevant contract variate (e.g. *minu*) and firm-year fixed effects. We examine some of these issues in section 3 below.[4]

The key result in this paper, that there is yet another source of pervasive bias to RTT estimates obtained via reduced form estimation, may lead the investigator to conclude that a safer way to proceed is via a fully specified structural model of wage shocks and worker mobility (see Buchinsky *et al.*, 2010, for a recent example of such a model). However one key finding of our work is that it is firm specific (heterogeneous) comovement that drives the biases we find, and not macro (aggregate) effects. A structural model with heterogeneous firm hiring (and firing) — as opposed to cyclically related hiring — may be hard to specify and identify empirically. Additionally, estimates gleaned from structural models may be only as good as the veracity of their underlying assumptions. As far as reduced form modeling goes, our paper has a clear message: to avoid substantial RTT bias one must not only control for worker-firm match quality, but also for equal treatment shocks.

The paper is laid out as follows. Section 2 revisits the traditional econometric model of RTT and the implications for wages. We outline the four traditional estimation methods of Topel (1991), Altonji and Shakotko (1987), Abraham and Farber (1987) and the addition of match fixed effects. We estimate RTT for these methods using Portuguese and German panel data and plot the corresponding RTT profiles together with that obtained using our proposed correction. We then offer an anatomy of the bias from a theoretical and empirical viewpoint. Importantly, here we show that the bias is driven by heterogenous (across firms) employment/wage comovements. Section 3 looks at the implications of our analysis when contractual variables play a role, in particular tenure related maro variables associated with wage contracts. Section 4 offers concluding comments.

2 Estimates of RTT: A Comparison of Traditional Methods and the Corrected Method

In this section we estimate RTT using the four traditional methods outlined above and compare the implied RTT profiles obtained with that using our proposed correction for firm employment/wage comovements. To begin with, we revisit the bias caused by unobserved match quality and outline the methods that were designed to deal with it. We call these four methods T (Topel), AS (Altonji and Shakotko), AF (Abraham and Farber) and the addition of match fixed effects, MFE. To do so we use a somewhat simplified archetypal model of

[4]There may of course be other sources of wage growth within the firm arising from wage contracts, such as backloading, for which observable controls are not available. We discuss these issues in section 3 below.

RTT. We assume that log wages w_{ijt} for worker i in firm j at time t are given by

$$w_{ijt} = \alpha + \beta\tau_{ijt} + \gamma E_{it} + \varepsilon_{ijt}, \tag{1}$$

$$\text{with} \quad \varepsilon_{ijt} = \theta_{ij} + \varpi_{jt} + u_{ijt} \quad , \tag{2}$$

where τ_{ijt} is the worker's tenure, E_{it} is her lifetime work experience. The error consists of job match quality θ_{ij} (which also may include a worker and firm fixed effect), an idiosyncratic error, u_{ijt}, which is assumed to be uncorrelated with the regressors (especially tenure) and an equal treatment wage component ϖ_{jt} — the innovation in our study. The coefficient β is the per year RTT within the firm.[5] The traditional problem (dealt with in T, AS etc.) arises when the job match quality θ_{ij} is correlated with worker i's tenure. When the match is good (high θ_{ij}), the worker's separation hazard may fall (see in particular Bowlus, 1995), and expected tenure will rise. This makes tenure endogenous and biases β upwards. The aspiration of the traditional RTT estimation methods is to estimate the causal effect of tenure on wages in the presence of unobserved match quality θ_{ij}. The point of this paper is to show that the existence of equal treatment wage elements (ϖ_{jt}) in addition to match quality, undermine this aspiration.

Topel's (1991) method is to first-difference incumbents' wages to remove the (presumed time invariant) match quality. In the absence of ϖ_{jt}, regressing these incumbent wage changes on an intercept would, in this model at least, produce a consistent estimate of $\beta + \gamma$, $\widehat{\beta + \gamma}$ say. In order to separately identify β and γ, Topel (1991) proposed estimating a second-stage regression of $w_{ijt} - (\widehat{\beta + \gamma})\tau_{ijt}$ on the worker's initial experience on entry to the firm. Provided the latter is not correlated with job match quality, an admittedly strong assumption,[6] this produces a consistent estimate of γ. Subtracting the latter estimate from $\widehat{\beta + \gamma}$ gives a consistent estimate of β. Altonji and Shakotko (1987) proposed an IV method whereby tenure is instrumented by the deviation of tenure from its spell mean $\widetilde{\tau_{ijt}}$. By construction this variable is orthogonal to (constant within spell) match quality. Again in the absence of ϖ_{jt} this would offer consistent estimates of β.[7] Abraham and Farber (1987) propose adding duration — the final ex post tenure of the worker at the firm — as a regressor. If workers with better matches have longer completed tenure — as the traditional bias story goes —

[5] In more general contexts where RTT are heterogenous across workers and/or firms, β could be interpreted as the average RTT or average treatment effect in the words of the experimental literature.

[6] Topel (1991) argued that more experienced workers are likely to form better matches, in line with "job shopping" models of search. If true, returns to experience will be overestimated in the second stage and tenure underestimated — his RTT estimates are a *lower bound*. He considers in detail two further sources of bias: Frequent job changers may be less productive in which case more able workers' initial experience will tend to be lower, leading to γ being underestimated. Secondly jobs offering low wage growth may survive with a lower probability than higher wage growth jobs. This could lead to an overestimate of $\beta + \gamma$. Topel (1991) gives evidence to suggest that these biases are not likely to be significant; we discuss the issue further in Section 2.5.

[7] As with Topel (1991) this requires that experience is not correlated with job match quality. If it is positively correlated, again presumably because of job shopping, then the estimate of γ will be biased upwards and that of β downward biased although, they argue that this effect is relatively small (see Altonji and Shakotko, 1987, pp. 450–453).

then controlling for completed tenure directly should eliminate the bias in β.[8] Finally the MFE method adds match fixed effects to the estimation process. This focuses on within match variation in tenure. As was the case with Topel, within match de-meaned tenure and de-meaned within match experience are the same variable and the latter must be dropped from the estimation. The result is that the coefficient on tenure becomes an estimate of $\beta + \gamma$. To estimate γ — and hence β — one would use Topel's second stage (above).

All of these methods ignore the existence of the equal treatment wage components ϖ_{jt}. If these components positively comove with firm employment then they will be negatively correlated with firm average tenure and this will induce downward bias in estimates of β. We propose to augment the MFE estimator with firm-year interaction fixed effects (FYFE). The FYFE will absorb the equal treatment wage components and eliminate the bias arising from wage/employment comovements. As with MFE and Topel we use a second stage regression of $w_{ijt} - \widehat{(\beta + \gamma)}\tau_{ijt}$ on the worker's experience at entry to the firm to obtain an estimate of γ. If it is true that more experienced workers do find better matches (and this effect does have significant traction) then the estimated γ will be upward biased and RTT downward biased. At worst therefore, the RTT profile produced by FYFE will be a lower bound for the true RTT.[9]

2.1 Data

We draw our data from the German BeH and the Portuguese QP. Before discussing our subsamples we give a brief overview of these two well known data sources. We then describe the samples and the "cleaning" operations we perform on them.

The BeH data set is organized by worker spells. A spell is a portion of a year spent at a single firm. For the BeH, if a worker stays with one firm throughout the year the average daily wage for that "spell" forms a single datapoint. If the worker moves to a second firm within the year there will be two spells that year; the average wage at each firm would form a separate datapoints for that year. By contrast, the QP is an annual survey that records data on each worker at only one point in the year (census date in March up to 1993 and in October from 1994 onwards). For QP then there is only one worker "spell" per year.

The BeH draws data from the totality of gainfully employed members of the German population who are covered by the social security system. Not covered are self-employed, family workers assisting in the operation of a family business, civil servants (Beamte) and regular students. The BeH covers roughly 80% of the German workforce. We focus solely on workers employed in states of the former West Germany. Plausibility checks performed by the social security institutions and the existence of legal sanctions for misreporting guarantee

[8] Under their assumption that initial experience is only correlated with match quality through total job duration.

[9] In the presence of the mechanism identified in this paper, an additional likely upward bias exists in the estimation of γ (and hence an additional downward bias in RTT). See footnote 18 below.

that the earnings data are very reliable — in contrast with interview based wage data such as that in, say, the PSID (for the US) or the SOEP (for Germany).

Unfortunately the BeH only documents total spell earnings and not hours worked in that spell. We therefore only consider full-time workers. Nearly all full-time workers in Germany work a standard number of hours per week so the average daily wage should be very closely related to the hourly wage. To calculate the daily real wage (in 2005 prices) we use Germany's Consumer Price Index (CPI). Another problem is that wages are censored at a maximum level equal to the contribution assessment ceiling of the compulsory pension insurance scheme.[10] Earnings spells with wages above or close to (within 1% of) the truncation point are dropped. We drop all spells that have missing tenure. This means a worker only enters the data when he joins a firm after Jan. 1, 1975.[11] For this reason and in order to match the data period used for Portugal, we drop the first 12 years and use worker spells dated at 1986 and beyond.

The QP covers all workers except the self-employed and those employed in the public sector; of course, the unemployed and the inactive are also not included. There are several wage variables, all of them expressed in monthly values (the most common type of pay in Portugal), including base wages, tenure-related payments, overtime pay, subsidies and 'other payments' (this latter category includes bonuses and profit- or performance-related pay). All QP wages have been deflated using Portugal's CPI and are expressed in 2005 euros. There is also information about normal hours and overtime hours per month. The benchmark measure of pay adopted in this study is based on the sum of all five types of pay divided by the sum of the two types of hours worked, resulting in a measure of total hourly pay. Tenure — in both datasets — is measured (in rounded years) as the current year minus the reported start year.

From the QP we sample all workers from the 127 largest firms that existed throughout the entire period 1986–2009. From the BeH we sample all full-time full-year worker spells from the 100 largest (West German) firms[12] that existed throughout the same time period.[13] The motivation for using large firms is to enable us to get good estimates of the ϖ_{jt} for subsequent analysis. An additional reason to focus on a relatively small number of firms is to allow a subsequent computation of diagnostic regressions (below) that involve more than $2n$ regressors with two dimensional fixed effects. Of course estimated RTT of workers

[10] In a sensitivity analysis in Snell et. al. (2016) we found that our core result — the downward bias when FYFE are not controlled for — was robust with respect to artificially censoring the highest wages in our already censored sample. This suggest the original censorship is not impacting our results.

[11] For the analysis we only use the years 1986–2009, but for the identification of firm entrants and the calculation of firm-tenure we use BeH data from 1975 onwards. However, we exclude all spells starting Jan. 1, 1975 because the tenure could be left censored.

[12] The BeH reports establishment level data and the QP firm level data. In the paper we refer to both as "firms".

[13] Focusing on full-year spells eliminates anomalies such as supposed full time workers working for two firms at the same time and workers who have short tenured jobs. It also gives a cleaner approach to estimating within firm wage growth - especially when we use first differences (Topel) in a regression. Finally having a maximum of one observation per worker per year makes the German sample more comparable to the Portuguese.

in large firms may not be representative of RTT that exist in the economy at large. But, in an earlier version of this paper (Snell et al. , 2016) we showed that our main result — that there is substantial downward RTT bias if you fail to control for equal treatment wage components — is robust with respect to changing the sample to a) one consisting of the 1000 largest firms and to b) a randomly drawn sample of 10000 (mostly small) firms.

Despite the small number of firms, the sample still yields around 3.3 million datapoint for Portugal and 12.8 million datapoints for Germany, around 5% and 3% of the total available data from the QP and BeH respectively over this period. We also observe a substantial proportion of workers in more than one firm — just under 5% of workers in the Portuguese data and just under 10% in the German. These are higher proportions than one would expect if workers joined firms randomly. This suggests that the labour markets within which these large firms operate have a high degree of segmentation from the rest of the labour market.

Table 1: Sample Summary Statistics[14]

	Portugal	Germany
Average log monthly wage (2005 Euros)	7.01	7.91
s.d. of log monthly wage	.637	.260
s.d. of log average annual firm employment	.882	.760
Average Tenure (Years)	12.9	9.26
s.d. of Tenure	10.2	7.28
Number of Worker Spells Per Firm Per Year	1172	5278
Number of Years Available (1986–2009)[15]	22	24
Number of Tenure Categories Available	51	36
Average Firm Size in QP and BeH in 1997	9.5	14.6

Table 1 offers some summary statistics from the two samples. It shows some stark differences in the two samples and in the two labour markets. Aside from average wages being very much lower in our Portuguese sample (as we would expect) wages are over twice as variable therein. Average tenure however is substantial in both samples. Average separation rates (which can be backed out from average tenure) are around 10% per year, considerably lower than the 30% level in the US (see for example Holbijn and Şahin, 2007). The sixth and last lines give the average firm sizes in our core sample of 100 large firms and in the wider "economy" (as recorded in the QP and BeH). Firms in general are smaller in Portugal than Germany. Our samples also indicate that variation in size may be higher in Portugal than in Germany. The stark differences in the labour markets is reassuring for our analysis; if we find similar results from both countries then those results will have greater external validity than those based on a single dataset.

[14] Worker level data for 1990 and 2001 are not available from the QP. For comparison purposes we present average establishment size in both cases. Average firm size in Portugal is marginally higher at 11.0

2.2 Implementation of the Methods and Estimates

We generalise the tenure function in (1) by allowing RTT to follow a quartic function.[16] Experience is modeled via a quadratic function[17] and we control for business cycles and common trends using year fixed effects in all methods. We control for time invariant worker heterogeneity using first differences in Topel, match fixed effects in MFE and FYFE and worker fixed effects in the other specifications. We now give specific implementation details method by method.

Topel: For the first stage in Topel we estimate the following regression using data only on incumbents:

$$\Delta w_{ijt} = \delta + \beta_2 \Delta \tau_{ijt}^2 + \beta_3 \Delta \tau_{ijt}^3 + \beta_4 \Delta \tau_{ijt}^4 + \gamma_2 \Delta E_{ijt}^2 + u_{ijt}.$$

The estimate of δ, $\widehat{\delta}$ say, gives an estimate of $\beta_1 + \gamma_1 + \mu_1$ where β_1 and γ_1 are the linear terms of the quartics in, respectively, tenure and experience and μ_1 is the linear trend. We then regress the levels "residual" $w_{ijt} - \widehat{\delta}\tau_{ijt} - \widehat{\beta}_2 \tau_{ijt}^2 - \widehat{\beta}_3 \tau_{ijt}^3 - \widehat{\beta}_4 \tau_{ijt}^4 - \widehat{\gamma}_2 E_{ijt}^2$ on E_{0jt} and $(\tau_{ijt} - tr_t)$ where E_0 is initial experience on joining the firm and tr is a time index. The coefficients from this second regression ($\widehat{\gamma_1}$ and $\widehat{\mu_1}$) are consistent estimates under our assumptions, respectively, of γ_1 and μ_1. The estimate of β_1 is then obtained as $\widehat{\beta_1} = \widehat{\delta} - \widehat{\gamma_1} - \widehat{\mu_1}$.[18]

MFE and *FYFE*: For the first stage in MFE we estimate

$$w_{ijt} = \beta_1^* \tau_{ijt} + \beta_2 \tau_{ijt}^2 + \beta_3 \tau_{ijt}^3 + \beta_4 \tau_{ijt}^4 + \gamma_2 E_{ijt}^2 + \theta_{ij} + u_{ijt} \tag{3}$$

using match fixed effects to control for match quality θ_{ij}. To estimate FYFE we add firm-year interaction fixed effects to (3). Due to the addition of match fixed effects the estimated linear tenure coefficient $\widehat{\beta_1^*}$ in MFE and FYFE is an estimate of $\beta_1 + \gamma_1$. Unlike Topel, where the linear tenure coefficient also include the effect of trend, here the deterministic trend is identified separately and absorbed in the year and firm-year fixed effects respectively.[19] To obtain a consistent estimate of γ_1 we regress the levels residual $w_{ijt} - \widehat{\beta_1^*}\tau_{ijt} - \widehat{\beta}_2 \tau_{ijt}^2 - \widehat{\beta}_3 \tau_{ijt}^3 -$

[16] We also tried adding a tenure zero dummy to the quartic to capture any additional wage effect of being a new hire that the quartic specification cannot easily capture; while we find that there is a significant pay increase in the first year, in line with previous work (e.g., Altonji and Shakotko, 1987, Table 1) the impact on RTT is small in both datasets; likewise the bias we find is virtually unchanged.

[17] In Snell et. al. (2016) we found that the bias in MFE (RTT from FYFE minus that from MFE) was virtually unchanged when we generalised the experience function to a quartic.

[18] Because δ, by the reasoning of the paper, is downward biased, for this to be a consistent consistent estimate of β_1 requires an additional assumption, that initial experience is uncorrelated with duration as well as match quality (if they are positively correlated then γ_1 will be upward biased and RTT downward biased). This applies equally to the second stage of MFE below. But if experience is only correlated with match quality via its correlation with duration as assumed in AF then this is not an additional assumption.

[19] Topel argues that the estimate of deterministic trends from levels are upward biased because of the secular tendency for worker quality to improve. For this reason he uses an extraneous trend estimate. However this critique does not apply to MFE and FYFE because in those specifications match quality of every worker is controlled for via match fixed effects.

$\widehat{\beta}_4\tau_{ijt}^4 - \widehat{\gamma}_2 E_{ijt}^2$ on initial experience. This coefficient is subtracted from $\widehat{\beta_1^*}$ to give our estimate of β_1.

AF: In AF we simply add the within match variate τ_{ij}^c — completed tenure — to the main regression. For workers whose tenure is incomplete —"ongoing" workers in 2009 — we may either use imputed values or the actual value of tenure in 2009. We experimented with two imputations: i) we assumed constant exit hazards after 2009 and ii) we used the sample of workers with completed tenures to compute the expected additional tenure of workers with τ years of tenure in 2009. Both imputation methods produced profiles virtually identical to that obtained from using the value of final tenure itself and so we simply present the profile from the latter.

AS: For AS we adjust the tenure terms by subtracting their respective within match means. For example τ_{ijt}^3 becomes $\tau_{ijt}^3 - \overline{\tau_{ijt}^3}$ where denotes within match mean. These variates are used as instruments for the tenure terms in a (2SLS) IV regression.

2.3 Results

The estimates and standard errors of the quartic tenure parameters, β_i , $i = 1, \ldots 4$, for the four "traditional" methods (T, AF, AS and MFE) together with the corrected method (FYFE) are presented in Table A1 of the online appendix. The coefficient estimates are quite hard to map into RTT itself — which is the object of interest here. More informative is the RTT tenure profiles implied by these estimates. We plot these profiles for values of tenure from 0 to 20 years in Figure 1 (Portugal) and Figure 2 (Germany).

The graphs show that AS, AF and MFE offer similar RTT estimates.[20] The methods themselves are in fact quite close. Both AS and MFE measure the tenure regressors in the same way i.e. as deviations from match mean. In fact in the absence of other regressors the 2SLS tenure estimates of AS would be identical to those of MFE. As far as AF goes, if completed tenure is a good proxy for match quality, then AF will also effect an approximate within spell demeaning of the regressors.

By contrast with the other three methods,Topel's method produces RTT estimates that are quite low and the dynamic pattern is also somewhat different. Of course Topel uses a

[20] The 10-year tenure effects for the traditonal estimators are broadly in line with what Altonji and Williams (2005) find for the U.S. in their reappraisal of earlier work. The fact that tenure profiles are falling at higher tenures is not unusual in the literature. For example in Altonji and Williams' replication exercise, when, as here, the time trend is controlled for using time dummies and a quartic in tenure is included, the IV1 estimates (AS here) have a falling tenure profile above 5 or 10 years depending on the specification, while Topel's method yields the same at 10 and 20 years in one of the two specifications reported (Table 3 and footnote 12). When they adjust the relative timing of wage and tenure measures they find the tenure effect is negative above 10 years for both AS and T (Table 5).

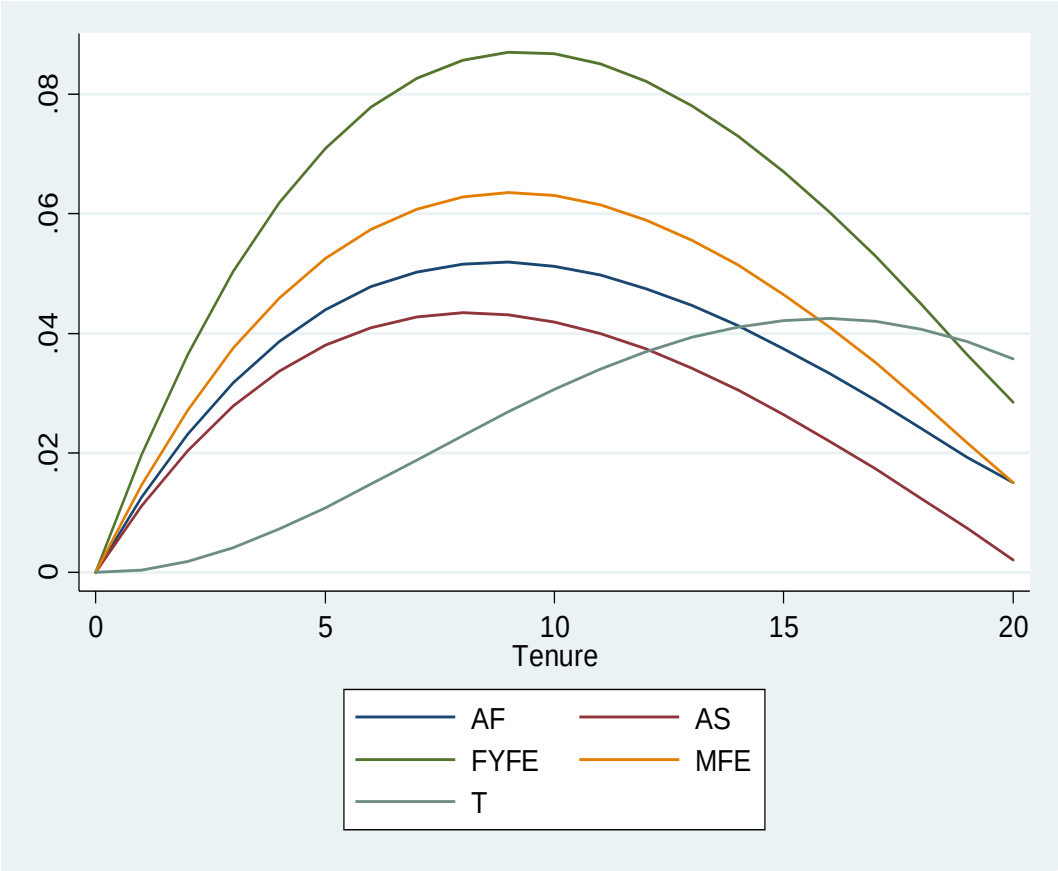


Figure 1: RTT Profiles from the 5 Methods (Portugal)

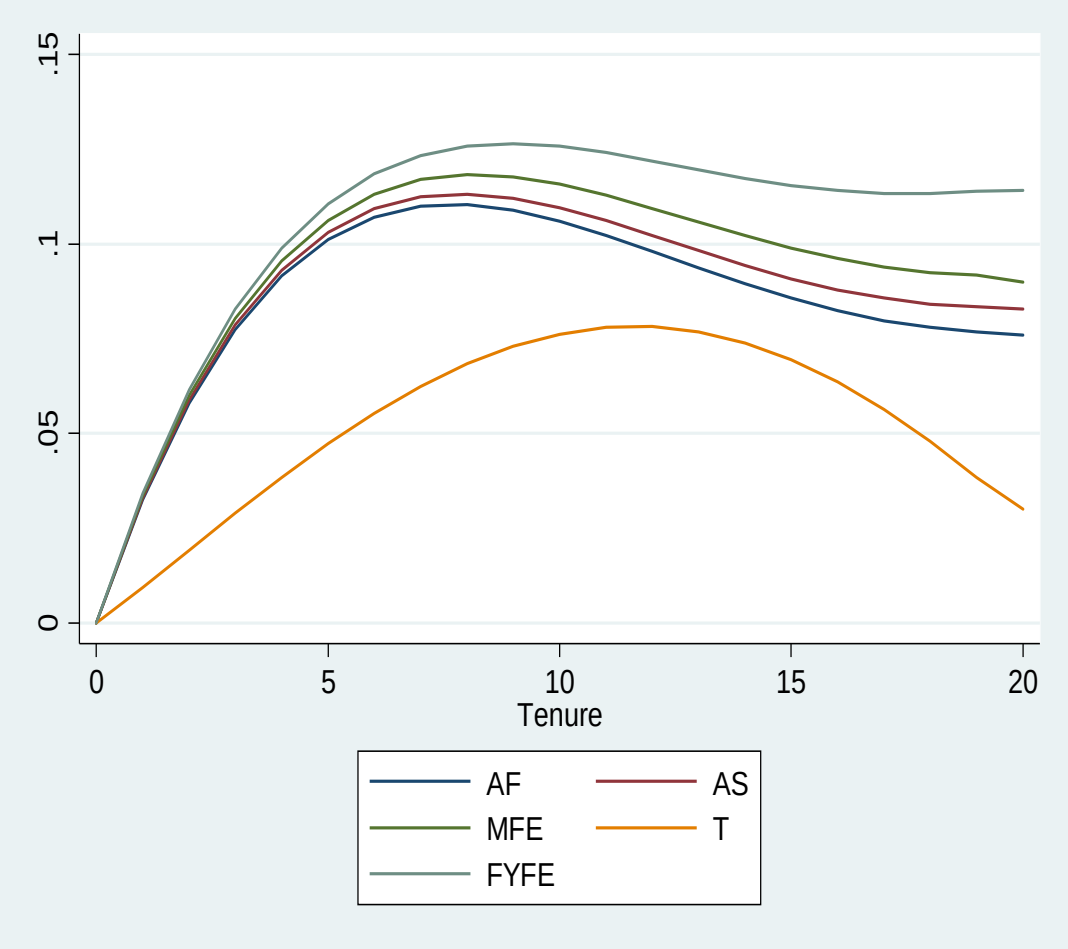


Figure 2: RTT Profiles from the 5 Methods (Germany)

first differenced specification in contrast to the levels of MFE, AF and AS - an important difference that sets Topel apart from the other methods. At the same time we should also point out that the standard errors of Topel's estimates are quite high for Portugal suggesting that the corresponding RTT schedule is not as precisely estimate as the others. For Germany, Topel's estimates are better determined but once again the corresponding RTT lie substantially below that of the other three methods and this is quite hard to rationalise.[21]

The key point however is that in both datasets the corrected RTT profile lies substantially above that of the other four methods.[22] The purest measure of the impact on RTT of adding firm-year fixed effects can be seen by looking at the vertical gap between FYFE and MFE because the two methods differ only by the application of our proposed correction. This shows a substantial bias in the case of Portuguese data with FYFE lying 2.4% of wages above MFE at 10 years of tenure – around 30% of the RTT level itself although the gap falls somewhat as tenure grows towards 20 years. For Germany it is the other way around. MFE's RTT lies only 1.3% of wages below FYFE at 10 years of tenure but the gap grows to around 2.5% of wages as tenure increases to 20 years. If we repeat these calculations using the average of the four traditional methods as a baseline then the "bias" is of course considerably larger.

We have shown that our corrected RTT profile (FYFE) lies substantially above that of the other four methods. We now try and expose and understand better the source of these differences.

2.4 An Analysis of the Source of The Bias

We argued above that positive comovement of firm employment and firm wages is a new (i.e., uninvestigated) source of bias in RTT estimates; when a firm's employment and wages rise (fall) together, its average tenure falls (rises) and tenure becomes endogenous. To get a better analytical handle on how this mechanism works we use a simple model which offers a "sketch" of the mechanism at work. The model has only a single regressor — tenure — with a regression error consisting only of equal treatment effects. Explicitly we consider the

[21] Reversing the original findings in the literature. Altonji and Williams (2005) argue that the reason for Topel's (1991) finding of a much higher RTT than previously estimated is to a substantial extent accounted for by his use of a secular wage trend from an alternative data source (using a CPS-based wage index rather than from the PSID panel he uses in his estimations, as was the case in Altonji and Shakotko, 1987, and Abraham and Farber, 1987, and here), and his use of lagged wages with current tenure. They also argue, in contrast to Topel, that individual heterogeneity biases the return to experience downwards in T (but not AS), and so RTT upwards, as discussed in footnote 6.

[22] Confidence bands are not displayed to avoid cluttering the graph; however a 95% confidence interval around Portugal's FYFE curve excludes all the other curves when tenure is below 18 years. The FYFE profile for Germany is less well defined and its confidence interval is wider; nonetheless it still excludes all of the other curves when tenure exceeds 12 years.

panel data regression of (log) wages on individual tenure

$$w_{ijt} = \alpha + \beta\tau_{ijt} + \varpi_{jt}, \tag{4}$$

where symbols are as previously defined. We ignore match fixed effects and the usual idiosyncratic regression error here because we wish to focus on the object of interest — bias caused by the existence of ϖ_{jt} and its comovement with firm employment at time t , L_{jt} say. We assume that the data comes from all workers in n large firms that exist over T years with total number of observations N $(= \sum_{t=1}^{T}\sum_{j=1}^{n} L_{jt})$. We discuss the interpretation of ϖ_{jt} below but for now and for illustrative purposes we take ϖ_{jt} to be (proportional to) a mean zero shock to firm profits.

Standard textbook theory tells us that OLS bias in the estimate of β will arise if tenure has a nonzero covariance with the error. The sample covariance of tenure with the regression error in (4) is

$$scov = \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n}\varpi_{jt}\sum_{i=1}^{L_{jt}}(\tau_{ijt} - \overline{\tau}),^{23} \tag{5}$$

where $\overline{\tau}$ is the sample mean tenure. We can rewrite the term in braces to get

$$scov = \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n}\varpi_{jt}\sum_{i=1}^{L_{jt}}\{(\tau_{ijt} - \overline{\tau}_{jt}) + (\overline{\tau}_{jt} - \overline{\tau})\}. \tag{6}$$

where $\overline{\tau}_{jt}$ is firm j's average tenure and $\overline{\tau}$ is the "long run" average tenure for all firms. The first braced summation term is by definition zero so we can simplify to get

$$scov = \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n}\varpi_{jt}\sum_{i=1}^{L_{jt}}(\overline{\tau}_{jt} - \overline{\tau}) \tag{7}$$

$$= \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n}\varpi_{jt}L_{jt}(\overline{\tau}_{jt} - \overline{\tau}). \tag{8}$$

Now suppose in year t there is positive comovement between firm j's hiring and its profit shock ϖ_{jt}. Effectively this means that firms currently experiencing above average profits (i.e..$\varpi_{jt} > 0$) will have above average employment, above average hiring and lower than average tenure. Hence $(\overline{\tau}_{jt} - \overline{\tau}_{t})$ will be negative and $\varpi_{jt}L_{jt}$ positive for such firms (vice versa for firms experiencing a negative profits shock). The net effect is to make *scov* — the OLS bias — negative

Note that the above logic would apply to a random sample (rather than a complete sample) of workers from these firms: a randomly chosen worker that has a higher than average

[23] In this illustration we abuse notation by indexing each firm's workers in the same way, i.e., worker $i = 1, ..., L_{jt}.$, when in fact the identity of workers at time t in firm j will vary from year to year.

wage is more likely to have come from a firm that has high levels of current employment (and high levels of hiring) than from one with low and that firm is more likely to have average tenure below the average for the economy as a whole. The bias argument goes through unchanged. One thing that does change when we have only a random sample of workers is our ability to reliably estimate and control for equal treatment effects ϖ_{jt}. In fact in random worker samples we are likely to see very few workers working at the same firm making identification and controlling for ϖ_{jt} practically impossible. This is one reason we chose a sample of firms rather than a sample of workers.

The preceding arguments were an analytical sketch of the main mechanism. In section A2 of the online appendix we develop a more formal model of the bias. Our benchmark model is (1) with (2). The key additional assumptions are that there are complete data on all workers in a small number of long lived large firms (offering a large number of data points in each year in each firm), that there is an exogenous worker exit/quit rate which we allow to be different in each firm, that a worker's initial experience on entering the firm is exogenous. The model also admits a completely general set of fixed effects. We find that the RTT bias is a weighted average (across firms) of the comovements between a firm's wage and its current and lagged employment levels. A key special case occurs when comovements between current wages and current employment are positive whilst those between current wages and lagged employment are zero. In this case the biases from each of the four methods are negative. When we generalise to allow current wage to comove with lagged employment as well this turns out to be relatively unimportant in determining the bias. It is the contemporaneous wage/employment comovement that matters most.

2.5 The Economic Mechanism Behind the Bias

Our primary claim is that the existence of equal treatment wage components that drive employment are an important source of RTT bias. We argue in Section 2.6 that it is movements in these components below the macro level that matter. We now try and justify that claim. First we discuss some models that are consistent with our approach. Then we look at the nature of the shocks we have identified and argue that they are consistent with our contention.

Consider a standard search-matching framework adapted to large firms (see Elsby and Michaels, 2013), with continuous bargaining.[24] Positive comovement of wages with employment requires for example the higher wages after a positive firm shock be associated

[24] To be consistent with the basic model outlined in (1) and (2) we could incorporate accumulation of general and job specific capital and random match quality, with all three translating multiplicatively into efficiency units of labor and hence wages. The latter would however depend on the bargaining protocol: The fact that a worker loses specific capital and idiosyncratic job match quality on leaving the firm would affect the outside option, so that the bargained wage may not be identical for each efficiency unit. (Elsby and Michaels, 2013, use the Stole-Zwiebel bargaining solution.) If shocks to firm productivity also affect individual productivity multiplicatively, they will affect log wages of all workers including new hires approximately equally.

with more matches being made/fewer separations; in Elsby and Michaels (2013) a positive/negative shock to a firm's productivity of sufficient size will lead to the firm increasing/decreasing its vacancies and hence hiring/laying off workers. A similar story could be told in a rent sharing or union model where positive shocks to a firm's price or productivity leads to higher employment, profits, and wages of all workers.

A number of wage posting models with on-the-job search exhibit positive tenure effects even in the absence of specific capital accumulation. For example Burdett and Coles (2003) show that with risk-averse workers wage-tenure contracts can arise, in which wages increase with tenure. The function of this backloading is to prevent turnover — firms cannot respond to outside offers but higher pay for higher tenured workers makes better outside offers less likely.[25] In equilibrium different firms start new workers at different points on the same tenure ladder. This leads to wage shopping and hence experience effects. Bagger et al. (2014) look at a related model but in which firms can match outside offers. Wages rise with tenure but stochastically, because the firm responds to outside offers (there is no point in backloading). These models for tractability typically do not have firm specific shocks of the type we have in mind (in Burdett and Coles, 2003, firms are identical, while in Bagger et al., 2014, they have different but fixed productivity). Nevertheless, in this general class of models, it would be expected that a positive firm productivity shock would increase the incentive for the firm to hire (by raising the wage profile and hence the utility of a contract offered to new hires) and to want to increase incumbent pay (to reduce turnover); so one would see highly correlated wage shocks across workers in the firm associated with an increase in employment (and vice versa for negative shocks).[26]

Models in which equal treatment (in the form of equal pay per efficiency unit) is imposed or derived, more straightforwardly lead to the empirical relationship hypothesised here (see, e.g., Snell and Thomas, 2010, Gertler and Trigari, 2009), when combined with a monopsonistic setting so that higher wages are needed to increase employment (or a competitive setting but with segmented labor markets so that positive industry shocks to productivity lead to higher industry wages and employment). In a model of on- and off-the-job search with equal treatment (so a firm cannot respond to outside offers as it pays all workers the same within a period), Moscarini and Postel-Vinay (2013) analyze the effect of aggregate shocks on wage contracts. When positive shocks occur, for example, larger firms expand more rapidly than smaller ones, and contract more rapidly in downswings (however they cannot consider idiosyncratic shocks as the equilibrium of the model can only be characterized when firm size ranking is preserved).

If mechanisms of the type described above are generating positive wage/employment comovements sufficient to underlie the bias, we would expect to find evidence that the wage

[25] Note that retention operates in the same direction: in these models, a decrease in firm wages following a negative shock, for example, will lead to workers with shorter tenure disproportionatly quitting (as they are more sensitive to outside offers), thus lengthening tenure. Likewise if the firm is laying off workers in response to a negative shock in a "last in, first out" model, tenure will lengthen.

[26] A related model to that of Bagger et al. (2014) that also has investment in specific and general training is Lentz and Roys (2015).

model we estimate generates a sufficient change in the present value of wages to attract new/retain older workers when positive shocks occur, and vice versa with negative shocks.[27] It is highly unlikely that transient shocks to wages will have any effect at all on attracting labor. By contrast permanent or highly persistent movements in a firm's wage will very likely affect its hiring, worker entry and worker exit.

Because our sample contains large numbers of workers per firm per year we can estimate each firm-year equal treatment component and examine its persistence/transience. Treating the estimated firm-year fixed effects from the FYFE specification as data we computed the first order autocorrelation coefficients (ρ say) in a balanced panel regression. The ρ values for Portugal and Germany were .92 and .99 respectively – suggesting unit root or near unit root behaviour.[28] Additionally we find that the residual from the FYFE, after eliminating the equal treatment shocks and match effects, was close to white noise (ρ values of $-.015$ and .115 respectively).[29] It seems then that – in our data at least – the two sources of permanent movements in a workers' wage within the firm appear to be the match effect (permanent by definition) and the equal treatment shocks. This finding is at odds with the assumption in Buhai and Teulings (2014) that it is within firm idiosyncratic wage shocks that drive the unit root in wages. In their model it is these shocks (together with a similar process for outside options) – not equal treatment shocks – that drive labour reallocations. Were it true that idiosyncratic shocks had this property and played this role then any attempts to estimate RTT via a reduced form (Mincer) method would be confounded at the outset because – almost by definition – we cannot control for these shocks in standard regression analysis. Our finding that idiosyncratic shocks appear to be white noise is important therefore; it is consistent with the idea that it is equal treatment shocks that drive the unit root behavior in wages instead.[30]

It would be interesting to see if the equal treatment shocks we have estimated correlate with firm productivity and/or its product price – a topic for future research. If this turned out to be true our results would be consistent with the hypothesis that these shocks drive

[27] This is particularly true of the wage-posting models which rely on wages to attract new workers/retain existing workers.

[28] Unit root behaviour of a worker's wage within a firm is a stylised fact in labour markets. See for example Buhai et al. (2014).

[29] Under the null hypothesis that the idiosyncratic wage components follow a unit root and that workers quit the firm when the value of this process falls below some value c*, we can show that the empirical autocorrelation coefficient still tends to unity, so our results strongly suggest that this can be rejected. However if the components are stationary (and again workers quit at some low threshold value) simulations suggest that the autocorrelation coefficient will underestimate the true degree of persistence.

[30] Topel (1991, pp.160-162) discusses the issue in some detail and finds no evidence that individual wage growth differences are related to contemporaneous mobility. Theoretically, if the process drivng a unit root in the wage reflects general human capital, then this should not affect mobility as outside options will move in tandem with inside returns. Likewise, timing is important: if it takes time to locate a new job after a negative wage shock, so that mobility is only affected after the period of the shock, there is no bias. Our point is that even if persistent wage shocks do affect contemporaneous mobility, so long as the persistence arises only through the equal treatment component and this is controlled for as we are proposing, there will be no bias.

hiring and that firms share rents with their workers; following a positive (permanent) shock to its product price or its productivity a profit maximizing firm would hire more workers and, under equal treatment or bargaining, the newly hired and the incumbents would get a share of the improved profits. In the next section we examine what our data has to say about the role of firm specific wage/employment comovements in generating the differences we see in Figures 1 and 2.

2.6 The Role of Firm Specific Wage/Employment Comovements

If the assumptions and arguments we make in this paper are correct then the FYFE method identifies the causal effect of tenure on wages. If initial experience is positively correlated with match quality then the FYFE method at the very least offers a lower bound on the causal effect of tenure on wages. In this section and henceforth we refer to differences between an estimated RTT profile from the FYFE method and another method as "bias". Use of this term is for convenience and it comes with the obvious caveat that it is only correct terminology if the assumptions and arguments we have made in the paper are true.

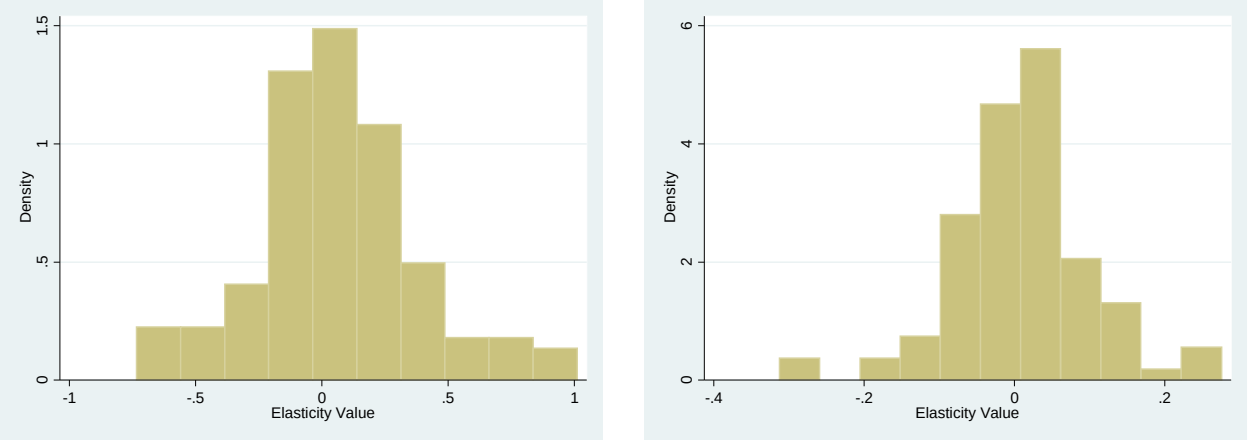
The contention of the paper is that comovement between firm wages and employment leads to a bias in the estimation of β in (1) using traditional methods, and we have found this bias to be negative and significant, implying that the comovement is positive. In principle, the positive comovement between firm wages and employment originating from the business cycle could be one source of bias in line with this logic. However in our estimates we had controlled for the business cycle via the addition of general year effects. In Snell et al. (2016) we found virtually no effect on the bias of not controlling for the business cycle in this way. The implication is that it is firm, locality or industry specific and not systemic firm wage/employment comovements that are causing the problem.

Nevertheless, to control for firm specific wage/employment comovements, the inclusion of current firm employment in the Mincer equation[31] is a possible alternative approach to adding FYFE. However it is easy to show that this will only remove the bias if the elasticities of the wage/employment relationships, are identical across firms. If there is a large amount of heterogeneity in these elasticities across firms, then this will not work.

Natural vehicles to investigate these issues empirically are the MFE and FYFE specifications; they are nested and only differ because of the addition of firm-year fixed effects. We did two exercises using these specifications. In the first we add (log of) firm employment and lagged firm employment to the MFE specification[32] allowing separate coefficients for each firm. The addition of these terms allows us to identify each firm's wage/employment

[31] Buhai et al. (2014) call the impact of firm employment on wages the firm "size" effect (although the traditional view of the firm size effect is a steady state notion). These traditional size effects would be typically absorbed using eihter firm fixed effects or match fixed effects.

[32] As before we allow for a quartic in tenure, a quadratic in experience and add year fixed effects.



(a) Portugal (b) Germany

Figure 3: Contemporaneous Firm Wage/Employment Elasticities

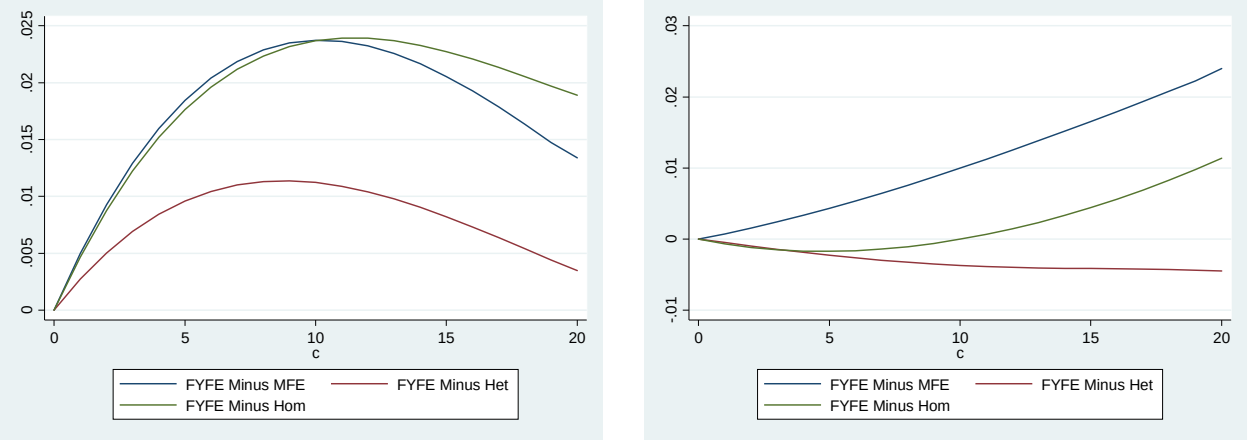
and wage/lagged-employment elasticities[33] and hence to see if these elasticities are heterogenous. Note that the addition of lagged employment terms is purely in order to obtain better estimates of the contemporaneous comovements; equation A1 in the online appendix shows that in the current context lagged employment is of second order importance to the bias, as discussed in Section 2.4. In the second exercise we add (log of) firm employment and lagged firm employment with a single (i.e., common across firms) coefficient. The idea here is that if the elasticities we found in the first exercise are homogeneous across firms then we would expect the addition of these two terms to eliminate much of the RTT bias we found in Figures 1 and 2. By contrast if there is substantial heterogeneity in the elasticities then the bias will remain. In this case we might expect the first exercise to eliminate much of the RTT bias. For clarity we call the specification in exercise one the heterogeneous specification and that in exercise two the homogeneous specification.

Histograms of the contemporaneous wage/employment elasticities obtained from the heterogeneous specification are plotted in Figures 3 (a) and (b) for Portugal and Germany respectively.[34]

The figures show two things: First that the elasticities are far more dispersed across firms in Portugal compared with Germany (the variance in Portugal is three times larger than in Germany) and second that the average elasticity — the key determinant of the bias in our

[33] In the analytical model considered in the online appendix the covariances driving the biases approximate (for small changes in wages and employment) these elasticities.

[34] The elasticities with respect to lagged employment were negative on average for Portugal but positive for Germany; histograms are shown in the online appendix. The variance was again three times higher in Portugal than Germany. As noted in the text the bias formula given in the annex predicts that in neither case do these lagged comovements matter very much in determining the bias.



(a) Portugal

(b) Germany

Figure 4: RTT Bias when Firm Employment is Controlled for

analytical model[35] — is much higher in the former than the latter (.06 in Portugal versus .01 in Germany). Given the arguments above we would expect the bias to be larger in Portugal than Germany — as indeed can be seen in Figures 1 and 2. More pertinently for the current discussion, we would expect the RTT profiles obtained from the homogeneous specification to be close to FYFE for Germany but not for Portugal. For Portugal we would expect only the heterogenous specification to deliver RTT estimates close to those of FYFE instead.

To examine these statements we compute the differences between the following RTT profiles for both countries: a) FYFE and the homogenous specification (Hom) and b) FYFE and the heterogenous specification (Het). These differences are plotted for Portugal and Germany in Figures 4 (a) and (b), respectively. For comparison purposes we also add a line representing the bias in MFE, i.e., the gap between FYFE and MFE in Figures 1 and 2. In these graphs the height above the x-axis represents the bias in each respective specification, i.e., the extent to which each respective specification fails to match the RTT generated by the FYFE specification. The "FYFE Minus MFE" line is the "baseline" bias, the "FYFE Minus Het" line is the bias from the heterogenous model and the "FYFE Minus Hom" line is the bias from the homogenous model.

We see that for Germany the homogeneous specification can eliminate quite a good proportion of the baseline bias; its profile lies halfway between the MFE line and the x-axis implying that about one half of the baseline bias has been removed. The heterogeneous specification line lies below but close to the x-axis — which we could interpret as a complete

[35] In the general version of the model where firms may have different sizes and rates of exit it is a weighted average of elasticities that matter. Only when firms are the same does the bias depend on the simple average of the elasticities. As we have only large firms in our sample we might expect them to be close in terms of size and possibly also in terms of quit rates.

removal of the bias. For Portugal things are very different. The homogeneous specification has virtually no impact on the bias — the "FYFE Minus Hom" line lies practically on top of the baseline. By contrast allowing heterogeneous comovements has far more leverage than it does in Germany — most of the bias is removed by controlling for heterogenous across firm wage/employment comovements. These results are consistent with what we predicted in the earlier discussion; the extent to which the bias may be removed by adding single coefficient employment terms depends on how homogenous cross firm wage/employment comovements are — where they are heterogenous adding employment terms with common cross firm coefficients have no impact on the bias.

2.7 Equal Treatment or Unequal Treatment?

Up to now we have focused on firm wage/employment comovements driven by wage "shocks" that are common to all workers. It is possible however that wage components of new hires alone may be causing bias and that these components are not present in incumbent wages. Suppose for example that the firm was hiring under conditions of monopsony. Suppose also that when profitability is high, hiring is high and new hires are brought in at a wage above that of incumbents. This would drive up the firm's average wage in that year and drive down the firm's average tenure. Once again we would get downward bias in RTT. But this effect is not an equal treatment effect — it is driven entirely by comovements between the new hire wage and employment. This new-hire-only effect works via the same mechanism as our equal treatment story but if it were to be the only mechanism behind the bias it suggests that a more efficient empirical procedure to remove it would focus on new hire wages only.[36]

In the light of the previous discussion, it would be interesting to see if augmenting FYFE with firm-year controls for newly hired workers only raises the RTT profile further. We call this augmented specification NHFY for convenience. Unfortunately NHFY will as a by-product also remove from the RTT profile the effects of wage growth during the first year of tenure. If the RTT gradient is steeper in the first year of tenure than in later years, removing it will move the overall RTT profile upwards. To allow for this and to be able to compare "like with like" we also strip out the initial tenure effect from the initial FYFE specification by adding a new hire dummy to it. Estimating FYFE with its new hire dummy and NHFY produced RTT profiles that were within .2% of each other (with NHFY being slightly higher in both countries).

The key takeaway of this exercise is that equal treatment wage components are the main driver behind RTT bias; adjusting additionally for movements in new hire wages has little

[36] If new hires receive a premium/discount in wages that is permanent — as would be consistent with models of full commitment by worker and firm — these will be absorbed in match fixed effects and will not affect estimates of RTT. It is short lived changes to the wages of a new hire, related to a firm's employment decisions, that we have in mind; for example, a premium in the first year of employment and thereafter being paid at some standard rate. Contracting models with limited commitment, for example, often have this short-run property (e.g., Beaudry and DiNardo, 1991, Rudanko, 2009).

incremental on the RTT profile

3 RTT and Implicit Wage Contracts

The purpose of this paper has been to derive unbiased estimates of the causal effects of tenure on wages: that is the effect on wages of experimentally increasing tenure by one year whilst keeping everything else in the economy (including for example outside options) constant. To achieve unbiasedness we have shown that it is necessary to control for equal treatment wage shocks as well as the more traditional unobserved match effects. Wages may vary with tenure for a number of reasons, not least because of returns to experience which are general market returns; but they also respond to internal offers. As discussed above, RTT estimates may be capturing the latter — a reward to the accumulation of specific capital, either human or physical. These returns might accrue to the worker for a variety of reasons: bargaining over quasi-rents for example, or a firm being prepared to respond to outside offers (distributed independently of the value of specific capital) to keep a worker with specific capital in the firm (see, e.g., Lentz and Roys, 2015).

However, the existence of quasi-rents (due to specific capital accumulation or search frictions), or the ability of firms to commit, may allow contracts in which wages do *not* correspond to marginal products in a time invariant fashion; a classic example would be backloaded wages to reduce turnover (as in Holmstrom, 1983). If there is no observable variable with which to control for contract driven wage growth (as would be implied by backloading say) then our estimates of RTT will include such effects. If this were true, only the results from a calibrated theoretical model could separately identify contract and human capital effects from estimated RTT. In this scenario our bias correction would yield consistent estimates of an "RTT+wage-contract" effect. Even in this scenario these estimates would be useful raw inputs to a calibration exercise of a theoretical model that attempts to separately identify the respective effects. If there is an observable control for wage contract effects it should be used in order to be able to identify pure firm specific human capital RTT. One class of contract models that do offer observable controls for wage contracts arises from Beaudry and Di Nardo's (1991) paper on implicit contracts.

Beaudry and Di Nardo (1991) developed a model where — modulo firm specific human capital — the minimum unemployment rate since the worker joined the firm ("*minu*" for short) was a sufficient statistic for within firm wage movements.[37] This spawned an empirical literature that added *minu* to Mincer equations to assess its importance. The findings

[37] In their analysis wages will be weakly increasing with tenure since wages are increasing with the tightest labor market conditions within the current job. Hagedorn and Manovskii (2013) argue that the results are consistent with a match quality, as opposed to an implicit contract, interpretation: better matches, which pay more, are more likely to survive periods of hightened job offers, proxied by cumulative low unemployment rates, and offer evidence to support this view. Bellou and Kaymak (2016) on the other hand find evidence for a history dependence in wages for stayers which suggests contracts do play a role.

in our paper have ramifications for this literature. *Minu* is intrinsically correlated with tenure — it falls in a weakly monotone fashion with it. Failing to control for positive firm wage/employment comovements biases the *minu* coefficient for much the same reasons as it biases RTT: higher firm wages associated with higher firm employment will lead to lower average firm tenure. Given that *minu* is negatively correlated with tenure it would be tempting to state that this bias is positive (towards zero for a negative coefficient). But in Snell et al. (2016) we argued that the inclusion of tenure in the regression complicates the bias and in general it cannot be signed. Nevertheless, using the Portuguese data we showed that adding firm-year fixed effects to a specification such as MFE that also includes *minu* dramatically affects our inferences; the coefficient on *minu* falls (in absolute value) from a highly significant value of $-.93$ to a borderline significant value of $-.29$. Whatever the sign of the bias in this context, equal treatment wage components should be controlled for; at best they are unwanted noise and at worst they cause bias.

Finally there is a recent empirical literature that tries to establish the extent to which the contract hiring wage is sensitive to conditions at the time of hiring. In this literature focus is on the significance of a measure of the state of the labour market (typically aggregate unemployment) and a "new hire dummy". If this variable — "*deltau*" for short — is found to have a significantly negative impact on wages it implies that firms take advantage of poor labour market conditions when hiring. As with *minu*, *deltau* is negatively correlated with tenure and once again failure to control for firm-year fixed effects would cause its estimated effect to be biased.

To sum up this discussion, controlling for the effects of wage contracts is crucial to be able to identify human capital RTT; but controlling for equal treatment components of wages is essential to get good estimates of both.

4 Conclusion

We have shown in this paper that the positive comovement of equal treatment wage components and firm employment causes significant bias in RTT. We showed that our equal treatment shocks are highly persistent (unit root or near unit root processes) and that controlling for them significantly changes the RTT estimates. This is important as we would expect only persistent wage shocks to drive firm quits and firm hiring. We concluded that match quality and our equal treatment shocks are two of a kind — persistent shocks to wages that impact a worker's tenure. Finally we found that controlling for these two shocks reduces the residual error to (near) white noise. This is consistent with the argument that adding firm-year fixed effects and match fixed effects to Mincer equations is necessary to control for those wage components that are jointly endogenous with tenure. This gives us some confidence that our bias corrected reduced form estimates of RTT are causal. We conclude with some additional ad hoc observations arising from our work.

First, if one is purely interested in effects that vary only with year and tenure then equal treatment shocks are "noise" and removing them seems a sensible thing to do. Once match quality is controlled for only the cross tenure/year movements in wages are relevant to estimating RTT; components of wages that are common to workers in firm j in year t cannot add information to this.

Second, our FYFE correction allows for the possibility that firms may have heterogeneous wage and employment co-trends. Fast (slow) growing and high (low) wage growth firms would have lower (higher) average tenure and higher (lower) average wages. This type of issue has been discussed before in the RTT literature but as far as we know it has not been analyzed.

Third, in this paper we focused on MFE as a "baseline" specification or method. But in fact we could add firm-year fixed effects to any of the three other methods to control for the bias we have identified.

Finally, the need to control for FYFE would seem to rule out the use of small random samples of workers to obtain unbiased RTT estimates. Such samples are unlikely to contain two workers in the same firm. Just how many workers per firm are required to remove the bias effectively is unclear and a subject for future research.

References

[1] Abraham, K. and H.S. Farber (1987), "Job Duration, Seniority and Earnings", American Economic Review, Vol 77. No. 3 June: 278–297.

[2] Altonji, J.G. and R.A. Shakotko (1987), "Do Wages Rise with Job Seniority", The Review of Economic studies, 54(3): 437–459.

[3] Altonji, J.G. and N. Williams (2005), "Do wages rise with job seniority: a reassessment", Industrial and Labor Relations Review, Vol. 58, No. 3 April: 370–397.

[4] Bagger, J., Fontaine, F., Postel-Vinay, F. and J.-M. Robin (2014), "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics", American Economic Review, 104(6): 1551–1596.

[5] Battisti, M. (2012),"Mobility and Wages in Italy: The Effects of Job Seniority" in *Labour Markets at a Crossroads: Causes of Change, Challenges and Need to Reform*, Lindberg, H. and N. Karlson (eds.), Cambridge Scholars Publishing: 205–232.

[6] Beaudry, P. and J. DiNardo (1991), "The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Micro Data", The Journal of Political Economy, 99(4): 665–688.

[7] Bellou, A. and B. Kaymak (2016): "Real wage growth over the business cycle: contractual versus spot markets," Discussion paper, Universite de Montreal, Working Paper.

[8] Bowlus, A.J. (1995), "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality", Journal of Labor Economics, 13(2): 335–350.

[9] Buchinsky, M., Fougere, D., Kramarz, F. and R. Tchernis (2010), "Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States", Review of Economic Studies, 77(3): 972–1001.

[10] Buhai, I.S. and C.N. Teulings (2014), "Tenure Profiles and Efficient Separation in a Stochastic Productivity Model", Journal of Business and Economics Statistics, vol. 32(2): 245–258.

[11] Buhai, I.S., Portela, M.A., Teulings, C.N. and A. van Vuuren (2014), "Returns to Tenure or Seniority?", Econometrica, 82(2): 705–730.

[12] Burdett, K. and M. Coles (2003), "Equilibrium Wage-Tenure Contracts", Econometrica 71 (5): 1377–404.

[13] Devereux, P.J., Hart, R.A. and J.E. Roberts (2013), "Job spells, employer spells, and wage returns to tenure", University of Stirling, Dept of Economics DP.

[14] Elsby, M.W.L. and R. Michaels (2013), "Marginal Jobs, Heterogeneous Firms and Unemployment Flows", American Economic Journal: Macroeconomics, January, Vol. 5(1): 1–48.

[15] Gertler, M. and A. Trigari (2009), "Unemployment Fluctuations with Staggered Nash Bargaining", Journal of Political Economy, 117(1): 38–86.

[16] Gertler, M., Huckfeldt, C. and A. Trigari (2014), "Unemployment Fluctuations, Match Quality and the Wage Cyclicality of New Hires", Cornell University Dept of Economics DP.

[17] Hagedorn, M. and I. Manovskii (2013): "Job selection and wages over the business cycle," The American Economic Review, 103(2): 771–803.

[18] Hall, R. (2005), "Employment Fluctuations with Equilibrium Wage Stickiness", American Economic Review, 95(1): 50–65.

[19] Hobijn, B. and A. Şahin (2007), "Job-Finding and Separation Rates in the OECD", Federal Reserve Bank of New York Staff Reports no. 298.

[20] Holmstrom, B. (1983), "Equilibrium Long-Term Labor Contracts." Quarterly Journal of Economics, 98(3): 23–54.

[21] Lentz R. and N. Roys (2015), "Training and Search On the Job", University of Wisconsin-Madison Discussion Paper.

[22] Michaillat, P. (2012), "Do Matching Frictions Explain Unemployment? Not in Bad Times", American Economic Review, 102(4): 1721–1750.

[23] Moscarini, G. and F. Postel-Vinay (2013), "Stochastic Search Equilibrium", Review of Economic Studies 80 (4): 1545–1581.

[24] Rudanko, L. (2009), "Labor Market Dynamics under Long-Term Wage Contracting", Journal of Monetary Economics, March, vol 56 (2): 170–183.

[25] Snell, A. and J.P. Thomas (2010), "Labor Contracts, Equal Treatment, and Wage-Unemployment Dynamics", American Economic Journal: Macroeconomics, 2(3): 98–127.

[26] Snell, A, Martins, P., Stueber, H. and J. Thomas (2016) " Bias in Returns to Tenure When Firm Wages and Employment Comove: A Quantitative Assessment and Solution", IZA Discussion Paper no. 9849.

[27] Topel, R. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority", Journal of Political Economy, 99(1): 145–176.