Unobserved individual and firm heterogeneity in wage and tenure functions: Evidence from German linked employer-employee data*

Thomas Cornelißen and Olaf Hübler†

This version: September 2006
Work in Progress.

Abstract

We estimate wage and job tenure functions that include individual and firm effects capturing time-invariant unobserved worker and firm heterogeneity. We use German linked-employer employee data (LIAB data set) and implement a memory-saving way to estimate the person and firm effects allowing us to estimate all firm effects that are identified in our sample (about 800 firm effects).

We find that both types of heterogeneity are correlated to the observed characteristics and that it is therefore warranted to include individual and firm fixed effects in both the wage and the job tenure equation.

We look primarily into the correlation of the unobserved heterogeneity components with each other. We find that high-wage workers tend to be high-tenure workers. At firm level, there seems to be a trade-off between wages and job stability: Low-wage firms tend to be high-tenure firms. High-wage workers and high-tenure workers, if they are not the same persons, seem to be matched to different types of firms. While high-wage workers tend to be matched to low-wage/high-tenure firms, high-tenure workers tend to be matched to high-wage/low-tenure firms.

*We thank the staff of the research data center (Forschungsdatenzentrum) of the Federal Employment Services (Bundesagentur für Arbeit) at the Institute of Employment Research (IAB) for their support and helpful comments in using the data. Financial support by the DFG (project no. HU 368/4) is gratefully acknowledged.

†Institute of Empirical Economics, University of Hannover, Königsworther Platz 1, 30451 Hannover, Germany, cornelissen@ewifo.uni-hannover.de, huebler@ewifo.uni-hannover.de

1 Introduction

This paper is concerned with two key features of the employment relationship: pay and job stability. We estimate the determinants of wages and job tenure by taking into account unobserved individual and firm heterogeneity using linked employer-employee data. The interest in doing this is, first, to get unbiased estimated of the structural parameters of the equations by controlling for unobserved heterogeneity and, second, to analyse how the unobserved individual and firm characteristics are correlated with the observed characteristics and with each other. In fact, the focus of the paper is on this second aspect.

With the availability of linked employer-employee data sets many researchers have investigated individual and firm effects in wage equations. To our knowledge, nobody has yet estimated both individual and firm effects in a job mobility equation. The novelty of the paper therefore lies in estimating tenure functions that control for unobserved worker and firm heterogeneity and by analysing the association of individual and firm effects from tenure equations with those of wage equations. While the empirical literature on individual and firm effects has looked into whether high-wage workers are employed in high-wage firms, we extend the research by asking whether high-wage workers are high-tenure workers, whether high-wage firms are high-tenure firms, and whether high-wage workers work in high-tenure firms etc[1].

The paper proceeds as follows. Section 2 develops expectations from a theoretical perspective about the association of individual and firm effects from wage and tenure functions. Section 3 refers to related literature. The data set is described in section 4 and the model and estimation problems are discussed in section 5. Empirical results follow in section 6 and section 7 concludes.

2 Theory

In this section we develop hypotheses about the interdependence of individual and firm effects in wage and tenure functions. The effects are those components of wages

[1] Abowd, Kramarz and Roux (2006) estimate simultaneous wage and mobility equations. While taking into account time-invariant unobserved person and firm heterogeneity in the wage equation, the mobility equation takes into account firm-specific effects but does not include person effects.

and tenure that are due to unobserved time-invariant individual and firm characteristics.

Individual and firm wage effects are often interpreted as reflecting person-specific and firm-specific productivity (Abowd et al. 2004) that is due to unobserved characteristics. These unobserved characteristics can be personal abilities and character traits for individuals and the quality of management for firms. The firm-specific wage effect does not in the first place have to reflect productivity, but it may also reflect the wage policy of the firm, the management's credo to incentives and their productivity effects. In other words, the firm wage effect may reflect whether the management believes that efficiency wage theory is relevant. In the empirical application, factors that are principally observable, but that are not available in the data set, can be captured by the unobserved heterogeneity effects as long as they are time-invariant during the period of observation. For example, with our firm data we cannot adequately control for the capital endowments of firms. As the capital endowment can be expected to be quite stable over the limited range of time covered by the data, the firm wage effect may also capture the productive effects of capital or of the capital intensity of production.

In the discussion we will refer to "good workers" and "good firms", whereby we mean workers and firms with a high person and firm effect estimated from the wage equation, i.e. workers and firms with unobserved time-invariant characteristics that lead to higher wages.

When thinking about the meaning of individual and firm effects in tenure functions, it is useful to consider that the decision to continue an employment relationship is a joint decision of the employer and of the employee.

If job stability and mobility where mainly driven by employees' decisions, we would interpret individual tenure effects as capturing preferences for job stability. Firm tenure effects would then capture firm characteristics that make employees willing to stay. This could be various kinds of working conditions.

If job stability and mobility where mainly driven by employers, individual tenure effects would reflect personal characteristics and abilities that lead employers to retain workers in the firm. These could be similar abilities as those that determine the individual wage (productivity) effect. Firm tenure effects would in this case capture the firm-specific need of fluctuation and adaptation of the skill-composition of their

work-force or the cost of fluctuation.

Based on these arguments on the meaning of individual and firm effects in wage and tenure functions we now discuss the interdependence of these effects.

2.1 Individual wage and firm wage effect

When workers and firms are heterogeneous in their productive capacity, a simple assignment model (Becker 1973) implies positive assortative matching between workers and firms. Abowd et al. (2004) derive a model in which the productive capacities of workers and firms lead to separate linear individual and firm effects in a log wage equation. Positive assortative matching can then be tested empirically by computing the correlation of individual and firm effects measured from a log wage equation. A positive correlation is compatible with positive assortative matching.

However, there are at least two possible explanations for negative assortative matching (Barth and Dale-Olsen 2003). First, as argued above, it may be that the firm effect captures the productivity effects of the capital endowment of firms. If a very high capital productivity (e.g. high-quality machines) and a very high worker productivity (i.e. workers with high ability) are substitutes, then one would expect to observe a negative correlation. Second, if individuals cannot distinguish between the unobservable individual and firm effects and base their behaviour only on the sum of both, then more productive workers are likely to be in a good-paying job earlier and to stop searching earlier. Low productivity worker will search longer and therefore end up more likely in high-productivity firms.

More elaborate models of labour markets with frictions (see Shimer 2001, Shimer and Smith 2000 and Postel-Vinay and Robin 2002) have different implications for positive or negative assortative matching depending on the assumptions.

2.2 Individual tenure and firm tenure effect

We are not aware of theoretical work that deals explicitly with assortative matching with respect to fluctuation and mobility of employers and employees. However, if we understand the individual tenure effect as a preference for mobility and the firm tenure effect as need for fluctuation, we would argue the following. Low-mobility

workers would derive disutility from working in high-fluctuation firms, and firms that wish to reduce fluctuation would find it costly to employ mobile workers. Therefore it seems that positive assortative matching is an efficient market allocation, as allocating high-mobility workers to high-fluctuation firms minimises the total cost to production.

If we understand the person tenure effect as a productivity effect and the firm tenure effect as a working conditions effect, then we would expect positive assortative matching, too, as firms with better working conditions manage to attract better workers.

2.3 Individual wage and individual tenure effect

As mentioned above, if all firms try to keep workers with good unobserved worker characteristics for a long time, then the tenure person effect would measure similar person characteristics as the person wage effect. In this case we expect high-wage workers to be high-tenure workers. However, in the empirical implementation we control for the wage rate when estimating the tenure effects. All good worker characteristics that are already compensated through higher wages can therefore not show up in the tenure effect. A positive correlation could therefore imply that good characteristics are partly compensated by wages and partly compensated by job stability.

On the other hand, high abilities also enable workers to have better outside opportunities and therefore may cause them to change employers more often. This would lead to a negative correlation between individual wage and individual tenure effect.

2.4 Firm wage and firm tenure effect

We argued that the firm tenure affect can be determined by the firm's need for and cost of fluctuation, or they may indicate good working conditions. Firms that have characteristics associated with high costs of fluctuation may pay high wages in order to keep fluctuation low. That would lead to a positive association of the firm wage and the firm tenure effect. A positive correlation would also result if firms that are

more efficient and productive than others offer higher wages but also offer better (non-wage) working conditions which generally reduce quits from their workforce.
On the other hand, firms need to compete on product markets with each other, which should reduce differences in the efficiency of production. One could also suspect a trade-off at firm level between wage costs and fluctuation in the opposite direction as in the preceding argument: in order to stay competitive firms that pay high wages require a more frequent adjustment of the skill composition of their workforce. A reason can be that paying higher wages than comparable competitors requires a more continuous and sustained process of investments and product innovation which also continuously requires an adaptation of the skill composition of the workforce, hence mobility.

2.5 Individual wage and firm tenure effect / Individual tenure and firm wage effect

The arguments for the remaining associations can be deduced from the arguments for the associations already discussed. For example, if high-wage workers are high-tenure workers and if high-wage firms are high-tenure firms, and if there is positive assortative matching, then we would expect also the individual wage effect to be positively related to the firm tenure effect, and the individual tenure effect to be positively related to the firm wage effect.
Suppose that high-wage firms are not high-tenure firms, because there is a trade-off at firm level between wages and fluctuation. Which worker types would then prefer the high-wage firms and which the high-tenure firms? High-wage workers might prefer high-tenure firms, because due to their individual abilities they can command relatively high wages even in a low-wage firm where at the same time they enjoy higher job stability. We would then have a positive association between individual wage and firm tenure effect.

3 Related literature

Several studies have estimated individual and firm effects in wage equations. In "High wage workers and high wage firms" Abowd, Kramarz and Margolis (1999) show that both person and firm effects are important, but that person effects are more important in explaining the variance in wages, as well as in explaining the firm size effect and the industry effect on wages. Firms that hire "high-wage workers" are more productive, more capital intensive and more high skilled labour intensive and have higher chances of survival. Firm that are "high wage firms" are more productive and more profitable as well as more capital intensive and more high skilled labour intensive, but they have a lower chance of survival. One key finding of Abowd, Kramarz and Margolis (1999) is that the correlation between individual and firm effects on wages is positive but not large (between 0.08 and 0.14), i.e. they find weak evidence for positive assortative matching. However, in that paper the authors use an approximative method, the reliability of which they test in a later paper (Abowd, Creecy, Kramarz 2002). In that later paper they compute the exact least squares solution to the problem and find the importance of person effects confirmed, but the correlation of person and firm effects turns into a negative correlation.

Practically all subsequent studies that have computed individual and firm effects in wage equations find negative correlations between the two effects (Andrews et al. 2006b, Alda 2006, Barth and Dale-Olsen 2003, Grütter and Lalive 2004, Goux and Aurain 1999). This has spurred a debate on economic (Abowd et al. 2004) and statistical (Andrews et al. 2006b) explanations for that (apparent) negative assortative matching.

We are aware of studies that take into account individual heterogeneity in quit and separations equations using models of binary choice (Anderson and Meyer 1994, Frederiksen 2004) but not of studies that estimate individual fixed effects in job tenure equations. Unobserved firm heterogeneity in tenure and job duration models is taken into account in Mumford and Smith (2002) and Gerlach and Stephan (2004, 2006). These studies use linked employer-employee data, but the data sets are not constructed as panels, i.e. they do not allow for the estimation of individual effects alongside with the firm effects. Furthermore, again due to data restrictions, the studies are focused on elapsed tenure as a dependent variable, whereas the more

appropriate measure is completed tenure if job stability is to be investigated. From the weak correlation of firm effects with observable person characteristics, Mumford and Smith (2002) conclude that there is no evidence that long-tenure workers sort into long-tenure workplaces.

Grotheer et al. (2004) analyse the determinants of job stability in German linked employer-employee data by estimating job duration models and taking into account the competing risks of unemployment versus job-to-job change. However, they do not control for unobserved heterogeneity. Boockmann and Steffes (2005) estimate job duration models taking into account competing risks in German linked employer-employee data. They control for unobserved firm heterogeneity but not for unobserved worker heterogeneity.

Abowd, Kramarz and Roux (2006) estimate simultaneous wage and mobility equations. They find that in 30-40% of the firms in their sample the individual effect from a wage equation reduces the probability of separations, while in 10% of the firms it increases separations. Other interesting findings include a positive correlation of the intercepts from firm-specific wage an mobility functions ("high-wage firms are high-mobility firms"), and the finding that large firms tend to hire workers with low individual wage effects. While taking into account unobserved time-invariant person and firm heterogeneity in the wage equation, the mobility equation only includes time-invariant unobserved firm heterogeneity.

To date, there seems to be no investigation into the determinants of job tenure or job mobility that controls for both, unobserved individual and firm heterogeneity with respect to mobility. However, omitting one or the other is likely to produce biased estimates (Abowd, Kramarz and Margolis 1999).

4 Data

We use the West German sample of the first version of the longitudinal model of the German linked employer-employee data set LIAB provided by the Institute for Employment Research (IAB). Alda et al. 2005 give an overview of the LIAB data set. This LEE data set links the survey data of the IAB establishment panel to employee registry data from the employment and unemployment benefit registry of

the German employment service. The individual characteristics that are provided by the registry data are few in comparison to the ample firm characteristics provided by the IAB establishment panel. A variable list and descriptive statistics of the variables we include into the analysis are provided in table 7 in the appendix.

In order to describe the data set design it is useful to define three types of firms. IAB firms are those that are covered by the IAB establishment panel survey, which started in 1993. A sub-group of these, namely all firms that have been part of the IAB establishment panel survey continuously at least from 1999-2001 or from 2000-2002 may be called LIAB firms. Thirdly, there are non-IAB firms, which are those that are not covered by the IAB establishment panel.

The employee side of the LIAB data version we use includes all persons that have been employed in any of the LIAB firms at least one day between 1996 and 2001. For all of those employees the complete set of employment and unemployment benefit spells between 1991 and 2002 is provided, regardless in which type of firm employment took place. This implies that there are many spells in non-IAB firms, i.e. spells for which apart from a firm identifier, no firm characteristics are available. It also follows from this design that, while for the LIAB firms between 1996 and 2001 practically the hole workforce is in the data set, for all non-LIAB firms there are typically only a few employees represented in the data. We therefore restrict our analysis to the set of LIAB firms.

As period of observation we chose 1996 to 2002 because for LIAB firms observed before 1996 it is not assured that the whole set of employees is in the data.

Using a sub-sample of firms reduces the chances to observe a worker who changes job in his new firm. The extent of observed worker mobility between firms is therefore limited. Out of about 650,000 workers we observe 3,500 workers in more than one firm ("movers").

The employee data comes in the format of employment and unemployment spell data. In this analysis we exploit only the employment spells. Each firm in Germany has to notify the social security authorities at least once every year (typically at the end of the year) about who is employed and how much each employee has earned on average since the last notification. The earnings information, however, is reported only up to the social security contribution threshold and we therefore have right-censoring of the earnings variable. Apart from the wage, firms also report the

education, sex, age, job position, profession, full-time status and other information of each worker (see variable list in table 7 in the appendix). A typical employment spell of a continuously employed person refers to the time period from the 1st of January to the 31st of December of a given year. In the case that a firm gives notifications during the calendar year separate spells are generated. There are several reasons for notifications, some of which are of more interest for our analysis (begin, end, interruption of employment etc.) and some of less interest (change of the health insurance company etc.). Some interesting events (change in the wage paid) give not rise to an extra notification, but they show up eventually in the next notification.

The tenure variable has to be constructed from the longitudinal employment information that is available for each employee back to 1991. Employment relationships that began before 1991 are coded in the data as beginning on the 1st of January 1991. Likewise, for employment relationships that continue after 2002 we do not observe the end date. Consequently, we have left-censored and right-censored employment spells in the data.

We define separations (terminations of employment relationships) by (i) interruptions of employment with the present employer of longer than 30 days (i.e. the separation is followed by a recall), (ii) changes of the employer identifier (i.e. the separation is followed by a job-to-job move) and (iii) no subsequent employment spell recorded (i.e. the separation is followed by unemployment or inactivity). It follows from our definition of a separation that a worker can have several employment spells with the same employer. About 13% of workers in the data set have experienced a recall[2].

The firm data has an annual panel format and the survey is recorded around the 30th of June of each year. For all employee spells that begin before the 30th of June of a given year we match the firm data from that year's survey. For the spells beginning after the 30th of June we match the firm data of the following year's survey. We configure two versions of the data set. The first version keeps several observations

[2]Evidence of Mavromaras und Rudolph (1995) based on the same underlying data source, albeit for the time period before 1990, shows that 12% of all newly started employment relationships in Germany are recalls. According to their findings, recalls occur mostly in sectors with seasonal fluctuations, and are more frequent for blue collar workers in the case of men and part-time workers in the case of women.

per employer-employee match and therefore captures the variation of time-varying characteristics during the match. There is at least one observation per year, but if there are several notifications during a year with economically interesting changes, then we keep these as different observations. We could establish an annual panel by choosing as the annual observation the spell that refers to the 30th June. But by doing this we would lose variation in the data and we would lose short employer-employee matches that last for periods of less than a year if they do not extend up to the 30th of June of a given year. We cannot afford to lose employer-employee matches because they may concern movers and therefore contribute to the identification of firm effects. Therefore we keep the single spells as the level of observation. Because we now have several observations per employment relationship, the tenure variable in this version of the data set is elapsed tenure.

The second version of the data set is one where we only keep one observation per employment relationship. The tenure variable in this case captures completed tenure. In order to retain one observation per employment relationship we retain either the value of a variable at the end of the employment relationship (tenure, wage, age), its mean value (firm size, export share, sum of investments, business growth) or its mode (most categorial variables). We replace some indicator variables by an indicator that captures whether the variable has been affirmative at least once during the employment relationship (we do so for indicators with respect to investment into IT, use of part-time work and use of fixed-term contracts). The incidence of recalls mentioned above implies that workers can have several employment relationships with the same employer, i.e. also "stayers" (those who are observed only at one employer) can have several observations of completed tenure[3].

We estimate wage and tenure equations in both versions of the data set. Version one of the data set is our preferred sample for the wage equation, because there is more variability in the data, and version two for tenure equations, because the variable of interest is completed tenure.

[3]When estimating wage equations it is doubtful whether one should start to count tenure from zero after a recall, because a worker is unlikely to lose all firm-specific human capital during the interruption. But when estimating job stability it seems justified to count recalls as separate employment relationships because employment relationships that are characterised by recalls are instable employment relationships. They may be so due to firm decisions, i.e. in the case of seasonal workers, or due to employee decisions, i.e. in the case of motherhood leaves.

In order to take possible biases in the estimation of tenure into account, we introduce further sample restrictions. Due to the design of the data set we worry about left- and right-censored employment relationships. We eliminate left-censored spells by restricting all tenure estimations in both versions of the data set to employment relationships that began after the 1st of January 1990[4].

The remaining right-censored employment spells cause interruption bias (Salant 1977), i.e. we observe elapsed tenure instead of completed tenure. In version one of the data set we deliberately estimate elapsed tenure and therefore do not adjust for this bias[5]. In version two, however, we principally observe complete tenure. Only those employment relationships that are ongoing after 2002 are right-censored. In order to minimise this problem we restrict tenure estimations in version two of the data set to all non-censored employment relationships plus all right-censored employment relationships of workers older than 55 years. We argue that for this age group, elapsed tenure is very close to completed tenure.

For the estimation of the wage equations we do not impose these restrictions, but alongside with the tenure variable we introduce an explanatory variable that indicates censoring of tenure. Both left-censoring and right-censoring of the employment relationships leads to right-censoring of tenure. Therefore the indicator variable is 1 for both kinds of censoring and 0 if no censoring is present.

We restrict our sample to full-time workers, because there is no information on the hours worked in the data set and wages of part-timers are therefore not comparable between workers, and for part-timers the information on the job position (blue-collar / white-collar) is missing. We restrict the minimum age to 16 and base the analysis on employees in regular employment defined as employment subject to social security contributions. With respect to missing values in the person and firm characteristics we base the analysis on the complete cases.

[4]Due to the data set design, employment relationships that begin between 1990 and 1996 can only be observed if they are ongoing after 1996. Therefore one could argue that we over-sample long job durations. On the other hand, keeping the employment relationships that started between 1990 and 1996 allows us to observe uncensored completed job durations up to 13 years. Basing the analysis only on those employment relationships that started after 1996 would reduce the maximum tenure observed to 7 years.

[5]A number of studies estimate elapsed tenure, e.g. Mumford and Smith (2002) and Gerlach and Stephan (2004,2006).

Table 1 gives an overview over all 4 estimation samples. In our preferred sample for the wage equation 650,000 person effects and around 800 firm effects are identified. In the preferred sample for the tenure estimation these are 260,000 person effects and 570 firm effects.

5 Model and Estimation

5.1 Non-linearity and fixed effects

We estimate the determinants of wages and tenure alongside with individual and firm effects that may be correlated with the observables.

Concerning the choice of our estimation method we are in a dilemma of choosing between non-linear models and of estimating and predicting individual and firm fixed effects. Some aspects of our estimation problem call for non-liner models. First, as described in the preceding section there is right-censoring of the wage variable which calls for a censored normal regression. Second, in order to analyse job stability with our tenure variable we would preferably want to employ a model of duration analysis with censoring. There are some duration models that can be expressed as linear models in the logarithm of duration and estimated by a linear regression model. Apart from leading to inefficient estimates as compared to the maximum likelihood method on a non-linear model, this procedure does not lend itself readily to take into account the right-censoring of employment spells (Lancaster 1990, pp.219).

The dilemma comes along as we intend to estimate explicitly person and firm fixed effects in a very large linked employer employee data set. Unfortunately, non-linearity poses a problem in this endeavour for two reasons.

First, in many cases fixed effects methods applied to non-linear models lead to inconsistent estimates due to the incidental parameter problem (Neyman and Scott 1948, Greene 2003, pp. 690). In some non-linear models the unobserved heterogeneity term can be eliminated by conditional, marginal or partial likelihood methods. This is the case in the conditional logit model (see for example Greene 2003, p. 698) and in some duration models (Lancaster 1990, pp.263). For censored dependent variables Honoré (1992, 1993) proposes two estimation methods. Even though these different

models do allow to find consistent estimates for the structural parameters, they do not allow to estimate the fixed effects themselves, which is our aim. Furthermore, these methods are developed to take into account one fixed effect and they do not seem to have been extended to include several effects, e.g. person and firm effects. This may in some cases actually be impossible, because even in the linear case there does not seem to be an algebraic transformation that allows to sweep-out and to recover both person and firm effects at the same time[6].

Second, the maximum likelihood estimator of non-linear models typically does not lead to an explicit set of normal equations, but it is solved through iterative methods. Iterative maximum likelihood estimation has a disadvantage when compared to least squares estimation in very large data sets. With large data sets it can be necessary to decompose the estimation problem. The least squares estimator can be decomposed observation-wise (Ritchie 1995), but this is not possible with the maximum likelihood estimator when it is solved iteratively.

These reasons restrict our analysis to linear models that are estimated by least squares techniques[7]. But at the same time we try to take into account the problems of possible biases of the tenure equations by an appropriate choice of the estimation sample (see previous section)[8].

[6]If it was only to sweep-out the heterogeneity one could estimate a fixed effect for each unique worker-firm match and thus only have one effect to control for. However, recovering the the person and firm effect from this match effect is not possible (Andrews, Schank, Upward 2006a).

[7]When it comes to the trade-off of taking into account non-linearity versus fixed effects, it may in some cases be more important to take into account fixed-effects. Although in a different context, Ferrer-i-Carbonel and Frijters (2004) find that including fixed-effects affected their results more strongly than taking into account the ordinal character of the dependent variable.

[8]In this paper we ignore the right-censoring of the wage variable. Other options would be dropping the censored cases or imputing the censored values by a censored normal regression. This model is a non-linear model where we would not be able to take into account the unobserved heterogeneity. While the first solution is not optimal because it can lead to sample selection biases, the second is not optimal, because if unobserved heterogeneity is important, the imputation model is biased.

5.2 The estimation of person and firm effects

We estimate a model of the form

$$y = X\beta + D\theta + F\psi + \epsilon, \tag{1}$$

where the dependent variable y is either the wage rate or tenure, X $(N^* \times K)$ is the design matrix of time varying characteristics; F $(N^* \times J)$ is the design matrix for the firm effect; and D $(N \times N)$ is the design matrix for the person effect. N^* is the number of person-years in the dataset, J is the number of firms, N is the number of persons, and K is the number of time varying regressors. The coefficient vector β captures the effects of observed time-varying worker and firm characteristics (including time effects). Our main quantities of interest are the individual effects θ and the firm effects ψ.

The assumption under which we estimate the model is that the error term is orthogonal to all regressors, including the individual and firm effects. This implies that the matching of workers to firms does not systematically depend on the shocks incorporated in ϵ, i.e. that mobility is exogenous (see Grütter (2006) for a discussion of this assumption and an estimation strategy that relaxes the assumption).

Unfortunately, there seems to be no simple algebraic transformation to sweep out individual and firm effects as for example the within-transformation ("time-demeaning") in the fixed effects model. It is possible, however, to include the firm effects as dummy variables and to sweep-out the person effects by the within-transformation. The within-transformation consists of subtracting the person mean from each observation. This eliminates the person effect. The transformed model may be written as

$$\tilde{y} = \tilde{X}\beta + \tilde{F}\psi + \tilde{\epsilon}. \tag{2}$$

When estimating this mixture of the within-transformation and the dummy variable model we encounter the problem that the design matrix becomes too large to fit the available computer memory (see also Andrews et al. 2006a, Abowd et al. 2002 and Grütter 2006 for methods to estimate person and firm effects in large data sets).

The design matrix of the time-demeaned model $(\tilde{X}, \tilde{F})$ has dimension $(N^* \times (K+J))$. In our largest sample we have approximately $N^* = 2,500,000$, $K = 50$ and a number of identifiable firm effects of $J = 800$. Consequently, the design matrix has 2.125

billion cells ($2,500,000 \cdot 850$). Assuming that 4 bytes of memory are needed to store each cell, the memory requirement amounts to 8.5 GB. However, the cross-product matrix $(\tilde{X}, \tilde{F})'(\tilde{X}, \tilde{F})$, which is eventually needed for estimation has only a dimension of $850 \times 850 \approx 700,000$ and requires memory of below 3 MB[9]. In Cornelißen (2006) it is developed how to construct $(\tilde{X}, \tilde{F})'(\tilde{X}, \tilde{F})$ without creating the F part of the design matrix. In our case, using this procedure lowers the memory requirement from 8.5 GB to below 1 GB.

Firm effects are identified through the mobility of workers between firms. Of the 1,941 firms in our largest sample, only 821 firms have "movers", i.e. workers that are observed in more than one firm of the data set. No firm effects can be estimated for the 1,120 firms without movers. The 821 firms with movers are divided into 21 groups of firms, which are defined such that firms within one group are connected by worker mobility, but firms of different groups are not connected by worker mobility (see Abowd, Creecy and Kramarz (2002) for an algorithm to determine the groups). If N_g is the number of persons in a group, and J_g the number of firms, then in each group $N_g - 1$ person effects, $J_g - 1$ firm effects and a group mean are identified. The $J_g - 1$ firm effects and $N_g - 1$ person effects within each group are arbitrary in the sense that they depend on which person and firm in each group serves as the reference. Therefore, person and firm effects should not be compared between different groups. One can normalise person and firm effects within each group by subtracting the mean person and the mean firm effect respectively, so that they sum to zero within each group and represent deviations from the group mean. Even after normalisation it is in our view not straightforward to compare person and firm effects between groups, because it is questionable whether a firm effect of $+1$ in a group with a very high group mean means the same as a firm effect of $+1$ in a group with a low group mean. Therefore, we base our subsequent analysis of the person and firm effects only on effects out of the same group. For this purpose we use the largest group which contains the majority of the observations. In our largest sample this group contains 90% of the observations and 772 of the 800 identified firm effects.

After the estimation of the person and firm effects we can study the correlations of

[9]But the X part of the design matrix will still require memory. Using the numbers given above, this would be an additional 500 MB

these effects with observable characteristics and the correlation of the effects among each other. Under the assumption that individual and firm effects are not correlated with the other regressors, Andrews et al. (2006b) and Abowd et al. (2004) show that $Corr(\theta, \psi)$ is biased downwards if there is true positive assortative matching. With arbitrary correlation between the unobserved heterogeneity and the other regressors, the sign of the bias cannot be determined a priory, but it is an empirical question. The formulae to compute the exact bias when there is arbitrary correlation between unobserved heterogeneity and observed characteristics require the inversion of an $N^* \times N^*$ matrix (Andrews et al. 2006b), which is computationally not feasible with the size of our data set. Andrews et al. (2006b) propose the alternative of assuming that the observable regressors are uncorrelated with the unobservable heterogeneity. This assumption does not really fit the framework of a fixed effects estimation where one explicitly allows for such correlation. Therefore we prefer not to compute the bias under this assumption but to exploit a different finding of Andrews et al. (2006b). They show that the bias in the estimation of the correlation decreases if the number of movers increases. We therefore compute the correlations based on a sub-set of all firm effects that are identified by at least 30 movers. We furthermore also use only those individual effects of persons for whom we have at least 2 observations, because person effects of persons who have only one observation are poorly estimated.

5.3 Selection of explanatory variables

At the present stage of our research we have specified the wage and tenure equations with the same set of regressors, which we see as key determinants of wages and job stability. In the continuation of this research we plan to include further important characteristics, such as whether a firm has a works council and whether a firm provides training to the work force. The studies of Grotheer et al. (2004) and of Boockmann and Steffes (2005) show that these firm characteristics have stabilising effects on employment relationships. The effects of the omitted variables may be taken up by the effects of the observed characteristics or by the unobserved heterogeneity effects, depending on whether the effects are correlated with the omitted

variables. We suppose that in our present specification, the effects of works council and employer provided training are to a large part taken up by other regressors, such as firm size and collective bargaining and not so much by the unobserved firm effects.

5.4 Endogeneity

Wages and job stability are likely to be determined simultaneously. Wages usually increase with the duration of the employment relationship as specific human capital is accumulated. On the other hand, the wage development is likely to be a determinant of the continuation of the employment relationship. The payment of high wages reduces quits and, if it can be interpreted as the quality of the match or as a measure of productivity, it is also likely to reduce layoffs. Simultaneity biases the estimates if both equations are estimated separately by ordinary least squares.
In the analysis of this paper we do not take simultaneity into account. However, in subsequent work we plan to estimate wages and tenure simultaneously. The challenge for simultaneous estimation is finding suitable instruments for the wage rate and tenure, in other words to specify the wage and tenure equation with different sets of regressors in order to achieve identification. Abowd and Kang (2002) discuss different models that deal with the endogeneity of wages and tenure.

6 Results

Table 2 reports results of the pooled and fixed effects regressions of the wage and tenure equations in the preferred samples.
Most coefficients of the pooled and fixed effects estimation of the wage equation in table 2 (first two columns) are according to expectations. We only mention a few effects that surprised us: Somewhat against our expectations is the steepness of the cubic function of age, reaching its maximum at about 25 years of age[10]. Furthermore, we find that not only EU foreigners but also other foreigners seem to earn

[10]We suspect that this result is driven by interactions between the highly non-linear tenure specification, the age specification and the year dummies. If we specify tenure linearly, we estimate a tenure coefficient of 0.022 and an age parabola that reaches its maximum at 48 years of age.

higher wages than comparable workers of German nationality. An explanation could be that workers who migrate, after controlling for education, job position and profession, are a positive selection of workers. A further result is that firms that use part-time work pay lower wages to their full-time staff (recall that only full-time workers are in the sample). Due to considerable fixed-costs per employee, firms that use part-time work have higher wage costs than firms that employ the same labour input only in terms of full-time work. We suspect that some of that cost is shifted onto full-time workers in the form of lower wages. According to our results, fixed-term work increases wages. We suspect that the flexibility gained through fixed-term work makes firms more competitive, because they can adapt better to fluctuations in demand for their product. This allows them to employ less workers on average and to pay higher wages per employee.

Turning to the comparison of the pooled wage estimation with the fixed effects wage estimation (first two columns of table 2), we see that introducing person and firm fixed-effects into the estimations has an effect on the coefficients. If a coefficient decreases after introducing fixed effects, we can conclude that it is biased upwards in the pooled wage regression, and if it increases it is downwards biased in the pooled regression. If there is an upwards bias in the pooled regression then we expect the regressor in question to be positively correlated to the person and/or firm effects. Likewise, if there is a downwards bias in the pooled regression, then we expect the regressor to be negatively correlated with the person and/or firm effects. The first two columns of table 3 report the correlations of the individual and firm wage effects with the observed characteristics. We now use these correlations to asses whether the biases that appear in the pooled wage regression in table 2 are due rather to unobserved person or unobserved firm heterogeneity or both.[11]

The coefficients of education (Vocational training and A-levels as well as University degree) and professional status (skilled-blue collar and white collar workers) appear upwards biased in the pooled wage regression. This can mean that workers of higher

[11] We use bivariate correlations. However, the omitted variable bias of the coefficient of a given observed characteristic depends on the correlation of that observed characteristic with the omitted characteristic holding all other observed characteristics constant. Therefore our procedure is not exact, but we use the bivariate correlations as first indicators and plan to look at partial correlations or regression coefficients later.

education and higher professional status are "better workers" or that they work in "better firms"[12]. The correlations reported in table 3 suggest that university graduates and white collar workers are "better workers" but not that they tend to work in "better firms", and that skilled blue-collar workers are not "better workers" but that they tend to work in better firms: University degree and white collar job position are positively related to the individual wage effect, but not to the firm wage effect. Being a skilled blue-collar worker is positively related to the firm wage effect but not to the individual wage effect.

Returning to the first two columns of table 2, we note that the reduction in the effect of firm characteristics on wages after controlling for fixed effects suggest that the coefficients of the following firm characteristics are upwards biased in the pooled wage equation: pay above collectively bargained wages, business expectations, investment sum, use of fixed-term work and adherence to both types of collective bargaining. Again, firms with these characteristics seem to be "better firms" or to employ "better workers". For all but firm-level collective contracts, the correlations in table 3 suggest that these firms are "better firms", because the observed characteristics are positively related to the firm wage effect. Firm-level collective contracts, however, are not implemented in "better firms" but in firms that employ "better workers". Applying a firm-level collective contract seems to be effective in attracting good workers.

The firm-size effect appears to be biased downwards in the pooled wage regression in the first column of table 2 when compared to the fixed-effects specification in the second column. The negative correlation of firm size with individual wage effects reported in table 3 suggests that the downward bias comes from the fact that larger firms tend to employ workers with lower person wage effects, n finding that is also reported by Abowd, Kramarz and Roux (2006) who use French data. In table 3 it is evident that larger firms have higher firm wage effects. But as we observe the downward bias in the pooled wage regression, the negative correlation with the person wage effect seems to outweigh the positive correlation with firm effects.

[12]Recall that we employ the term "good workers" and "good firms" to workers and firms with high person (firm) effects estimated from the wage equation, i.e. workers (firms) with unobserved time-invariant characteristics that lead to higher wages, and can be assumed to indicate higher productivity.

The correlations of the person and firm wage effects with the observed characteristics in table 3 reveal more interesting associations: there is a positive correlation of male sex and of being a non-EU foreigner with firm wage effects. Part of the gender wage differential and of the positive wage effect of non-EU foreigners reported in the pooled wage regression can therefore be explained by sorting into high-paying firms. In the case of non-EU foreigners the higher firm wage effect seems to outweigh the lower person wage effect.

We have no immediate explanation for the finding that firms that invest into IT tend to have lower firm wage effects and to employ workers with lower individual wage effects. A possibility is reversed causality: Firms with lower productive capacity need to invest into IT technology in order to stay competitive.[13] We find that in general sectors are stronger correlated with firm wage effects than with person wage effects. Firms with high firm wage effects seem to be more often in sectors such as investment goods and credit and banking, and less often in the restaurant and hotel as well as the education and publishing sector.

The last two columns of table 2 report the pooled and fixed effects estimation of completed tenure. Tenure increases with wages, and men and foreigners seem to have more stable employment relationships. Workers with vocational training have more stable employment relationships than those with no vocational training. This effect vanishes, however, after controlling for fixed effects. The coefficient appears upwards biased in the pooled regression. Workers with vocational training seem to have unobserved characteristics associated with job stability, or they work in firms with unobserved characteristics associated with lower fluctuation. The correlations in table 3 (last two columns) suggest that it is the person effects and not the firm effects that cause the upward bias of the coefficient on vocational training. The opposite seems to be the case for university graduates. The regression coefficients

[13] If we had a sufficiently long period of observation the negative correlation might disappear because investments could lift the productivity of the firm to a higher level in the future and thus lift the firm wage effect. If the negative correlation would not disappear in a data set with a very long period of observation, then we would have to interpret this in the sense that the productivity effects of investments are only temporary, i.e. they do not increase the firm effect, and that firms with lower firm effects therefore have to reinvest more often into IT than firms with higher firm wage effects.

in table 2 reveal that after controlling for fixed effects the coefficient on university education rises, suggesting that university graduates have unobserved personal characteristics associated with more mobility or that they work in firms with unobserved characteristics associated with higher fluctuation. The negative correlations of university education with both person and firm tenure effects in in the last two columns of table 3 suggest that both is the case.

Returning to the last two columns of table 2, we note that with the exception of the sum of investments, the coefficients on firm characteristics all increase after controlling for fixed effects, some turning from negative to positive. The coefficients on these firm characteristics are obviously biased downward in the pooled regression. This suggests that firms with more or higher values of the observed firm characteristics in question either employ more mobile worker types, measured in terms of individual effects, or they themselves have unobserved characteristics associated with higher fluctuation, measured in terms of firm effects. The correlations of the firm characteristics with person and firm tenure effects in the last two columns of table 3 reveals that both is the case for most of the firm characteristics: large firms, firms that pay above collectively bargained wages, firms with good business expectations, firms that invest into IT, firms that use part-time or fixed-term work and firm that adhere to sector-level collective contracts have on average lower person and firm tenure effects.

In the last two columns of table 3 we find that male sex is correlated with higher individual tenure effects but with lower firm tenure effects. The higher stability of male employment found in the pooled regression of completed tenure in table 2 seems to be a person effect and not to be due to sorting of men into more stable firms. When looking at the correlations of industries with firm tenure effects, table 3 reveals that high firm tenure effects are most common in the mining and energy sector, and least common in the investment goods sector.

In table 6 in the appendix the regression results for wages and tenure in the respective non-preferred samples are presented.

In the wage equation estimated in data set version two (first two columns of table 6) the coefficients of some firm characteristics are insignificant, and some are significant but display the opposite sign when compared to the estimation in data set version

one discussed before (table 2)). These include firm size (in the fixed effects specification) and investments into IT. That some coefficients become insignificant in the estimation in data set version two can be due to the loss of variability in the data that comes about from retaining only one observation for each employment relationship (see section 4 for the explanation of our two versions of the data set). Different signs may come from the fact that for some regressors, the one observation retained for each employment relationship has been constructed as the mean value over the whole employment relationship (firm size, export share, sum of investments, business growth), the mode (most categorial variables) or as an indicator variables that captures whether the variable has been affirmative at least once during the employment relationship (investment into IT, use of part-time work and use of fixed-term contracts). Varying this procedure still leaves scope to find better specifications, which will be part of our further research.

The equation of elapsed tenure (last two columns of table 6) is in some respects in line with the estimation of completed tenure discussed before(table 2), but there are also very marked differences in single results. For example, in the elapsed tenure regression the difference between male and female workers is much smaller, and university graduates seem to have lower job stability. An explanation for the differences can be that the estimation of elapsed tenure is likely to give different results from completed tenure if the composition of the workforce changes over time. For example, an increase of the share of university graduates in the labour market leads to more university graduates observed with values of low tenure, as they start their careers. This may well explain the negative sign on the coefficient of university graduates in the pooled elapsed tenure equation in columns 3 of table 6, but it has nothing to do with lower job stability of university graduates. We therefore prefer not to interpret in detail the coefficients of the elapsed tenure equation. However, in the following it turns out that the correlations among the unobserved components of the wages and tenure equations are quite similar, whether we base the calculation on elapsed or on completed tenure.

Tables 4 and 5 report the correlations of individual and firm effects estimated from the wage and tenure functions presented above. We report partial correlation coefficients holding the observed characteristics age, sex and nationality constant,

because these characteristics are not included in the fixed effects estimation and their effects are therefore taken up by the person and firm effects[14].

The results from the two tables are similar and can be summarised as follows (the correlations in parentheses are taken from table 5):

1.) High-wage workers tend to be high-tenure workers (0.26).

2.) High-wage workers tend to work in low-wage firms (-0.12).

3.) High-wage workers tend to work in high-tenure firms (0.11).

4.) High-tenure workers tend to work in low-tenure firms (-0.28).

5.) Low-wage firms are high-tenure firms (-0.13).

(Whereby "high-wage" and "high-tenure" are shortcuts for high unobserved time-invariant wage and tenure effects.)

The correlation not mentioned in these result is the one between individual tenure effect and the firm wage effect. It is the lowest correlation in each of the two tables 4 and 5 and its sign varies between the two tables. As it is not robust over the two specifications we do not include it into the interpretation of the results.

Result 2 is consistent with findings from other studies (Alda 2006, Andrews et al. 2006b, Abowd et al. 2002). The fact that the correlation of individual and firm effects from the same equation is likely to be biased (Andrews 2006b, Abowd et al. 2004), casts some doubt on whether the negative correlations in results 2 and 4 are really due to negative assortative matching, or whether they are statistical artefacts.

[14]Even tough age is not time-constant, the within-variation of age is an increase of 1 for each person each year, similar to a time trend. The between-variation of age is taken up by the person effect. This would also be the case if age was explicitly included into the fixed effects estimation. As is visible in table 3, age is highly correlated with the individual wage and the individual tenure effect. Correlating the individual wage and the individual tenure effect without holding age constant gives us very high correlations of >0.9 but these are essentially spurious correlations. Therefore we hold age and the other time-invariant observed characteristics constant when computing the correlations in tables 4 and 5.

Therefore, results 2 and 4 should be interpreted with caution, as we cannot determine the exact size and the direction of the bias[15].
The correlations across equations are unlikely to be biased, as worker and firm effects from different equations do not directly depend upon each other. Our key results are therefore results 1, 3 and 5, namely that high-wage workers are high-tenure workers and that they work in high tenure-firms, which are low-wage firms. This last result contradicts the findings of Abowd, Kramarz and Roux (2006), who find that low-wage firms are high-mobility firms[16].

We propose the following interpretation of the results: Low-wage firms compete via cost reduction strategies, whereas high-wage firms compete via other strategies (e.g. product innovation), which require a more frequent adjustment of the skill composition of their workforce in order to stay competitive. Therefore, high-wage firms are low-tenure firms. Workers derive utility from both, high wages and high job stability. High-wage workers are endowed with a higher income potential. Therefore, high-wage workers (especially if they are not at the same time high-tenure workers) chose to use part of it in order to buy higher job stability and therefore prefer employment in low-wage firms, which are high-tenure firms. Alternatively, due to their innate ability, high-wage workers simply reach a satisfactory wage level more quickly and therefore stop searching for higher-paying firms more quickly. In order to receive a satisfactory pay level, low-wage workers have to search longer and will therefore have greater chances to find a high-paying firm.
High-tenure workers may be workers with characteristics that make firms try to keep them in the firm (see discussion in section 2). As such, high-tenure workers can expect a relatively high job tenure even in low-tenure firms and therefore the high-tenure workers (especially if they are not at the same time high-wage workers)

[15]Andrews et al. (2006b) and Abowd et al. (2004) estimate the size of the biases in their estimations under simplifying assumptions and find that the size of the bias is not large enough to turn the negative correlations into positive ones.

[16]Major differences between the study of Abowd, Kramarz and Roux (2006) and our study are that they estimate wages and mobility simultaneously while we estimate separate equations. Furthermore, their measure of mobility is the probability of separations while elapsed job tenure is held constant, whereby our mobility equation is estimated with completed job tenure as dependent variable, and they include no individual fixed effect for mobility.

chose to enjoy the higher wages of low-tenure firms.

Result 1 suggests that quite frequently high-wage workers are high-tenure workers. In this case, we cannot say which type of firm they would be matched to, because according to result 3 they would be matched to high-tenure firms and according to result 4 they would be matched to low-tenure firms. A way to learn more about this would be to include the observed and the unobserved characteristics into a multivariate analysis of explaining who tends to work in high-tenure firms and who tends to work in low-tenure firms.

7 Conclusion

Our aim was to learn about unobserved individual and firm heterogeneity in wage and tenure functions. We have estimated individual and firm effects that capture time-invariant unobserved heterogeneity in wage and tenure equations and related these effects to observed characteristics and to each other.

We have circumvented the computer memory restrictions that are often encountered when identifying a large number of individual and firm effects in linked employer-employee data sets by applying a memory-saving way to compute the full least squares solution of the problem. Thus we could estimate all identified firm effects in our sample, which are about 800 firm effects.

Our results match the results of previous studies in that we have found a negative correlation between the individual and firm effects from the wage equation. We have found a similar negative correlation between the individual and firm effects from the tenure equation. Both results are consistent with negative assortative matching, but we have interpreted these results with caution because they can be subject to statistical bias. When relating the effects from different equations to each other, we have found that high-wage workers tend to be high-tenure workers. At firm level, there seems to be a trade-off between high wages and high job stability: high-wage firms are low-tenure firms, probably because a high-wage policy requires a more frequent adjustments of the skills of the work force and hence mobility.

We have found that high-wage workers (if they are not at the same time high-tenure workers) tend to work in high-tenure firms, which tend to be low-wage firms. We

have interpreted this as a strategy of high-wage workers to use their income potential to buy job security. High-tenure workers, on the other hand (if they are not at the same time high-wage workers), tend to be matched with the opposite type of firms, i.e. with low-tenure firms, which tend to be high-wage firms.

Our analysis can be refined in several ways. In future research we plan to take the simultaneity of wages and tenure into account by estimating a system of simultaneous equations. This will require to find different specifications for the wage and tenure equations because exclusion restrictions are needed for identification. Testing more specifications will also shed light on the question of how robust the correlations between the unobserved effects are over a range of different specifications.

Tables

Table 1: Overview of estimation samples

Sample:	Wage 1	Wage 2	Tenure 1	Tenure 2
Level of observation:	Notification (version 1)	Employment relationship (version 2)	Notification (version 1)	Employment relationship (version 2)
Restriction:	No	No	Yes [a)]	Yes [b)]
Observations:	2,540,084	762,941	1,612,944	312,320
No. Persons:	648,085	655,098	466,962	264,057
thereof movers:	3,588	3,632	3,134	1,544
thereof >1 obs.:	523,013	86,076	354,132	36,012
No. Firms:	1,929	1,941	1,918	1,871
thereof with movers	816	821	796	612
connected groups	20	21	27	43
Identified firm effects	796	800	769	569

[a]) Restriction: Begin of employment relationship after 1st January 1990.

[b]) Restriction: Begin of employment relationship after 1st January 1990 and right-censored employment spells in 2002 of over 55 year-olds only.

Table 2: Estimation of wage and tenure equations (preferred samples)

Dependent Variable:	Log Wage		Log Tenure	
Sample	Wage 1		Tenure 2	
Model:	Pooled	P+F FE	Pooled	P+F FE
Wage	-	-	0.006	0.004
	-	-	89.6	22.5
Tenure	0.059	0.046	-	-
	266.8	158.1	-	-
Tenure squared/100	-0.347	-0.167	-	-
	-185.5	-86.2	-	-
Ten. cens. (dummy)	0.005	-0.145	-	-
	7.2	-60.0	-	-
Age	0.078	-	0.600	-
	97.7	-	83.7	-
$Age^2/100$	-0.152	-0.108	-1.339	-1.273
	-76.7	-22.3	-72.4	-7.8
$Age^3/1000$	0.099	0.059	0.986	0.963
	61.6	15.6	65.0	7.4
Male	0.203	-	0.083	-
	349.6	-	13.4	-
EU foreigner	0.017	-	0.087	-
	15.0	-	6.6	-
Other foreigner	0.011	-	0.052	-
	12.6	-	5.9	-
Voc. Training	0.065	0.064	0.221	0.033
	108.0	17.5	34.2	0.5
Voc. Training and A-levels	0.113	0.081	-0.025	-0.005
	85.9	11.3	-1.8	0.0
University	0.211	0.157	0.115	0.234
	208.8	22.1	9.7	1.8
Skilled blue-collar	0.061	0.013	0.224	0.268
	92.9	6.0	27.6	3.5
White collar	0.174	0.078	0.196	0.163
	165.6	24.6	15.4	1.5
Log firm size	0.022	0.041	-0.054	0.828
	101.4	24.2	-22.4	15.4
Firm pays > tarif wage	0.039	0.002	-0.253	-0.165
	62.7	2.6	-36.4	-5.1
Business expectations	0.024	0.019	0.223	0.405
	11.9	9.5	9.4	4.1
IT investments (dummy)	0.0004	0.002	0.399	0.569
	0.6	2.4	46.2	20.5
Investments (in 10 mill. Euros)	0.002	0.001	0.004	-0.014
	64.0	33.2	9.0	-5.8
Firm uses part-time work	-0.007	-0.003	0.353	0.711
	-6.0	-2.3	23.6	14.2
Firm uses fixed-term work	0.008	0.001	0.333	0.666
	12.1	0.7	38.7	22.3
Sector-level coll. contract	0.039	0.003	0.159	0.771
	33.2	1.7	13.3	12.1
Firm-level coll. contract	0.055	0.009	-0.062	0.244
	40.3	4.5	-4.3	3.2
Constant	2.413	-	-9.652	-
	231.9	-	-106.7	-

Note: All specifications include year, sector and profession dummies.
Reference categories are: Female, German nationality, No vocational training, unskilled blue-collar, no IT investment, no part-time, no fixed-term, no coll. contract.
T-values indicated below coefficients, "P+F FE": Person and firm fixed effects.

Table 3: Correlation of unobserved time-invariant effects with observed characteristics

Effect:	Individual wage		Firm wage		Individual tenure		Firm tenure	
Sample:	Wage 1		Wage 1		Tenure 2		Tenure 2	
Wage	-		-		0.141	***	-0.107	***
Tenure	0.334	***	0.029	***	-		-	
Age	0.911	***	-0.057	***	0.979	***	0.119	***
Male	0.116	***	0.142	***	0.073	***	-0.114	***
EU foreigner	0.005	***	0.012	***	0.006		-0.046	***
Other foreigner	-0.091	***	0.030	***	0.006		-0.030	***
Voc. Training	-0.004	***	-0.002	**	0.038	***	-0.033	***
Voc. Training and A-levels	-0.047	***	-0.023	***	-0.102	***	0.025	***
University	0.078	***	-0.002	**	-0.042	***	-0.016	***
Skilled blue-collar	-0.013	***	0.081	***	0.080	***	-0.095	***
White collar	0.190	***	-0.149	***	-0.005		0.214	***
Skilled manual	-0.025	***	0.056	***	0.014	***	-0.105	***
Technical	0.129	***	0.030	***	0.037	***	0.036	***
Unskilled services	-0.010	***	-0.091	***	0.088	***	0.053	***
Skilled services	-0.013	***	-0.086	***	0.041	***	0.149	***
Semi-Professional	0.007	***	-0.103	***	-0.033	***	0.164	***
Professional	0.045	***	-0.079	***	-0.023	***	0.070	***
Unskilled administrative	0.039	***	-0.196	***	-0.027	***	0.148	***
Skilled administrative	0.057	***	-0.014	***	-0.024	***	-0.012	**
Manager	0.094	***	-0.007	***	0.034	***	-0.011	**
Log firm size	-0.045	***	0.113	***	-0.016	***	-0.657	***
Firm pays > tarif wage	-0.007	***	0.283	***	-0.047	***	-0.546	***
Business expectations	-0.035	***	0.079	***	-0.033	***	-0.123	***
IT investments (dummy)	-0.014	***	-0.023	***	-0.031	***	-0.061	***
Investments (in 10 mill. Euros)	-0.057	***	0.096	***	-0.015	***	-0.498	***
Firm uses part-time work	-0.014	***	-0.056	***	-0.027	***	-0.091	***
Firm uses fixed-term work	0.017	***	0.015	***	-0.070	***	-0.259	***
Sector-level coll. contract	-0.066	***	0.311	***	-0.055	***	-0.361	***
Firm-level coll. contract	0.076	***	-0.321	***	0.055	***	0.360	***
Year 1996	0.067	***	0.027	***	0.064	***	-0.021	***
Year 1997	0.065	***	-0.054	***	0.020	***	0.013	**
Year 1998	0.033	***	-0.045	***	-0.004		0.081	***
Year 1999	-0.028	***	0.043	***	-0.008		0.003	
Year 2000	-0.030	***	-0.028	***	-0.003		0.033	***
Year 2001	-0.059	***	0.058	***	-0.005		0.040	***
Agriculture and forrestry	- a)		- a)		- a)		- a)	
Mining and energy	-0.042	***	0.041	***	0.041	***	0.320	***
Ressource processing	0.018	***	0.027	***	-0.066	***	0.070	***
Investments goods	-0.044	***	0.315	***	0.019	***	-0.530	***
Consumption goods	0.025	***	-0.065	***	- a)		- a)	
Construction	- a)		- a)		- a)		- a)	
Retail	0.034	***	-0.359	***	-0.029	***	0.208	***
Logistics and Communications	-0.009	***	-0.130	***	0.083	***	0.165	***
Credit and banking	-0.003	***	0.159	***	0.002		-0.080	***
Insurance	0.063	***	-0.046	***	0.017	***	-0.085	***
Restauration and hotel	0.014	***	-0.169	***	-0.064	***	0.122	***
Education and publishing	0.065	***	-0.290	***	-0.029	***	0.136	***
Health sector	-0.027	***	-0.050	***	-0.017	***	0.293	***
Liberal professions	-0.024	***	-0.040	***	-0.015	***	0.101	***
Other services	0.002	**	-0.090	***	- a)		- a)	

Note: Significance at the *** 1-% level, ** 5-% level, * 10-% level.

a) Correlation not available because in the largest group there is no firm in the sector with > 30 movers and persons observed > 2 periods.

Table 4: Correlation of unobserved effects in dataset version one (holding age, sex, nationality constant)

	IW	IT	FW	FT
IW	1			
IT	0.34	1		
FW	-0.30	0.01	1	
FT	0.02	-0.26	-0.02	1

All correlations significant at the 1-% level, N=650,155.
Based on effects of group 1, only firm effects with > 30 movers,
only person effects with > 2 observations.
IW: individual wage effect, IT: individual tenure effect
FW: firm wage effect, FT: firm tenure effect

Table 5: Partial correlation of unobserved effects in dataset version two (holding age, sex, nationality constant)

	IW	IT	FW	FT
IW	1			
IT	0.26	1		
FW	-0.12	-0.03	1	
FT	0.11	-0.28	-0.13	1

All correlations significant at the 1-% level, N=10,353.
Based on effects of group 1, only firm effects with > 30 movers,
only person effects with > 2 observations.
IW: individual wage effect, IT: individual tenure effect
FW: firm wage effect, FT: firm tenure effect

Appendix

Table 6: Estimation of wage and tenure equations (non-preferred samples)

Dependent Variable:	Log Wage		Log Tenure	
Sample	Wage 2		Tenure 1	
Model:	Pooled	P+F FE	Pooled	P+F FE
	-	-	0.010	0.005
	-	-	253.6	120.1
Tenure	0.079	0.157	-	-
	149.5	58.0	-	-
Tenure squared/100	-0.475	-1.247	-	-
	-115.0	-42.7	-	-
Ten. cens. (dummy)	0.053	0.027	-	-
	23.9	2.5	-	-
Age	0.083	-	0.684	-
	44.5	-	197.5	-
$Age^2/100$	-0.171	-0.685	-1.501	-1.423
	-36.7	-15.8	-168.5	-82.9
$Age^3/1000$	0.116	0.503	1.068	1.002
	31.1	14.4	144.9	71.1
Male	0.229	-	-0.006	-
	152.5	-	-2.5	-
EU foreigner	0.025	-	0.046	-
	7.9	-	8.9	-
Other foreigner	0.012	-	0.033	-
	5.4	-	9.2	-
Voc. Training	0.069	0.011	0.172	0.115
	44.1	0.5	63.4	9.7
Voc. Training and A-levels	0.118	0.031	-0.078	0.144
	35.5	0.7	-14.0	6.2
University	0.240	0.105	-0.243	0.286
	88.4	2.4	-53.4	12.8
Skilled blue-collar	0.040	-0.030	0.079	0.073
	22.5	-1.4	25.7	9.2
White collar	0.163	0.044	0.029	-0.017
	57.1	1.4	5.6	-1.5
Log firm size	0.018	-0.039	0.013	0.193
	30.8	-2.2	13.1	30.9
Firm pays > tarif wage	0.024	-0.046	-0.168	-0.013
	14.4	-4.5	-61.1	-4.2
Business expectations	-0.009	0.086	0.034	-0.108
	-1.3	2.6	4.1	-16.8
IT investments (dummy)	-0.030	-0.021	0.0125	-0.041
	-11.8	-2.3	4.1	-15.4
Investments (in 10 mill. Euros)	0.002	0.002	0.000	-0.003
	20.4	4.2	2.6	-18.6
Firm uses part-time work	-0.047	-0.059	0.056	0.012
	-10.5	-3.4	10.8	2.5
Firm uses fixed-term work	0.015	-0.012	0.008	0.001
	6.4	-1.2	2.8	0.3
Sector-level coll. contract	0.064	0.025	0.089	0.032
	22.7	1.2	18.0	5.1
Firm-level coll. contract	0.071	0.019	-0.092	-0.025
	21.1	0.8	-15.9	-3.4
Constant	2.399	-	-10.010	-
	98.3	-	-228.3	-

Note: All specifications include year, sector and profession dummies.
Reference categories are: Female, German nationality, No vocational training, unskilled blue-collar, no IT investment, no part-time, no fixed-term, no coll. contract.
T-values indicated below coefficients, "P+F FE": Person and firm fixed effects.

Table 7: Summary statistics

Sample:	Wage 1		Wage 2		Tenure 1		Tenure 2	
Observations:	2,540,084		762,941		1,612,944		312,320	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Wage	96.653	30.72	94.921	37.34	92.101	32.43	80.033	42.01
Log wage	4.510	0.41	4.447	0.58	4.449	0.46	4.220	0.73
Tenure	1.323	1.28	5.936	4.67	0.788	1.33	2.609	2.86
Tenure squared/100	5.984	3.95	0.571	0.64	3.830	3.11	0.150	0.27
Log tenure	0.514	0.50	1.141	1.48	0.244	0.32	0.117	1.55
Ten. cens. (dummy)	0.365	0.48	0.277	0.45	0.000	0.00	0.000	0.00
Age	40.207	10.04	40.677	10.98	37.644	9.98	37.879	12.11
$Age^2/100$	17.173	8.31	17.752	9.18	15.167	7.99	15.815	9.93
$Age^3/1000$	7.732	5.47	8.219	6.14	6.499	5.12	7.181	6.55
Male	0.778	0.42	0.746	0.44	0.749	0.43	0.680	0.47
EU foreigner	0.035	0.18	0.034	0.18	0.036	0.19	0.037	0.19
Other foreigner	0.075	0.26	0.078	0.27	0.083	0.28	0.096	0.29
Voc. Training	0.643	0.48	0.618	0.49	0.608	0.49	0.551	0.50
Voc. Training and A-levels	0.034	0.18	0.039	0.19	0.041	0.20	0.044	0.20
University	0.105	0.31	0.103	0.30	0.123	0.33	0.102	0.30
Skilled blue-collar	0.249	0.43	0.238	0.43	0.227	0.42	0.196	0.40
White collar	0.412	0.49	0.409	0.49	0.410	0.49	0.400	0.49
Skilled manual	0.176	0.38	0.161	0.37	0.163	0.37	0.138	0.34
Technical	0.126	0.33	0.110	0.31	0.113	0.32	0.086	0.28
Unskilled services	0.097	0.30	0.097	0.30	0.101	0.30	0.106	0.31
Skilled services	0.023	0.15	0.024	0.15	0.024	0.15	0.027	0.16
Semi-Professional	0.038	0.19	0.044	0.20	0.045	0.21	0.054	0.23
Professional	0.015	0.12	0.017	0.13	0.019	0.14	0.022	0.15
Unskilled administrative	0.044	0.20	0.047	0.21	0.047	0.21	0.059	0.24
Skilled administrative	0.164	0.37	0.166	0.37	0.157	0.36	0.150	0.36
Manager	0.018	0.13	0.017	0.13	0.020	0.14	0.018	0.13
Log firm size	7.213	1.45	7.144	1.50	7.047	1.49	6.839	1.52
Firm pays > tarif wage	0.675	0.47	0.637	0.48	0.651	0.48	0.609	0.49
Business expectations	0.017	0.11	0.014	0.08	0.017	0.12	0.011	0.10
IT investments (dummy)	0.881	0.32	0.939	0.24	0.875	0.33	0.898	0.30
Investments (in 10 mill. Euros)	4.555	9.47	4.355	8.52	4.285	9.44	3.117	7.28
Firm uses part-time work	0.965	0.18	0.982	0.13	0.961	0.19	0.969	0.17
Firm uses fixed-term work	0.862	0.34	0.917	0.28	0.856	0.35	0.875	0.33
Sector-level coll. contract	0.870	0.34	0.860	0.35	0.854	0.35	0.849	0.36
Firm-level coll. contract	0.095	0.29	0.095	0.29	0.103	0.30	0.102	0.30
Year 1996	0.126	0.33	0.056	0.23	0.100	0.30	0.092	0.29
Year 1997	0.118	0.32	0.056	0.23	0.103	0.30	0.100	0.30
Year 1998	0.129	0.33	0.076	0.27	0.125	0.33	0.133	0.34
Year 1999	0.164	0.37	0.093	0.29	0.166	0.37	0.173	0.38
Year 2000	0.161	0.37	0.084	0.28	0.170	0.38	0.167	0.37
Year 2001	0.176	0.38	0.091	0.29	0.194	0.40	0.180	0.38
Agriculture and forrestry	0.001	0.04	0.002	0.04	0.002	0.04	0.003	0.05
Mining and energy	0.022	0.15	0.030	0.17	0.022	0.15	0.036	0.19
Ressource processing	0.185	0.39	0.168	0.37	0.184	0.39	0.162	0.37
Investments goods	0.420	0.49	0.394	0.49	0.387	0.49	0.333	0.47
Consumption goods	0.053	0.22	0.057	0.23	0.058	0.23	0.070	0.25
Construction	0.014	0.12	0.016	0.13	0.018	0.13	0.026	0.16
Retail	0.039	0.19	0.039	0.19	0.044	0.21	0.052	0.22
Logistics and Communications	0.059	0.24	0.057	0.23	0.057	0.23	0.050	0.22
Credit and banking	0.035	0.18	0.041	0.20	0.030	0.17	0.028	0.16
Insurance	0.010	0.10	0.012	0.11	0.009	0.10	0.013	0.11
Restauration and hotel	0.014	0.12	0.015	0.12	0.020	0.14	0.022	0.15
Education and publishing	0.023	0.15	0.025	0.16	0.026	0.16	0.034	0.18
Health sector	0.057	0.23	0.068	0.25	0.069	0.25	0.084	0.28
Liberal professions	0.012	0.11	0.016	0.12	0.016	0.13	0.020	0.14
Other services	0.004	0.06	0.004	0.07	0.005	0.07	0.009	0.10

References

Abowd J. and Kang C. (2002): "Simultaneous Determination of Wage Rates and Tenure", mimeo.

Abowd J., Kramarz F. and Margolis D. (1999): "High Wage Workers and High Wage Firms", Econometrica, 67(2), pp. 251-334.

Abowd J., Creecy R. and Kramarz F. (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data", Technical Paper No. TP-2002-06, U.S. Census Bureau.

Abowd J., Kramarz F., Lengerman P. and Pérez-Duarte S. (2004): "Are good workers employed by good firms?", mimeo.

Abowd J., Kramarz F., Roux S. (2006): "Wages, mobility and firm performance: advantages and insights from using matched worker-firm data", The Economic Journal, 116, pp. F245-F285.

Alda, H. (2006): "Beobachtbare und unbeobachtbare Betriebs- und Personeneffekte auf die Entlohnung", Beiträge zur Arbeitsmarkt- und Berufsforschung Nr. 298, IAB, Nürnberg.

Alda, H., Bender S., Gartner H. (2005): "The linked employer-employee dataset of the IAB (LIAB)", IAB Discussion Paper No. 6/2005.

Anderson, P.M. and Meyer B.D. (1994): "The Extent and Consequences of Job Turnover", Brookings Papers on Economic Activity: Microeconomics, 177-248.

Andrews M., Schank T., Upward R. (2006a): "Practical fixed effects estimation methods for the three-way error components model", erscheint demnächst in: The Stata Journal

Andrews M., Gill L., Schank T., Upward R. (2006b): "High wage workers and low wage firms: negative assortative matching or statistical artefact?", Discussion Paper No. 42, Lehrstuhl für Arbeitsmarkt- und Regionalpolitik, Friedrich-Alexander-Universität Erlangen-Nürnberg.

Barth E. and Dale-Olsen H. (2003): "Assortative matching in the labour market? Stylised facts about workers and plants" mimeo.

Becker, G. (1973): "A Theory of Marriage: Part I", Journal of Political Economy, 81(4):813-846.

Boockmann, B. and Steffes S. (2005): "Individual and Plant-level Determinants of Job Durations in Germany", ZEW Discussion Paper No. 05-89, Mannheim.

Cornelißen, T. (2006): "A memory saving fixed effects estimation of the two-way error component model in large data sets", mimeo.

Frederiksen, A. (2004): "Explaining Individual Job Separations in a Segregated Labour Market" Industrial Relations Section Working Paper No. 490, Princeton University.

Gerlach K. and Stephan G. (2004): "Betriebszugehörigkeit, Betriebseffekte und industrielle Beziehungen", in: Struck O. and Köhler C. (Eds.) 2005: "Beschäftigungsstabilität im Wandel? Empirische und theoretische Befunde" Rainer Hampp Verlag, München.

Gerlach K. and Stephan G. (2004): "Individual tenure and collective contracts", IAB Discussion Paper No. 10/2005.

Goux, D. and Maurin E. (1999): "'Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data"', Journal of Labor Economics, 17(3), pp. 492-533.

Greene W. (2003) "Econometric Analysis", Prentice Hall, New Jersey.

Grotheer M., Struck O., Bellmann L. and Gewiese T. (2004): "Determinanten von Beschäftigungsstabilität. Chancen und Risiken von 'Entrants' im ost-westdeutschen Vergleich" in: Struck O. and Köhler C. (Eds.) 2005: "Beschäftigungsstabilität im Wandel? Empirische und theoretische Befunde" Rainer Hampp Verlag, München.

Grütter M. and Lalive R. (2004) "The Importance of Firms in Wage Determination", IEW Working Paper No. 207, Institute for Empirical Research in Economics, University of Zurich.

Grütter M. (2006) "The anatomy of the wage structure", Verlag Dr. Kovac, Hamburg.

Honoré B.E. (1992): "Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects", Econometrica 60, 533-565.

Honoré B.E. (1993): "Orthogonality conditions for tobit models with fixed effects and lagged dependent variables", Journal of Econometrics 59, 35-61.

Lancaster, T. (1990): "The Econometric Analysis of Transition Data", Cambridge University Press, Cambridge.

Mumford K. and Smith P.N. (2004): Job Tenure in Britain: Employee Characteristics versus workplace effects", Economica 71, 275-298.

Neyman J. and Scott E. L. (1948): "Consistent estimates based on partially consistent observations", Econometrica, 16(1), 1–32.

Postel-Vinay F. and Robin J.-M. (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", Econometrica, 70(6), 2295-2350.

Ritchie, F. (1995): "Efficient Access to Large Datasets for Linear Regression

Models", Working Paper 95/12, Department of Economics, University of Stirling.

Salant S.W. (1977): "Search theory and duration data: a theory of sorts", Quarterly Journal of Economics, 91(1), 39-57.

Shimer R. (2005): "The Assignment of Workers to Jobs in an Economy with Coordination Frictions", Journal of Political Economy, 113(5), 996-1025.

Shimer R. and Smith L. (2000): "Assortative Matching and Search", Econometrica, 68(2), 343-370.