

DISCUSSION PAPER SERIES

IZA DP No. 10447

**A Simple Method to Estimate
Large Fixed Effects Models Applied to
Wage Determinants and Matching**

Nikolas Mittag

DECEMBER 2016

DISCUSSION PAPER SERIES

IZA DP No. 10447

A Simple Method to Estimate Large Fixed Effects Models Applied to Wage Determinants and Matching

Nikolas Mittag
CERGE-EI and IZA

DECEMBER 2016

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

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

ABSTRACT

A Simple Method to Estimate Large Fixed Effects Models Applied to Wage Determinants and Matching*

Models with high dimensional sets of fixed effects are frequently used to examine, among others, linked employer-employee data, student outcomes and migration. Estimating these models is computationally difficult, so simplifying assumptions that are likely to cause bias are often invoked to make computation feasible and specification tests are rarely conducted. I present a simple method to estimate large two-way fixed effects (TWFE) and worker-firm match effect models without additional assumptions. It computes the exact OLS solution including estimates of the fixed effects and makes testing feasible even with multi-way clustered errors. An application using German linked employer-employee data illustrates the advantages: The data reject the assumptions of simpler estimators and omitting match effects biases estimates including the returns to experience and the gender wage gap. Specification tests detect both problems. The results suggest that firm fixed effects, not match effects, are the main channel through which job transitions drive wage dynamics, which underlines the importance of firm heterogeneity for labor market dynamics.

JEL Classification: J31, J63, C23, C63

Keywords: multi-way fixed effects, linked employer-employee data, matching, wage dynamics

Corresponding author:

Nikolas Mittag

CERGE-EI

Politických vězňů 7

Prague 1, 111 21

Czech Republic

E-mail: nikolas.mittag@cerge-ei.cz

* I would like to thank Dan Black, Randy Filer, Jeffrey Grogger, Stepan Jurajda, Francis Kramarz, Bruce Meyer and participants at the IAB, RWI, EALE and the Czech Economic Society for comments and suggestions as well as Stefan Bender and the staff at the IAB for the excellent cooperation with the data. All errors are mine.

1. Introduction

This paper introduces a fast and simple method to estimate the two-way fixed effects model (TWFE) and the match effects model when the number of fixed effects makes standard estimation techniques infeasible. Following the seminal article by Abowd, Kramarz and Margolis (1999), models with two large sets of fixed effects are frequently applied to linked employer-employee data to examine a wide range of topics in labor economics (see Abowd, Kramarz and Woodcock 2008 for an overview) as well as in other fields where agents from two heterogeneous populations interact. Recent examples of such applications include student-teacher data (e.g. Jacob and Lefgren 2008, Kramarz, Machin and Ouazad 2008) and doctor-patient data (e.g. Bennett, Hung and Lauderdale 2010). The match effects model extends the TWFE by including the interaction of the two fixed effects, i.e. a match or job fixed effect. Theories of matching between employers and firms (e.g. Jovanovic 1979, Mortensen 1978) usually allow for complementarities between workers and firms, but empirical specifications with only individual and firm fixed effects do not allow productivity to depend on how workers and firms are matched. Looking beyond the labor market, match effects are of interest whenever the way in which agents are paired with each other matters, which is likely in many interactions, for example between doctors and patients, students and teachers and the obvious case of the marriage market.

However, the computational complexity of both models makes specification and estimation difficult. Even in the simpler TWFE the number of fixed effects often exceeds one million, which makes it impossible to use standard estimation methods. Instead, researchers often estimate models that impose additional restrictions to reduce computation, such as random or mixed effects models. These restrictions are usually neither justified by theory nor tested or supported by the data, so they are likely to cause bias. Similarly, only few applications consider or test for match fixed effects, even though their omission causes bias if the pairing of agents matters. Consequently, it is important to be able to estimate fixed effect specifications for both the TWFE and the match effects model to avoid such

simplifications or at least test their validity. Finally, other estimation strategies often compute the exact OLS slopes, but do not calculate important parameters such as the fixed effects or the variance matrix, which limits the usefulness of the results as well as tests of their significance and model specification

The method I propose in this paper solves or greatly simplifies the computational problem. It yields the exact OLS solutions including estimates of the fixed effects and the covariance matrix of the slope coefficients for both the TWFE and the match effects model.¹ Besides the convenience of faster computation and lower memory requirements, a first advantage is that being able to calculate the exact OLS slopes even in large applications avoids the bias from assumptions that are only made to make computation feasible. Next, estimates of the fixed effects are often of substantive interest in their own right. The method introduced here simplifies the decomposition into a firm-, individual- and match component. In addition, obtaining the correct variance matrix of the slopes is crucial for tests of significance and model specification, yet most previous solutions only estimate the coefficients without standard errors. They are usually bootstrapped, which amplifies the computational problem. Even worse, to my knowledge, there is no practical way to consistently bootstrap the covariance matrix if errors are not independent and identically distributed. Yet in most applications it is likely that there is two-way clustering, for example because errors are correlated within both individuals and firms. The method in this paper can be used to obtain the asymptotic covariance matrix in the presence of clustering, which solves these problems and simplifies specification tests. Finally, the method reduces the computational burden of estimating both the TWFE and the match effects model, which makes tests for the presence of match effects based on the restricted and unrestricted model feasible. The omission of match effects can bias estimates and there is ample evidence that they matter in common applications (e.g. Garen 1989, Jackson 2013, Woodcock 2008, Woodcock 2015). Since tests for their presence are well known, a

¹ Implementations in Matlab and Stata are available from my website and ssc.

likely reason why they are rarely conducted is the computational cost of estimating two complex models, which the method in this paper makes negligible for models of common sizes.

Using the method to analyze wage determinants and matching with linked employer-employee data from Germany illustrates these advantages. The first part of the analysis confirms that the restrictions of computationally simpler estimators do not hold. In addition, match effects matter and their omission leads to substantial downward bias in the returns to experience and education, and overstates the gender wage gap by almost 100 percent. Standard specification tests detect these problems and are simple to conduct based on the method I present. The second part of the analysis illustrates that using the method presented in this paper to decompose wages into individual- firm- and match effects can yield interesting insights. I show that conditions at the time a match is formed such as previous labor market status and age of the individual are systematically related to both the subsequent firm and match effects, i.e. they predict whether an individual obtains a job with a higher or lower wage than a randomly assigned job. Estimating this relationship not only helps to explain the bias from omitting the match effects, it also provides suggestive evidence on the mechanisms that make job transitions important for subsequent wages. For example, the impact of the labor force status before a new match is larger than the effect of its duration, which favors explanations for the long term negative effects of non-employment between jobs on wages based on models of search and signaling over those based on human capital. The analysis also reveals large gender differences in how matching changes over the life-cycle: both firm and match effects initially increase for both genders, but eventually decrease in age for men while they continue to increase for women. The differences are mainly driven by the firm fixed effects, as opposed to match effects, which suggests that differential selection into the labor market rather than statistical discrimination is the primary cause of these gender differences. This underlines that estimates of the match effects can be of interest even if one does not want to interpret them as parameters of an underlying structural model.

The next part of this paper describes the two models, estimation problems and common methods. Part 3 introduces a new estimation strategy and compares it to previous approaches. Part 4 applies the method to analyze wage determinants and matching using German linked employer-employee data.

2. The Two-Way Fixed Effects and the Match Effects Model

A common specification in panel data models is the unbalanced two-way fixed effects model which includes a set of fixed effects for primary units indexed by $i=1,\dots,N$ and secondary units indexed by $j=1,\dots,J$. Thus, there are N primary units and J secondary units. Applications of this model include, among others, matched employer-employee data (in which the units are individuals and firms, i.e. a fixed effect for each individual and each firm is included; see e.g. Abowd and Kramarz 1999) and student-school data (including fixed effects for pupils and schools; see e.g. Kramarz, Machin and Ouazad 2008). The model is defined by:

$$y_{ijt} = x_{ijt}\beta + \theta_i + \psi_j + \varepsilon_{ijt} \quad (1)$$

Where y_{ijt} is unit i 's (scalar) outcome at time t , x_{ijt} is a $1 \times K$ vector of time-varying observed covariates, β is a vector of coefficients, θ_i and ψ_j are time-invariant scalar fixed effects and ε_{ijt} is the error term. Let T_i indicate the number of observations on primary unit i , i.e. the subscript t runs from 1 to T_i for unit i . Similarly, let F_j stand for the number of observations on secondary unit j , so that the total number of observations is $N^* = \sum_{i=1}^N T_i = \sum_{j=1}^J F_j$. The model allows both T_i and F_j to vary between units, so the panel does not have to be balanced. In matrix notation, the model can be expressed as

$$y = X \beta + D_\theta \theta + D_\psi \Psi + \varepsilon \quad (2)$$

$\begin{matrix} N^* \times 1 & N^* \times K & K \times 1 & N^* \times N & N \times 1 & N^* \times J & J \times 1 & N^* \times 1 \end{matrix}$

Where y is an $N^* \times 1$ vector of outcomes, X is an $N^* \times K$ matrix of observable time-varying covariates, D_θ is the $N^* \times N$ matrix of indicators for the primary unit, D_ψ is the $N^* \times J$ matrix of indicators for the secondary unit and ε is the $N^* \times 1$ vector of error terms. The parameters of the model are β , the $K \times 1$ vector of slopes, θ , the $N \times 1$ vector of fixed effects for the primary units and Ψ , the $J \times 1$ vector of fixed

effects for the secondary units. The model is usually estimated by OLS invoking the standard conditional mean independence assumption:

$$E(\varepsilon|X, D_\theta, D_\Psi) = 0 \quad (3)$$

To ease notation, define

$$\begin{aligned} T &= D'_\theta D_\theta = \text{diag}(T_1, \dots, T_N) \\ &_{N \times N} \quad N \times N^* \quad N^* \times N \\ F &= D'_\Psi D_\Psi = \text{diag}(J_1, \dots, J_J) \\ &_{J \times J} \quad J \times N^* \quad N^* \times J \\ K &= D'_\theta D_\Psi \\ &_{N \times J} \quad N \times N^* \quad N^* \times J \end{aligned} \quad (4)$$

So that T is an $N \times N$ diagonal matrix with the number of observations on primary unit i as the i^{th} diagonal element and F is a $J \times J$ diagonal matrix with the number of observations on secondary unit j as the j^{th} diagonal element. Element (i, j) of the $N \times J$ matrix K indicates how many observations on primary unit i belong to secondary unit j , e.g. how many periods individual i worked for firm j .

The match effects model (see e.g. Woodcock 2008, 2015) is an extension of this model which includes an interaction between the two fixed effects:

$$y_{ijt} = x_{ijt} \beta + \theta_i + \psi_j + \lambda_s + \varepsilon_{ijt} \quad (5)$$

Where the index $s=1, \dots, S$ is for notational convenience only as it is determined by i and j : $s = f(i, j)$. The match effects model adds an effect to the TWFE within which both the firm and individual fixed effect are nested. As discussed in Woodcock (2015), the mean of the match effects within each i and j is not identified and has to be normalized. In matrix notation, the model can be expressed as

$$y = X \beta + D_\theta \theta + D_\Psi \Psi + D_\lambda \lambda + \varepsilon \quad (6)$$

$N^* \times 1$ $N^* \times K$ $K \times 1$ $N^* \times N$ $N \times 1$ $N^* \times J$ $J \times 1$ $N^* \times S$ $S \times 1$ $N^* \times 1$

D_λ is the $N^* \times S$ matrix of indicators for matches between the two units and λ is the $S \times 1$ vector of match fixed effects. The mean independence assumption of the error term in this case becomes

$$E(\varepsilon|X, D_\theta, D_\Psi, D_\lambda) = 0 \quad (7)$$

This paper deals with the case in which both sets of fixed effects include a large number of units, making estimation by standard techniques infeasible. In the remainder of this paper, I discuss these models in terms of matched employer-employee data, i.e. I refer to y as (log) wages, θ and Ψ as individual- and firm-fixed effects and λ as match-fixed effects. I define the secondary unit as the smaller unit in the sense that $J < N$ without loss of generality (since labeling of the units is arbitrary). I assume that there are fewer firms than individuals for ease of exposure, but the estimation method extends to other applications with the variables defined analogously.

Abowd, Creedy, and Kramarz (2002) discuss identification in the TWFE. They show that all worker and firm effects within each connected group (groups of firms with realized mobility) are identified up to one normalization: The level of one set of fixed effects is only identified relative to the other set of fixed effects in each group. The estimation strategy below excludes an overall intercept and constrains the individual fixed effects within each group to sum to zero. Other normalizations are easy to implement. The match effects model additionally includes the interactions between firm and individual fixed effects. The mean match effect for each individual and firm is not identified. Intuitively, the average match quality is an invariant characteristic of a firm or individual by construction, so it cannot be separately identified from the person and firm effect. Therefore, match effects are usually constrained to sum to zero for each individual and each firm; see Woodcock (2015) for a discussion. This normalization has convenient computational properties and other normalizations can easily be implemented after estimation. While this makes the match effects orthogonal to θ and Ψ by construction, they are not necessarily orthogonal to X . Thus, omitting match effects in a model without regressors does not cause any bias, but if the model includes X omitting them is likely to bias $\hat{\beta}$ as well as $\hat{\theta}$ and $\hat{\Psi}$.

Estimation Problems and Common Solutions

Standard OLS estimation of the models above is computationally infeasible, because it requires the inverse of a $(K + N + J) \times (K + N + J)$ matrix for the two-way fixed effects model $((K + N + J +$

$S) \times (K + N + J + S)$ for the match effects model). In typical datasets, this number easily exceeds one million. Calculations with matrices of this size require tremendous amounts of main memory² and computational power³. Each of the two problems by itself makes standard estimation methods infeasible. I provide a brief overview of alternative estimation methods below. For more detailed discussions of these strategies and other methods (such as random or mixed effects) see Abowd, Kramarz and Woodcock (2008) and Andrews, Schank and Upward (2006).

In some applications (e.g. Bennett, Hung and Lauderdale 2010, Markussen and Røed 2015) only the slope coefficients β are of interest while the fixed effects are nuisance parameters. One has to control for the fixed effects, because they are correlated with the covariates, but their estimates are not of interest. For the match effects model, subtraction of match specific means sweeps out all fixed effects and the slopes can be estimated by OLS on the transformed data. This can easily be done using standard software so the method in this paper yields only minor advantages. It is more common, however, that the researcher is interested in the slopes of the TWFE model only. The TWFE model is a restricted version of the match effects model (it restricts all match effects to equal zero), so the strategy above still yields consistent estimates. However, they are inefficient estimates of the TWFE slopes since the model contains millions of superfluous parameters. Additionally, regressors that do not vary within a match cannot be included, because they are purged when subtracting match specific means. Consequently, it is only a viable strategy to estimate the slopes if efficiency is not a concern, there are no match specific regressors and one has good reasons to believe that all match effects are indeed zero. The algorithm I propose yields efficient estimates of the TWFE slopes that are identical to the OLS estimates even if there are match specific regressors.

² E.g., a matrix with one million rows and columns stored in double precisions requires 8 TB of working memory.

³ To my knowledge, the matrix inversion algorithms in conventional software require computation of cubic order. Computation would still be infeasible if the current lower bound in Le Gall (2014) were practicable.

In other cases, at least one set of the fixed effects is of interest. Several parameters of interest can be calculated from the estimates of the individual and firm fixed effects (e.g. Abowd, Kramarz, Roux 2006) and teacher fixed effects are commonly used as measures of value added (e.g. Rockoff 2004). Estimates of the match effects are of interest in search models (Woodcock 2015), matching of teachers and schools (Jackson 2013) as well as international migration (Grogger and Hanson 2011). To estimate the TWFE in these applications, the researcher needs to estimate the full model given by equation 2. Woodcock (2015) shows that the match effect model with the normalization that is also used in this paper can be reduced to solving a problem of size $N+J$ by first calculating match means and then decomposing them into the three sets of fixed effects. So both models can be estimated by solving a problem of similar size, but this problem is too large for conventional methods.

A common approach to make the problem computationally tractable is to make additional orthogonality assumptions. Abowd, Kramarz and Margolis (1999) propose estimators for the TWFE that provide unbiased estimates if one set of fixed effects is orthogonal to the other set of fixed effects and the observables X , conditional on other variables Z . Others calculate the fixed effects from the residuals of a regression (e.g. Chetty, Friedman and Rockoff 2011⁴), which requires them to be uncorrelated with the other set of fixed effects and X . These assumptions are similar to those of random- or mixed-effects models, which are easy to compute. For example, Woodcock (2015) proposes a mixed-effect estimator that allows correlation with X , but requires the sets of fixed effects to be orthogonal. Such assumptions are made for computational rather than substantial reasons and are usually rejected when tested, so their advantages are likely to be outweighed by the bias they introduce.

Approaches to compute the exact OLS slopes reduce computational complexity by solving the normal equation without inverting the cross-product matrix. For example, several algorithms such as the conjugate gradient algorithm (CGA, see Abowd, Creecy and Kramarz 2002, Ouazad 2007) or alternating

⁴ The final version (Chetty, Friedman and Rockoff 2014) uses a different model and does not face this problem.

projections (Guimarães and Portugal 2010, Gaure 2013, Correia 2014) that iteratively search for a solution to the normal equations have been proposed. Iterative solutions cannot be used to calculate standard errors, which can at best be obtained by computationally intensive procedures such as the bootstrap. This does not work with multi-way clustering and exacerbates the problem that iterative methods are very slow. Another approach reduces the dimensionality of the matrix to be inverted by performing a first-difference (e.g. Abowd, Kramarz and Margolis 1999) or within-transformation (e.g. Andrews, Schank and Upward 2006, Cornelissen 2008) on the larger set of fixed effects. This reduces the size of the matrix to $K + \min(N, J)$, but both approaches still have to invert a very large matrix.

3. A Computationally Simple Estimation Strategy

The estimation strategy I present below is related to the three approaches in the previous section in that it uses a transformation to obtain the slopes, reduces the dimensionality of the matrix and then solves a system of linear equations rather than inverting the matrix. It yields the exact OLS estimates including all fixed effects and the asymptotic covariance matrix without additional assumptions and requires less computation and memory. Contrary to iterative solutions, it yields standard errors of the slopes even if errors are not independent and identically distributed. It does not calculate standard errors of the fixed effects, which can still be bootstrapped. I first discuss the estimation of the match effects model, then show how to apply it to the TWFE model and finally briefly compare it to other strategies.

Estimating the match effects model

Simplifying the estimation for both models rests on three simple properties of OLS:

1. Partial Regression (e.g. Yule 1907)
2. If $\hat{\beta}^{OLS}$ are the OLS coefficients on X from a regression of y on $[X Z]$, then the OLS coefficients on Z from regressing $(y - X\hat{\beta}^{OLS})$ on Z and regressing y on $[X Z]$ are numerically identical.
3. The residuals sum to zero for every firm, individual and match.

Property 1 and 3 are well known, and proofs can be found in most econometric text books. Property (2) is often shown in the context of partial regression (e.g. Greene 2008 p. 27); a simple proof is included in appendix A. For the application in this paper, Z should be thought of as the design matrix of the fixed effects, e.g. $Z = [D_\theta \ D_\psi]$ for the TWFE model.

To estimate the match effects model, one first needs to estimate $\hat{\beta}^{ME}$, the slopes from the match effects model defined by eq. 6. This can be done by an OLS regression of deviations of y_{ijt} and x_{ijt} from their match means. This is equivalent to running a partial regression, i.e. regressing the residuals from a regression of y on the three sets of fixed effects on the residuals from regressions of each column of X on the three sets of fixed effects. Because individual and firm fixed effects are constant within each match, these residuals are easily computed as the deviations of y_{ijt} and x_{ijt} from their match means (see e.g. Woodcock 2015 for proof). The estimates, residuals and standard errors (corrected for the degrees of freedom) from these partial regressions are numerically identical to those that would be obtained from OLS estimation of the full model (Yule 1907).

$\hat{\beta}^{ME}$ from the previous step is the OLS estimate, so estimates of the fixed effects can be obtained by a regression of $\tilde{y} = y - X\hat{\beta}^{ME}$ on the two sets of fixed effects by property (2). The normalization that match effects sum to zero within each firm and individual allows one to omit the match effects in this regression because the model does not contain X . In a regression that includes the three sets of fixed effects only, the effect of omitting the match effects is given by the usual formula for omitted variable bias: $[[D_\theta \ D_\psi]'[D_\theta \ D_\psi]]^{-1}[D_\theta \ D_\psi]'D_\lambda\lambda$. This bias is always exactly 0 by construction, because the normalization of having match effects sum to zero within firm and individual implies that $D_\theta'D_\lambda\lambda = 0$ and $D_\psi'D_\lambda\lambda = 0$ hold by construction within sample, so one can compute the match effects separately. Note that this requires the first step of purging the effect of X , because the match effects need not be exactly orthogonal to the firm and individual effects after conditioning on the covariates.

Estimates from a regression of \tilde{y} on the two sets of fixed effects are then given by the standard formula:

$$\begin{pmatrix} \hat{\theta}^{ME} \\ \hat{\Psi}^{ME} \end{pmatrix} = \begin{pmatrix} T & K \\ K' & F \end{pmatrix}^{-1} \begin{pmatrix} D'_\theta \tilde{y} \\ D'_\Psi \tilde{y} \end{pmatrix} = \begin{pmatrix} T & K \\ K' & F \end{pmatrix}^{-1} \begin{pmatrix} T\bar{y}_i \\ F\bar{y}_j \end{pmatrix} \quad (8)$$

where \bar{y}_i and \bar{y}_j are vectors of individual and firm means of \tilde{y}_{it} and T, F and K are defined by eq. 4. To calculate the firm fixed effects (or, more generally, the smaller set of fixed effects), one only needs the lower blocks of the inverse of the partitioned matrix. These blocks can be obtained by applying the formula for the inverse of a partitioned matrix (see Theil 1971 section 1.2):

$$\begin{aligned} \hat{\Psi}^{ME} &= \left[-(F - K'T^{-1}K)^{-1}K'T^{-1} \quad (F - K'T^{-1}K)^{-1} \right] \begin{pmatrix} T\bar{y}_i \\ F\bar{y}_j \end{pmatrix} \\ &= (F - K'T^{-1}K)^{-1}F\bar{y}_j - (F - K'T^{-1}K)^{-1}K\bar{y}_i \\ &= (F - K'T^{-1}K)^{-1}(F\bar{y}_j - K\bar{y}_i) \end{aligned} \quad (9)$$

T and F are diagonal matrices and K has at most S non-zero elements, which makes them sparse. All three matrices contain only integers. Consequently, the matrices and vectors in this expression are simple to obtain and can be stored efficiently. The only remaining computational difficulty is inverting the matrix in the first brackets. However, it is of size $J \times J$, so it is already much smaller than the $(K + N + J + S) \times (K + N + J + S)$ matrix that needs to be inverted for the standard OLS formula. It is symmetric positive definite, so its Cholesky factorization can be used rather than calculating the inverse matrix. One can also use iterative algorithms such as the CGA to solve the much smaller $J \times J$ system of equations implied by equation 9. In most applications, memory is the binding constraint. This favors the CGA approach with its much lower memory requirement. However, if sufficient memory is available, the Cholesky factorization is likely to be faster. In case there are multiple connected groups, $(F - K'T^{-1}K)$ is block diagonal with one block for every connected group. Consequently, the problem can be solved group by group. This makes the problem considerably easier, because the computational complexity increases by the square of the size of the matrix. However, it is not possible without subtracting $X\hat{\beta}$ to get rid of X . Including X adds K rows and columns to the matrix, so it is no longer block-diagonal.

Residuals sum to zero for each individual, firm and match, so the individual fixed effects can be recovered from the estimated firm fixed effects and the individual means. The individual mean contains the average firm effect of the firms for which individual i worked, weighted by match length. The vector of these averages is equal to $T^{-1}K\hat{\Psi}^{ME}$, so that the OLS estimate of the individual fixed effects is:

$$\begin{aligned}
\bar{y}_i &= \bar{X}_i \hat{\beta}^{ME} + \hat{\theta}^{ME} + T^{-1}K\hat{\psi}^{ME} && \Leftrightarrow \\
\hat{\theta}^{ME} &= \bar{y}_i - \bar{X}_i \hat{\beta}^{ME} - T^{-1}K\hat{\psi}^{ME} && \Leftrightarrow \\
\hat{\theta}^{ME} &= \tilde{y}_i - T^{-1}K\hat{\psi}^{ME}
\end{aligned} \tag{10}$$

Finally, the match effects can be computed from the match means:

$$\begin{aligned}
\bar{y}_s &= \bar{X}_s \hat{\beta}^{ME} + \hat{\theta}_i^{ME} + \hat{\psi}_j^{ME} + \hat{\lambda}_s^{ME} && \Leftrightarrow \\
\hat{\lambda}_s^{ME} &= \bar{y}_s - \bar{X}_s \hat{\beta}^{ME} - \hat{\theta}_i^{ME} - \hat{\psi}_j^{ME} && \Leftrightarrow \\
\hat{\lambda}_s^{ME} &= \tilde{y}_s - \hat{\theta}_i^{ME} - \hat{\psi}_j^{ME}
\end{aligned} \tag{11}$$

Where \bar{y}_s and \tilde{y} are $S \times 1$ vectors containing the match means of y_{ijt} and $y_{it} - X_{it}\hat{\beta}^{ME}$ and \bar{X}_s is the $S \times K$ matrix containing the match means of x_{ijt} . Both calculations are computationally trivial.

Estimating the TWFE model

The same algorithm can be applied to the TWFE model with one caveat: In the TWFE model, the OLS predictions based on the fixed effects are different from the match means, so the estimates $\hat{\beta}$ obtained in the first step are not the same as the OLS estimates, even if there really are no match effects. This makes the partial regression step to obtain estimates of β more complicated: One needs to run a partial regression by regressing y and each column of X on the two sets of fixed effects and use the residuals to obtain $\hat{\beta}^{TW}$, the coefficient on X in the TWFE model. Implementing this is greatly simplified by the fact that these regressions have the same covariates (the two sets of fixed effects only) as the regression of \tilde{y} on the two sets of fixed effects that is solved by equation 8. Consequently, one can use the same simplification as above by solving equation 9 and 10 and repeating this for each column of X in place of y .

This is a simple way to implement the transformation proposed by Wansbeek and Kapteyn (1989) and Davis (2002). Rather than inverting the whole $(N+J) \times (N+J)$ matrix, it only requires computation of the Cholesky factorization of $(F - K'T^{-1}K)$. This is the same matrix that is later used to compute the firm fixed effects, so it has the same desirable properties. More importantly, its Cholesky decomposition only needs to be computed once and can be reused in the auxiliary regressions for y and all elements of X as well as to obtain the estimates of the firm fixed effects from \tilde{y} after $\hat{\beta}^{TW}$ has been calculated. This advantage does not apply when using the CGA to solve equation 9, which has to be repeated for each covariate. In most cases, however, solving equation 9 with the CGA is a matter of seconds.

After this step, estimation proceeds exactly as in the match effects model: One obtains estimates of $\hat{\beta}^{TW}$ from a partial regression, subtracts the fitted values from y to obtain \tilde{y} and obtain the firm fixed effects either using the CGA or the Cholesky factorization of $(F - K'T^{-1}K)$ from the partial regressions to solve equation 9. The individual effects are then given by equation 10.

Comparison to Other Methods

Programs that implement the method in Matlab and Stata are available online. This section briefly compares them to other implementations and discusses two key advantages of the approach proposed here. To illustrate the computational advantages of the approach above, Table 1 reports the computation times of different algorithms as the number of individuals and firms increase. Further detail on the simulation setup is provided in Appendix C. Such comparisons should be taken with caution, because computation time depends heavily on the implementation of the algorithm. For comparability, I focus on Stata implementations of the within transformation, the CGA and alternating projections.⁵

The within transformation proposed by Andrews, Schank and Upward (2006) and implemented by Cornelissen (2008), also reduces the size of the matrix that must be decomposed. I additionally exploit

⁵ Implementations for other software include the Fortran version of the CGA, an implementation of the alternating projection algorithm in R and the Matlab version of the methods introduced here. Differences in the underlying software, memory requirements and the practicability make such comparisons even less likely to be meaningful.

the simple structure of the matrix that needs to be inverted to set it up directly and wipe out the K rows for X , which makes the structure of the problem simpler, and solve the smaller problem by the CGA rather than directly. Column 3 of Table 1 shows that this speeds up computation dramatically, particularly as the size of the problem grows. The algorithm does not work for large problems, because the within transformation is also substantially more memory intensive than iterative solutions.

Table 1: Computation Time (in Minutes) of Different Algorithms Using Simulated Data

		(1)	(2)	(3)	(4)	(5)	(6)
<i>Individuals</i>	<i>Firms</i>	<i>Match Effects</i>	<i>TWFE</i>	<i>felsdvreg</i>	<i>a2reg</i>	<i>reg2hdfe</i>	<i>reghdfe</i>
100000	1000	0.2	0.2	0.6	0.5	29.1	0.6
1000000	10000	2	2	40	5	302	7
10000000	100000	19	26	-	51	-	73

Note: Column 1 reports computation time from estimating the match effects model using the algorithm described here, the remaining columns estimate the TWFE. Column 2 uses the method in this paper. Column 3 uses the within transformation described in Cornelissen (2008), column 4 uses the CGA (Abowd, Creedy and Kramarz, 2002) implemented by Ouazad (2007). Columns 5 and 6 use the alternating projections methods by Guimarães and Portugal (2010) and Correia (2014). All times are the average of 5 iterations in minutes. See appendix C for details on the simulation setup. There are 5 covariates and 10 observations per individual. Individuals move to a randomly chosen firm either once (with probability 0.16) or twice (with probability 0.04). All simulations were run single-threaded on a computer with 74 GB of main memory.

Compared to the CGA approach by Abowd, Creedy and Kramarz (2002) implemented in Ouazad (2007), the main advantage of the approach proposed here is that the size of the problem is reduced from $K + N + J$ to J equations and parameters. Even in the most unfavorable case, my approach cuts the size of the problem in half. The results in column 4 show that this cuts computation time by a factor of two to three. However, the CGA scales well, so both approaches can solve large problems, although the approach proposed here required only about half the memory in this application.

Columns 5 and 6 report times from two implementations of the alternating projections approach by Guimarães and Portugal (2010) and Correia (2014). Arcidiacono et al. (2012) and Gaure (2013) also propose solutions based on alternating projections, all of which iteratively solve a problem of size $K + N + J$. An advantage of these solutions is that they are easier to extended to more than two sets of fixed effects and non-linear models. They are based on estimation routines available in all statistical programs, which makes them easy to implement. However, Columns 5 and 6 confirm that this comes at the expense of a much higher computation cost. The implementation by Guimarães and Portugal (2010)

maintains low memory requirements, but becomes very slow as the size of the problem grows. The program by Correia (2014) requires more memory, but scales better in terms of computation time.

In summary, the simulations underline the substantially faster computation and the lower memory requirements of the approach proposed here. In addition to the sizeable differences in computational costs, there are two key advantages of my approach as discussed below: First, it avoids making any assumptions for computational reasons only, such as orthogonal effects or the exclusion of match effects, and hence avoids the bias if these assumptions fail. Second, it estimates important parameters that other approaches do not compute. Most importantly, it yields estimates of the fixed effects, which are often of substantive interest and the asymptotic variance matrix of $\hat{\beta}$.

While the reduction in computational resources needed enables us to estimate larger models, most empirical applications can in principle be estimated by other programs (see e.g. Card, Heining and Kline 2013 for an exception). However, many papers use estimators that make restrictive assumptions with the main benefit of decreasing computational complexity such as random and mixed effects models (e.g. Jackson 2013, Woodcock 2015) or other orthogonality restrictions (e.g. Card, Heining and Kline 2013). In a similar vein, the TWFE is a constrained match effects model that restricts all match effects to equal zero. It is important to test these restrictions, because they cause bias if they do not hold (Woodcock 2008). This can be done by testing whether the slope coefficients change (which requires estimating their variance matrix) or by comparing the restricted TWFE to the unrestricted match effects model (which requires estimating both models). However, such tests are rarely conducted, even though the restrictions these models impose are almost always rejected when they are tested. A likely reason for this is that it is often impossible or prohibitively costly to compute these tests. The estimation procedure here estimates the variance matrix of the slopes and makes it easy to estimate the unrestricted model, thereby making both tests feasible.

The solution further simplifies model specification and hypothesis testing, since it makes it easy to compute the correct asymptotic covariance matrix of $\hat{\beta}$ even if ε_{ijt} is not independent and identically distributed.⁶ All programs I am aware of require partial regression or the bootstrap to obtain standard errors for the TWFE. This exacerbates the problem that iterative methods are slow, because the problem has to be solved K times for partial regressions and a large number of times to bootstrap them.⁷ The fact that it is so difficult to obtain the variance matrix of the slopes may help to explain why most applications keep specification tests to a minimum or avoid them altogether despite the fact that misspecification in these models is just as problematic as in OLS problems of regular size. The method I propose here yields the block of the inverse of the $X'X$ matrix corresponding to $\hat{\beta}$, so it is simple to compute the asymptotic variance matrix regardless of whether ε_{ijt} is iid across observations or not. Most papers calculate SEs under the iid assumption even though it is frequently violated in panel data and can lead to seriously distorted rejection rates even when including fixed effects (Kezdi 2004). In most applications, it is reasonable to assume that errors are correlated within both units for which fixed effects are included, that is, to assume that errors are correlated both within firms and individuals. The bootstrap is infeasible in such cases, because it is basically impossible to preserve the connectedness of the sample when re-sampling (see Woodcock 2015 Appendix B for subsampling connected groups). However, two-way clustered standard errors of the slopes are simple to obtain based on the transformed covariance matrix from the partial regression to obtain $\hat{\beta}$ above. One can use the formula to calculate $V(\hat{\beta})$ from one-way clustering matrices in Cameron, Gelbach and Miller (2011), which is computationally simple. In the two-way clustering case, it reduces to

⁶ See Cattaneo, Jansson and Newey (2016) for a discussion of whether standard asymptotic apply. However, the alternative asymptotic formulas also require the (upper block) of $([X D_\theta D_\psi]'[X D_\theta D_\psi])^{-1}$ (see e.g. Cattaneo, Jansson and Newey, 2016, forthcoming), so the problems discussed below apply in the same way.

⁷ Abowd, Kramarz and Roux (2006, appendix B) suggest that SEs could be obtained by a regression of $y - \hat{\theta} - \hat{\Psi}$ on X , which yields $\sigma_\varepsilon^2(X'X)^{-1}$ instead of the correct OLS variance matrix, $\sigma_\varepsilon^2([X D_\theta D_\psi]'[X D_\theta D_\psi])^{-1}$. However, the SEs in Abowd, Kramarz and Roux (2006) are not calculated this way and are correct.

$$V(\hat{\beta}) = (\tilde{X}'\tilde{X})^{-1}[\hat{B}^J + \hat{B}^I + \hat{B}^S](\tilde{X}'\tilde{X})^{-1} \quad (12)$$

\tilde{X} is the transformed data matrix that is obtained as a byproduct of the partial regression above and \hat{B}^J , \hat{B}^I and \hat{B}^S are the $\tilde{X}'\hat{\varepsilon}\hat{\varepsilon}'\tilde{X}$ -matrices obtained from clustering at the firm, individual and match level. Clustering on both units allows the residuals to be arbitrarily correlated within matches, so no additional clustering is necessary for the match effects model.

4. Application to Wage Determinants and Matching in Germany

I use the method introduced above to examine wage determinants and matching. The first part of the application shows that the data reject the assumptions of computationally simpler models and estimates of both coefficients and fixed effects are substantially biased when omitting match effects. The methods proposed above make it simple to test for and avoid such misspecification. The second part of the application illustrates that analyzing the estimated fixed effects can help to examine interesting questions. While it is primarily an illustration, the results underline that conditions preceding a new job, such as whether there was intermittent unemployment, predict the firm and match fixed effect of the subsequent job, i.e. how the wage of an individual differs from what the same person would receive at a randomly chosen job. This helps to explain the bias I find and underlines that estimates of the fixed effects can provide evidence on economic mechanisms and theories, even if they are not interpreted as structural parameters. For example, the results favor search and signaling over human capital loss as explanations for long term wage losses after unemployment and selection over statistical discrimination as the cause of gender differences. Thus, decomposing the fixed effects into individual, firm and match components is a useful advantage of the method in this paper.

I use linked employer employee-data from Germany (LIAB mover model 9308) which the German Institute for Employment Research (IAB) created by linking social security records to panel data on firms. The firm data stem from the IAB Betriebspanel (Fischer et al. 2008), a panel based on yearly interviews with managers of the firms starting in 1993. The IAB created a linked employer-employee dataset by

matching this data to social security records of the individuals working at the firms in the panel (Jacobebbinghaus 2008). The data cover 1993 to 2008 and are designed to estimate models with individual and firm fixed effects. They only include firms that employed at least one worker who also worked for another firm in the panel, i.e. only firms for which the firm effect is identified. They contain all workers who moved between firms in the panel. Since the individual data stem from social security records, they include very accurate employment biographies. This makes it possible to create biographic covariates such as exact work experience, tenure, year and age at which a person entered the labor market as well as information on job transitions and unemployment. Summary statistics for all variables I use are in appendix E. I follow the common practice to restrict the sample to the largest connected group, which contains 99% of all observations. In addition to all movers (713,559 individuals), the IAB sampled up to 500 employees from each firm. I exclude individuals from the former GDR (East Germany) who entered the labor market before 1990, because their records may not be complete. Appendix D provides further information on the data, potential issues such as topcoding as well as the sample I use. The final sample contains 9,891,519 observations from 3,068,373 individuals working at 24,323 firms. I first use the algorithm described above to estimate a rich wage regression. I regress the log of person i 's daily wage at firm j in time period t (w_{ijt}) on characteristics of person i (x_{ijt}^I), and firm j (x_{ijt}^F) as well as year, firm, individual and match fixed effects:

$$\log(w_{ijt}) = \theta_i + \Psi_j + \lambda_s + \phi_t + x_{ijt}^F \beta^F + x_{ijt}^I \beta^I + \varepsilon_{ijt} \quad (13)$$

Summary statistics for this regression are in Table A1 in appendix E. To analyze the bias from omitting the match effects, I also estimate the TWFE model, which omits λ_s from equation 13. I allow for two-way clustering, i.e. errors can be arbitrarily correlated within firms and individuals. Since income is topcoded at the social security limit for some individuals, I include a dummy for top-coded observations in all specifications. As a robustness check, I repeat the entire analysis excluding individuals with any university education, which reduces the problem of topcoding. The results are similar, so I focus on

selected results from the match effects and TWFE models using the full sample. Full results of all models are in appendix F. I do not estimate more restrictive random and mixed effects models, since the results show that the restrictions they impose do not hold.

I then use the estimates of the fixed effects from the models above in the following regressions:

$$\begin{aligned}\hat{\Psi}_j &= z_j^F \delta + \eta_j^F \\ \hat{\theta}_i &= z_i^I \gamma + \eta_i^I\end{aligned}\tag{14}$$

where z_j^F are time invariant characteristics of firm j and z_i^I are time invariant characteristics of individual i . Summary statistics are in Table A2 and Table A3 in appendix E. Such models can be used to examine which permanent characteristics make some firms pay high wages and some individuals receive high wages. Greene (2011) discusses the assumptions that are required to interpret them as causal. However, the main objective of these regressions here is to show that the estimated fixed effects and their relation to time invariant individual characteristics differs systematically between the TWFE and the match effects model to emphasize the importance of correctly specifying equation 13 even if only the fixed effects are analyzed such as in models of teacher value added.

The last part of the analysis underlines that it is useful to be able to estimate the fixed effects and decompose them into firm and match effects. The sum of the firm and match fixed effect is the expected permanent difference between the wage of individual i in the actually realized match and the wage of the same person in a randomly chosen job. Several studies have found that events at or preceding the start of a new job (e.g. recessions or mass layoffs) have permanent effects on wages. Being able to estimate the fixed effects and relating them to pre-match characteristics, such as whether the match was preceded by an unemployment spell, allows me to assess whether these effects are partly due to an effect of conditions when matches are formed on the way individuals are matched to jobs. Decomposing the permanent wage difference into a match and a firm effect reveals whether and how these conditions affect working for lower or higher paying firms and how it compares to their

effects on match quality. In particular, I regress match effects and the corresponding firm effects on pre-match characteristics:

$$\begin{aligned}\hat{\lambda}_s &= z_s^M \pi_1 + \eta_s^M \\ \hat{\Psi}_s &= z_s^M \pi_2 + \nu_s^M\end{aligned}\tag{15}$$

where $\hat{\Psi}_s$ is the estimated firm fixed effect of the firm that match s corresponds to and z_s^M is a vector of variables measured at the time match s is formed. The first regression provides information on the characteristics that predict a good fit between the individual and the firm, while the second regression examines what leads to employees being matched with high-wage or low-wage firms. Summary statistics are in Table A4 in appendix E.

Results

Table 2 reports selected coefficients from the wage regressions defined by equation 13.⁸ The first question of interest is whether match effects are important and if so, whether specification tests allow us to detect and avoid misspecification of the fixed effects structure. The difference in the explained sum of squares between the TWFE and the match effects model in table A5 indicates that, as Woodcock (2015) shows for the US, match effects explain a relevant part of the overall wage dispersion in Germany. Formally, an F-Test based on this difference in the explained sum of squares rejects the TWFE in favor of the match effects model (p-value of 0). However, this test is only valid under the assumption that the errors are independent, which is unlikely to hold. In practice, it is likely preferable to test whether match effects matter by conducting a Hausman-type test for equality of the slope coefficients between the two models. This test can only detect match effects that are correlated with the covariates, but still rejects the TWFE here. Conducting this test requires estimating the variance matrices of the slope coefficients, thereby emphasizing an advantage of the method proposed here that automatically yields an estimate of the variance matrix.

⁸ Table A5 in appendix F reports the full results.

Table 2: Regressions of log of Daily Wage on Time Variant Characteristics, Selected Coefficients

	Match Effects	TWFE
Part time job	-0.3395*** (0.0046)	-0.4446*** (0.0059)
Number of years in current establishment	0.0099*** (0.0017)	0.0066*** (0.0002)
Experience in years	0.1579*** (0.0026)	0.1171*** (0.0014)
Experience in years ²	-0.0032*** (0.0001)	-0.0047*** (0.0002)
Experience in years ³	0.000097*** (0.000004)	0.00016*** (0.000006)
Experience in years ⁴	-0.000001*** (0.00000005)	-0.000002*** (0.00000008)
Age at end of year ²	-0.0022*** (0.0001)	-0.0039*** (0.0001)
Age at and of year ³	0.000039*** (0.000002)	0.000063*** (0.000002)
Age at end of year ⁴	-0.00000027*** (0.00000001)	-0.00000039*** (0.00000001)

Notes: Number of observations: 9792405. The regressions also include total number of employees, the fraction of employees working part time, fraction of female employees, DHS employment growth index, number of employees that left, number of new employees and dummies for the following variables: business volume, business volume per employee, investment per employee, if they wanted to hire, but did not, whether total employment increased, calendar years and whether an observation was topcoded. Table A5 in Appendix F reports all coefficients. Note that the linear age term is perfectly collinear with the individual and time dummies and therefore omitted from the model. Standard errors are clustered at the firm and individual level. Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

In addition to allowing for a more credible test of model specification, estimating the variance of the slope coefficients is also important to answer substantive questions. Many applications are primarily interested in the slope parameters. So another important question is whether the omission of match effects biases any particular coefficients, i.e. whether it leads to wrong conclusions. Conducting hypotheses tests and inference on these coefficients requires estimating their variance matrix. Comparing the results from the TWFE and the match effects model in table 2 reveals substantive differences: The returns to age, experience and tenure are biased in the TWFE, which suggest a flatter experience and tenure profile⁹. Part of the large difference between part- and full-time jobs in the TWFE is also due to the omission of match effects. The formulas in Woodcock (2008) show that this bias stems from the partial correlation of X with the match effects: For example, part-time jobs tend to be worse

⁹ For example, at 20 years of experience, the log wage difference in the experience profile between the two models is 1.07, which implies that the TWFE underestimates the accumulated returns by 47 percent. The difference increases further at higher levels of experience.

matches than full-time jobs, so failing to control for match quality overstates the wage difference between part- and full-time jobs.

Table 3: Regressions of Individual FE on Time Invariant Individual Characteristics, Selected Coefficients

	Match Effects	TWFE
Female	-0.0559*** (0.0032)	-0.1035*** (0.0029)
<i>School education and vocational training</i>		
Secondary / intermediate school w/o completed vocational training	-0.2325*** (0.0042)	-0.2605*** (0.0038)
Upper secondary school w/o completed vocational training	-0.3391*** (0.0152)	-0.3914*** (0.0136)
Upper secondary school with completed vocational training	0.2338*** (0.007)	0.1849*** (0.0065)
Completion of a university of applied sciences	0.3861*** (0.007)	0.3147*** (0.0064)
College / university degree	0.5251*** (0.006)	0.4205*** (0.0053)
Missing	-0.0248*** (0.0061)	-0.1286*** (0.0056)

Omitted Category: *School education and vocational training: Secondary/intermediate school with completed vocational training*

Notes: Number of observations: 3062118. Regressions also include dummies for nationality (grouped), the year in which an individual first entered the labor market, the age at first employment and an intercept, but coefficients are not reported. Table A7 in Appendix F reports all coefficients. Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

The results of regressions of the individual fixed effects on time invariant individual characteristics in Table 3 confirm that the TWFE model leads to biased estimates not only of the slope coefficients, but also affects the fixed effects.¹⁰ Most importantly, the estimated gender wage gap in the TWFE is almost twice as large as the 5% gap in the match effects model and the education gradient is steeper in the match effects model. This shows that the bias from misspecification spreads to the estimates of the fixed effects. Woodcock (2015) derives formulas for the bias in the fixed effects caused by the omission of match effects that clarify the sources of this bias. Since individual and match fixed effects are orthogonal by construction, the bias arises only because omitted variable bias spreads from time-varying characteristics. For example, the omission of match effects negatively biases the effect of part time jobs. Since females are more likely to work in part time jobs, part of this bias spreads to the individual fixed effects and thus the gender wage gap. Despite the fact that it is an indirect effect, the

¹⁰ Table 3 reports only selected coefficients. See table A7 in appendix D for the full results.

bias is substantial. This emphasizes that even if one is only interested in the fixed effects, one should not be casual about the choice of regressors and the fixed effect structure. In addition to wage gaps, match effects matter, for example, for teacher value added (Jackson 2013). Yet, they are rarely included in these models.

The fact that there are sizeable biases that spread to other parameters underlines the importance of careful model specification tests even of the parts of the model that may not be of primary interest to the researcher. Specification tests are even more important for the random and mixed effects models that are often estimated instead of fixed effects models to simplify computation. The random effects model assumes that the random effects are uncorrelated with the variables in the model. This assumption does not hold here, since otherwise the omission of match effects would not lead to bias. Mixed effects models rely on the weaker assumption that the true covariance between the three sets of fixed effects is zero. This assumption is rejected by the data as well, the correlation between the estimated firm and individual fixed effects is -0.37 in the match effects model and -0.10 in the TWFE and both are significantly different from zero. This correlation is known to be biased downwards (Andrews et al. 2008) and provides only limited evidence about the underlying search model (e.g. Eckhout and Kircher 2011). However, together with the change in the slope parameters in table 3, the results provide clear evidence that the computational simplicity of random and mixed effects models comes at the expense of biased estimates. The method presented above simplifies computation and tests of the corresponding fixed effects models and thereby helps to avoid this bias.

Besides the slope coefficients, in many applications estimates of the fixed effects are of interest. Studies of the sources of inequality (e.g. Card, Heining and Kline 2013) or teacher value added (e.g. Rockoff 2004) are directly interested in the estimated fixed effects. Many other applications compute functions of the fixed effect, such as studies of wage differentials (e.g. Woodcock 2015), wage dynamics (e.g. Abowd, Kramarz and Roux 2006) or teacher mobility (Jackson 2013). However, estimating the fixed effects is

computationally burdensome and therefore often not done or they are estimated under orthogonality assumptions that simplify computation, but are likely to result in bias. The method in this paper simplifies computation of the fixed effects without further assumptions.

In order to illustrate this advantage, I analyze how pre-match characteristics affect the permanent wage components of the subsequent job, i.e. the firm and match fixed effect. Since the sum of the two effects is the expected difference between the wage of the job an individual accepted and the wage of the same individual at a randomly chosen job, this provides evidence on the circumstances that lead to matches with permanently high or low wages. Permanent wage effects of the conditions under which matches are formed have been documented for, among others, employment after mass layoffs (e.g. Jacobson, LaLonde and Sullivan 1993), economic conditions during job search (e.g. Oreopolous, von Wachter and Heisz 2012) and age of the employee (e.g. Dustman and Pereia 2008, Topel and Ward 1992). Most previous papers restricted attention to jobs after a particular event, such as mass layoffs, that affect only a small fraction of individuals in order to isolate the effect of a specific factor (e.g. involuntary unemployment in case of mass layoffs) on subsequent wages. Being able to estimate both firm and match fixed effects for a large sample of movers, I look at these permanent wage effects from a different angle. Rather than restricting the sample to specific job transitions, I use all individuals with two or more jobs and exploit the rich information in the IAB data to measure the circumstances before employment. So unlike most of the existing literature, this analysis does not focus on a specific type of mechanism or job transition (e.g., transitions following a mass layoff). It employs rich data with comprehensive coverage to illustrate how the fixed effects produced by the method introduced in this paper can be used to shed light on the importance of various mechanisms.

To do so, I estimate equation 15, which relates the match and firm effect of a job to pre-match characteristics.¹¹ The results are in table 4:¹² The first column provides evidence on the conditions under which job transitions lead to good matches between worker and firm, i.e. matches with wages that are unusually high for both the worker and the firm. The second column examines which characteristics lead to a worker being matched with a high-wage firm, i.e. a firm that pays high wages to all employees.

I first examine the long run effects of the type and length of the previous labor market status. Several papers have found lasting negative effects of work interruptions and unemployment (e.g. Albrecht et al. 1999, Burda and Mertens 2001, Jacobson, LaLonde and Sullivan 1993), while the gains from job-to-job transitions have been found to increase in the duration of previous employment (e.g. Abowd, Kramarz, Roux 2006). The detailed biographic information in the IAB data allows me to construct indicator variables for the type of transition (job-to-job, non-employment-to-job, training-to-job and initial job)¹³ as well as the number of days in the previous state. Type and duration of the previous spell are key factors in the models of search, signaling and human capital depreciation that are used to explain these long run wage effects. See e.g. Kroft, Lange and Notowidigdo 2013, Pissarides 1992 and Schoenberg 2007, for discussions and Schmiuder, van Wachter and Bender (2016) for an analysis of the effects of unemployment duration that goes far beyond the illustration here. The coefficients on the dummies for the type of transition in table 4 are as expected. The omitted category is job-to-job transitions, so initial employment and employment after an episode without employment is at lower paying firms than job-to-job transitions (by 8 and 6 percentage points) and as most search models would predict, initial employment is a slightly worse match.

¹¹ The regression only uses individuals for whom non-zero match effects are identified, i.e. movers. It includes all of their jobs including initial employment and employment after work interruptions. I do not consider job transitions within the same firm, since they are likely to be different in terms of job search and signaling.

¹² Table A8 in Appendix F reports all coefficients.

¹³ Transition type is defined using labor market status 8 days before the current match to avoid misclassifying job-to-job transitions with a short break.

Table 4: Regression of Match and Firm FE on Pre-Match Characteristics, Selected Coefficients

	Match Effect	Firm Effect
<i>Employment status 8 days before current match</i>		
Previous spell was benefits or gap	-0.0012 (0.0039)	-0.057*** (0.0107)
Apprentice/trainee at other Firm	0.0074 (0.0155)	-0.0164 (0.0488)
No previous record	-0.04*** (0.0066)	-0.0796*** (0.0176)
Number of days in labor market status 8 days before current match (main effect)	-0.000001 (0.000001)	0.000009*** (0.000002)
...if previous spell was benefits or gap (interaction)	0.000001 (0.000003)	0.000003 (0.000008)
...if previous spell was training (interaction)	0.000016 (0.000019)	0.000009 (0.000058)
Number of years of benefit receipt up to beginning of current match	0.0027* (0.0012)	-0.0301*** (0.0032)
Part time job (at beginning of match)	-0.1187*** (0.0047)	-0.6257*** (0.0151)
Female	-0.0529 (0.0433)	-0.376*** (0.1055)
Years since first employment at beginning of current match	0.0024** (0.0009)	-0.0046* (0.0022)
...if female	-0.0024 (0.0014)	-0.0132*** (0.0036)
Years since first employment squared	-0.000129*** (0.00000008)	-0.000058 (0.00000021)
...if female	0.000082 (0.00000013)	0.000457*** (0.00000036)
Age at beginning of current match	0.0024 (0.0015)	0.0364*** (0.0038)
...if female	0.0043 (0.0025)	0.0221*** (0.0059)
Age squared	-0.00002 (0.000019)	-0.000387*** (0.000048)
<i>Match Count</i>		
2	-0.0063 (0.0043)	0.0612*** (0.0117)
...interacted with female dummy	0.0071 (0.0065)	0.0027 (0.0207)
3	-0.0229** (0.0087)	0.0958*** (0.0236)
...interacted with female dummy	0.0208 (0.0142)	-0.0253 (0.052)
4	-0.0576* (0.0252)	0.1722*** (0.0401)
...interacted with female dummy	0.0807 (0.0562)	0.1501 (0.1395)
R^2	0.0343	0.2189

Omitted Categories: *Emp. Status 8 days before current match:* Employment at other Firm; *Match Count:* 1

Note: Number of observations: 665080. Estimated match and firm fixed effects are from main regression on full sample. The regression also includes dummies for the 5th and 6th match count, but the sample contains few individuals with 5 or more matches, so the coefficients on these dummies are only reported in the appendix and should be interpreted cautiously. It also includes an intercept and dummies for the year the match started, but coefficients are not reported. Table A8 in Appendix F reports all coefficients. Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

In addition to the type of the previous spell, its duration may matter for several reasons: It may reflect accumulation or depreciation of human capital and longer spells may allow for a more thorough job

search or send a signal to potential employers. The model includes interactions of the type of spell with its duration to allow the effect of duration to differ across spell types. The omitted category is job-to-job transitions, i.e. the main effect is the effect of an additional day at the previous employer in a job-to-job transition and the interaction terms capture the difference from this effect if the previous spell was unemployment or training. The positive coefficient on the main effect is consistent with previous findings that the gains from switching jobs increase with the length of the current job, but my estimates show that whether an individual had a job or not is more important than its duration: A job-to-job transition preceded by a 10-year employment spell leads to a meager 3 percent higher firm effect than a job-to-job transition after a short employment spell.

To analyze the effect of unemployment duration, the model contains both the number of years of benefit receipt accumulated up to the beginning of the current spell and the interaction of the length of the previous spell with the non-employment dummy. The former captures (365 times) the effect of an additional day of unemployment at any earlier time, while the coefficient on the interaction is the added effect if this additional day of unemployment was part of the last spell. As expected by signaling or human capital models, an additional accumulated year of unemployment leads to employment at a firm that pays slightly lower wages (by 3 percent).¹⁴ However, the interaction term indicates that the duration of the current unemployment spell does not reduce subsequent wages any further. The main difference between the two measures is that current unemployment duration also reflects the benefit of longer job search, so one may take this as evidence that the benefits of search duration are smaller than the harmful signaling effect of longer unemployment.

Overall, the fact that spell type is more predictive of subsequent wages than duration provides evidence that signaling plays a more important role in explaining permanent wage effects at job transitions than human capital loss. The latter would either predict a smoother pattern if human capital depreciates over

¹⁴It also leads to a slightly better match effect, although the effect is practically negligible (< 0.3 percent per year).

time or a similar effect for all types of job transitions if human capital is lost because it is firm specific. The results are also supportive of common search models: Longer periods of benefit receipt are associated with slightly higher complementarity between worker and firm, the initial match is slightly worse and, while insignificant and small, the coefficients on the duration of a current unemployment spell are both positive.

While involuntary job transitions have long lasting negative effects, job transitions are also an important source of wage growth as shown by Topel and Ward (1992) for young men in the US. Dustmann and Pereira (2008) use data from Germany and the UK to show that these gains decrease later in the life-cycle. They suggest that this pattern is consistent with declining benefits of job search. The results in table 4 clarify that permanent wage changes at job transitions are primarily due to moving to higher paying firms rather than obtaining better matches between employer and employee: The small coefficients and the low R^2 in column 1 of table 4 show that the relation between match quality and observable pre-match characteristics is weak compared to firm effects. This indicates that if wage gains at job transitions are indeed driven by job search, individuals primarily search for high paying firms rather than high paying matches. Consequently, firm heterogeneity not only plays an important role in wage dispersion (Gruetter and Lalive 2009) and rising inequality (Card, Heining and Kline 2013), but is also a key part of wage increases over the life-cycle.

The literature on wage growth at job transition has focused on men, but gender differences in mobility, search and earnings (e.g. Altonji and Blank 1999, Hunt 2002, Loprest 1992) are well documented. I include both males and females in the analysis, but interact all key variables with gender to allow coefficients to differ by gender. The results show that job changes are an important source of wage growth for both males and females, but their life-cycle match profiles differ substantively. For males and females, both firm and match effect increase with age and time in the labor market initially. That is, all else equal a transition at age 30 will lead to a better match and a higher paying firm than a transition at

age 25. For males, they peak between 30 and 45 (depending on age at entry) and both components decline with age from there on. For females, the profile of the permanent wage components is steeper and increases over the entire life-cycle. Thus, contrary to males, for females job changes continue to play an important role for wage growth later in their careers. The fact that match effects are systematically related to age and that this relation is stronger for females helps to explain the bias in the returns to experience and the gender wage gap when omitting the match effects in the TWFE above.

Gender differences in the match effect could be due to statistical discrimination, but the contribution of the match effect is small. The majority of the difference comes from the fact that, contrary to males, the firm effect continues to increase in age at the transition even after age 40. This is unlikely to be due to statistical discrimination, since these firms pay higher wages to all employees. A more plausible explanation for the steeper wage profile for females is a greater impact of selection into the labor market for females (see e.g. Hunt 2002, Jurajda and Hamgart 2007). Lower labor force attachment may make females on average less likely to accept jobs at low wage firms and drop out of the labor market instead, thereby leaving a more selected sample of females matched with high-wage firms in the data.

5. Conclusion

Models with two or more large sets of fixed effects are heavily used in a wide range of fields such as labor economics, education, health or migration. While their advantages are widely recognized, their use is limited by computational difficulties of estimation. Researchers often apply estimation strategies that are likely to lead to bias or do not calculate important parameters in order to make computation feasible. This at best limits the scope of their analyses and often casts doubt on the validity of their conclusions. In addition, the computational cost is likely to be partly responsible for the fact that researchers rarely make full use of these large data sets: Their size would easily allow including match effects or at least testing for them, conducting thorough specification tests and taking into account that errors are rarely independently distributed. A likely reason for this is that current estimation routines

make it costly to estimate match effects and the variance matrix that is needed for specification tests. Even worse, it is basically impossible to obtain clustered standard errors.

I propose a simple method to estimate the TWFE and the match effects model in large data sets that solves or greatly reduces the computational complexity that causes these problems. The method produces the exact OLS estimates without any additional assumptions. Compared to other exact methods, it not only offers advantages in terms of speed and computational resources needed, but also provides estimates of the fixed effects and the variance matrix of the slopes that can allow for multi-way clustering. Using it to analyze wage determinants and matching in Germany underlines the importance of solving the computational problems: The results reject the orthogonality assumptions of random and mixed effects models, so their computational simplicity is likely to come at the expense of biased estimates. Standard specification tests also reject the TWFE in favor of the match effects model and adding match fixed effects substantively changes parameter estimates, including the estimated fixed effects. The method I propose makes these specification tests feasible or greatly facilitates them. My results show that such tests are capable of selecting the correct model and thereby necessary to avoid the sizeable biases I find. The application underlines that it is valuable to estimate the three sets of fixed effects, because analyzing them can provide insights on economic questions regardless of their status as structural parameters. In particular, I find that the long run effects of labor market status before starting a job are driven by the type of status and not by its duration, which favors explanations in terms of signaling and search. Firm fixed effects play a more important role in gains from switching jobs than match effects and the life-cycle profile of such gains differs substantively by gender. Since this effect is driven by the firm fixed effect, it is likely to be due to selection rather than statistical discrimination.

References

- Abowd, J.M. and F. Kramarz 1999, The Analysis of Labor Markets Using Matched Employer-Employee Data, in: O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3(B) Chapter 26. North Holland, Amsterdam, pp. 2629-2710.
- Abowd, J. M., F. Kramarz, and D. N. Margolis, 1999, High wage workers and high wage firms. *Econometrica* 67 (2), 251-334.
- Abowd, J. M., R. H. Creecy, and F. Kramarz, 2002. Computing person and firm effects using linked longitudinal employer-employee data. Census Bureau Technical Paper TP-2002-06.
- Abowd, J. M., F. Kramarz, and S. Roux, 2006, Wages, Mobility and Firm Performance: Advantages and Insights from Using Matched Worker-Firm Data. *The Economic Journal*, 116(512), F245–F285.
- Abowd, J.M., F. Kramarz and S.D. Woodcock, 2008, *Econometric Analyses of Linked Employer-Employee Data*, in: L. Mátyás, P. Sevestre (Eds.), *The Econometrics of Panel Data*, Chapter 22. Springer, Heidelberg, pp. 727-759.
- Albrecht, J. W., P. Edin, M. Sundstrom and S. B Vroman. 1999. Career Interruptions and Subsequent Earnings: A Reexamination Using Swedish Data. *The Journal of Human Resources*, 34(2), 294–311.
- Altonji, J. G., and R.M. Blank. 1999. Race and Gender in the Labor Market. In: O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics*, Vol. 3, Chapter 48. Amsterdam: Elsevier B.V., pp. 3143–3259.
- Andrews, M., T. Schank and R. Upward, 2006. Practical fixed-effects estimation methods for the three-way error-components model. *Stata Journal*, 6(4), 461-481.
- Andrews, M., L. Gill, T. Schank and R. Upward, 2008. High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society A*. 171(3), 673-697.
- Arcidiacono, P., G. Foster, N. Goodpaster, and J. Kinsler, 2012. Estimating Spillovers using Panel Data, with an Application to the Classroom. *Quantitative Economics*. 3(3), 421–470.
- Bennett, D., C. L. Hung, and T. L. Lauderdale, 2015. Health care competition and antibiotic use in Taiwan. *The Journal of Industrial Economics*, 63(2), 371-393.
- Burda, M.C., and A. Mertens, 2001. Estimating wage losses of displaced workers in Germany. *Labour Economics*, 8(1), 15–41.
- Cameron, C. A., J. B. Gelbach and D.L. Miller, 2011, Robust Inference With Multi-Way Clustering. *Journal of Business and Economics Statistics* 29 (2), 238-249.
- Card, D., J. Heining, and P. Kline, 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3), 967-1015.
- Cattaneo, M. D., M. Jansson, and W. K. Newey, forthcoming. Alternative Asymptotics and the Partially Linear Model with Many Regressors. *Econometric Theory*.

- Cattaneo, M. D., M. Jansson, and W. K. Newey, 2016. Inference in Linear Regression Models with Many Covariates and Heteroskedasticity. Working Paper.
- Chetty, R., J.N. Friedman, and J.E. Rockoff, 2011. The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. NBER Working Paper 17699.
- Chetty, R., J.N. Friedman, and J.E. Rockoff, 2014. Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9), 2633–2679.
- Cornelissen, T., 2008. Using Stata for a memory-saving fixed-effects estimation of the three-way error-components model. *Stata Users Group German Stata Users' Group Meetings 2008* 07.
- Correia, S., 2014. "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects," *Statistical Software Components S457874*, Boston College Department of Economics, revised 25 Jul 2015.
- Davis, P. 2002. Estimating multi-way error components models with unbalanced data structures. *Journal of Econometrics* 106, 67-95.
- Dustman, C. and S.C. Pereira, 2008. Wage growth and job mobility in the United Kingdom and Germany. *Industrial & Labor Relations Review*. 61(3), 374-393.
- Eeckhout, J. and P. Kircher, 2011. Identifying sorting - in theory. *Review of Economic Studies*. 78, 872–906.
- Fischer, G., F. Janik, D. Müller and A. Schmucker 2008, The IAB establishment panel - from sample to survey to projection. *FDZ Methodenreport 01/2008* (en).
- Garen, J.E. 1989. Job-Match Quality as an Error Component and the Wage-Tenure Profile: A Comparison and Test of Alternative Estimators. *Journal of Business & Economic Statistics*, 7(2), 245-252.
- Gaure, S., 2013. OLS with multiple high dimensional category variables. *Computational Statistics & Data Analysis*, 66, 8–18.
- Greene, W.H. 2008, *Econometric Analysis*, 6th Edition. Prentice Hall, New York.
- Greene, W.H. 2011, Fixed Effect Vector Decomposition: A Magical Solution to the Problem of Time Invariant Variables in Fixed Effects Models. *Political Analysis* 19(2), 135-146.
- Grogger, J. and G.H. Hanson, 2011, Income maximization and the selection and sorting of international migrants. *Journal of Development Economics* 95, 42-57.
- Gruetter, M. and R. Lalive, 2009. The importance of firms in wage determination. *Labour Economics*, 16(2), 149–160.
- Guimarães, P. and P. Portugal, 2010. "A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects". *Stata Journal*, 10(4), 628-649.
- Hunt, J., 2002. The transition in East Germany: When is a ten per cent fall in the gender pay gap bad news. *Journal of Labor Economics* 20(1), 148–169.

- Jackson, C.K. 2013, Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *Review of Economics and Statistics*, 95(4), 1096–1116.
- Jacob, B.A., and L. Lefgren, 2008. Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education. *Journal of Labor Economics*, 26(1), 101–136.
- Jacobebbinghaus, P. 2008, LIAB-Datenhandbuch, Version 3.0. FDZ Datenreport, 03/2008 (de).
- Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan, 1993. Earnings Losses of Displaced Workers. *American Economic Review*, 83(4), 685–709.
- Jovanovic, B. 1979. Firm-specific Capital and Turnover. *Journal of Political Economy* 87(6), 1246-1260.
- Jurajda, Š., and H. Harmgart, 2007. When do female occupations pay more? *Journal of Comparative Economics*, 35(1), 170-187.
- Kezdi, G. 2004, Robust Standard Error Estimation in Fixed-Effect Panel Models. *Hungarian Statistical Review. Special English Volume 9*, 95-116.
- Koch, I. and H. Meinken 2004. The Employment Panel of the German Federal Employment Agency. *Journal of Applied Social Science Studies* 124(2), 315-325.
- Kramarz, F., S. Machin and A. Ouazad 2008. What Makes a Test Score? The Respective Contributions of Pupils, Schools, and Peers in Achievement in English Primary Education. IZA Discussion Paper 3866.
- Kroft, K., F. Lange, & M.J. Notowidigdo 2013. Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment. *The Quarterly Journal of Economics*. 128(3), 1123-1167.
- Le Gall, F., 2014. Powers of Tensors and Fast Matrix Multiplication, 1–28. Retrieved from <http://arxiv.org/abs/1401.7714v1>
- Loprest, P.J. 1992. Gender Differences in Wage Growth and Job Mobility. *The American Economic Review*, 82(2), 526–532.
- Markussen, S. and K. Røed, 2015. Social Insurance Networks. *Journal of Human Resources*, 50(4), 1081-1113.
- Mortensen, D.T. 1978. Specific Capital and Turnover. *The Bell Journal of Economics* 9(2), 572-586.
- Ouazad, A., 2007. Program for the Estimation of Two-Way Fixed Effects, available at <http://personal.lse.ac.uk/ouazad/>.
- Oreopoulos, P., T. von Wachter and A. Heisz, 2012. The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics*, 4(1), 1–29.
- Pissarides, C.A. 1992. Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics*, 1371-1391.
- Rockoff, J. E. 2004. The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review*, 94(2), 247–252.
- Schönberg, U. 2007. Wage growth due to human capital accumulation and job search: a comparison between the United States and Germany. *Industrial & Labor Relations Review*, 60(4), 562-586.

Theil, H. 1971, *Principles of Econometrics*. John Wiley and Sons, New York.

Topel, R.H. and M.P. Ward 1992, Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics* 107(2), 439-479.

Wansbeek, T. and A. Kapteyn 1989, Estimation of the Error-Components Model with Incomplete Panels. *Journal of Econometrics* 41, 341-361.

Woodcock, S.D, 2008, Wage Differentials in the presence of unobserved worker, firm and match heterogeneity. *Labour Economics* 15, 772-794.

Woodcock, S.D. 2015, Match Effects. *Research in Economics*.

Yule, G.U. 1907, On the Theory of Correlation for any Number of Variables, Treated by a New System of Notation. *Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character* 79 (529), 182-193.

SUPPLEMENTAL MATERIALS: APPENDICES

Appendix A: Proof of OLS Property 2

The second property used to simplify the estimation follows from the OLS normal equations:

$$\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z \end{pmatrix} \begin{pmatrix} \hat{\beta}^{OLS} \\ \hat{\delta}^{OLS} \end{pmatrix} = \begin{pmatrix} X'y \\ Z'y \end{pmatrix}$$

As $\hat{\beta}^{OLS}$ is assumed to be known, the second line is a system of equations which is uniquely solved by $\hat{\delta}^{OLS}$. Rearranging and solving for $\hat{\delta}^{OLS}$ yields

$$\begin{aligned} Z'X\hat{\beta}^{OLS} + Z'Z\hat{\delta}^{OLS} &= Z'y && \Leftrightarrow \\ Z'Z\hat{\delta}^{OLS} &= Z'(y - X\hat{\beta}^{OLS}) && \Leftrightarrow \\ \hat{\delta}^{OLS} &= (Z'Z)^{-1}Z'(y - X\hat{\beta}^{OLS}) \end{aligned}$$

Regressing $(y - X\hat{\beta}^{OLS})$ on Z as proposed by property 2 yields:

$$\hat{\delta} = (Z'Z)^{-1}Z'(y - X\hat{\beta}^{OLS}) = \hat{\delta}^{OLS}$$

which shows that the estimates from the auxiliary regression are numerically equal to $\hat{\delta}^{OLS}$, the estimates obtained from the full regression.

Appendix B: Summary of the computational steps in the algorithm for the case of multiple groups

1. Identify connected groups. Abowd et al. (2002) describe this algorithm. This is not necessary when using the CGA.
2. Calculate individual, firm and spell means.
3. Estimate the slope coefficients:
 - For the TWFE-model, do the WK transformation for y and X using the Cholesky factorization of $[F - K'T^{-1}K]$, which can be calculated separately by group and stored for later use or by applying the CGA to every variable. Regress the transformed y on the transformed X to obtain the slopes.
 - For the match effects model, run OLS on the deviations of y and X from spell means
4. Calculate SEs of the slopes and required test statistics based on the residuals from step 3.
5. Obtain $\bar{\tilde{y}}_i, \bar{\tilde{y}}_j$ and $\bar{\tilde{y}}_s$ by subtracting $\bar{X}\hat{\beta}^{OLS}$ (where the mean of X is taken over the appropriate index) from the individual, firm and spell means from step 2.
6. Calculate the firm fixed effects for each connected group separately using formula 9. This can be done by using the Cholesky factorization of $[F - K'T^{-1}K]$ (in case of the TWFE-model, it has already been calculated for step 3) or the CGA.
7. Use the firm effects to calculate the individual and match effects.

Appendix C: Simulation Setup

The simulations are based on randomly generated data with 5 time-varying covariates. The covariates in the simulation are drawn from the following joint normal distribution:

$$X \sim N \left[\begin{pmatrix} 5 \\ -6 \\ 0.5 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 9 & 5 & 2 & 3 & 4 \\ 5 & 9 & 1 & 7 & 3 \\ 2 & 1 & 9 & 2 & 1 \\ 3 & 7 & 2 & 9 & 4 \\ 4 & 3 & 1 & 4 & 9 \end{pmatrix} \right]$$

The parameters are arbitrary and there is no serial correlation in the covariates. The individual and firm fixed effects are correlated with the mean of X_1 of the respective individual or firm. Each individual is observed for 10 time periods and individuals move to a randomly chosen firm once with probability 0.16 or twice with probability 0.04. The dependent variable is then generated from the simple linear model

$$y_{it} = x_{it}\beta + \theta_i + \Psi_j + \varepsilon_{it}$$

with $\beta = (1,2,3,4,5)$. ε_{it} is a random normal error term, with the standard deviation fixed at 1/6th of the standard deviation of the resulting dependent variable. For each size of the problem, I run each algorithm on the same five randomly generated data sets and record the average time of computation. All simulations are carried out single-threaded on a server with 74 GB of main memory.

Code to replicate these simulations is available from my website and also provides further detail on the setup. The setup involves several arbitrary choices, such as the number and correlation of covariates, and several simplifications to make data generation feasible, such as random, but high mobility (to ensure connectedness). Computation time depends on features of the data such as the ratio of J to N and the connectedness, but I have no reason to believe that the setup here favors a particular algorithm. The program can easily be modified to examine different setups and I expect similar speed gains, since my implementation avoids some of the most computationally demanding steps of other programs.

Appendix D: Data Description

As discussed above, the LIAB data I used is based on a firm survey conducted by the IAB. It dates back to 1993 (1996 for the former East Germany) and is a stratified random sample of establishments in Germany. There are 43,617 firms in total and between 4265 and about 16,000 firms per year, with a large part of the variation explained by successive expansion of the panel. The matched employer-employee data is created by linking the survey data to administrative records from the German social security system. Fischer et al. (2008) and Jacobebbinghaus (2008) contain a detailed discussion of the data and how it was created, so I only discuss the details and problems that matter for the application in this paper below. In particular, I discuss issues raised by the coverage of social security records, the sampling of individuals and the problem that earnings are topcoded at the social security limit.

While there are good reasons to believe that the records provide very accurate information on earnings and labor market histories, a downside is that they do not include work that is not subject to social security such as self-employment. Approximately 75-80% (Koch and Meinken 2004) of employment in Germany is subject to social security and the data is only representative of this population. In addition, social security records for people from the former GDR do not exist prior to 1990. The data I use starts in 1990, so it does not affect the sampling of individuals, but their labor market histories will not be accurate if they were employed prior to 1990. Consequently, I exclude people from the former GDR who first entered the labor market before 1990.

In order to make the data representative of this population, I adjust the weights from the firm survey. The firm survey includes weights based on sample and population distributions that adjust for non-response and the non-random sampling. These weights are designed to make survey estimates representative of the universe of German firms, which does not extend to the sample of individuals. This is due to the fact that the individual data contains all workers that move between the firms in the sample (713,559) and up to 500 randomly selected employees of each firm who did not move or moved

to firms that are not in the sample. If the firm had fewer than 500 employees subject to social security, all employees are selected. Using the number of individuals at each firm that the IAB used as the sampling frame, I adjust the weights from the firm survey to make the resulting sample representative of the population of employees subject to social security. However, I only use these weights in the regressions with the fixed effects as dependent variables and do not use weights in the main regressions, as they increase the computational complexity considerably. I condition on all variables that were used to create the weights, so the only consequence of this is inefficiency

A final problem with social security data is that earnings are topcoded at the social security limit for some individuals, because employers can report the social security limit instead of the actual income if the latter exceeds the former. This limit is different for the states that belonged to the former GDR and varies by year. The exact values can be obtained from the website of the IAB, in 2007, for example, it was € 63,000/year (East: € 54,600). This affects 6.9% of the sample (4.7% after weighting). In order to account for this problem, the analysis below contains a dummy if an observation is topcoded. Additionally, I repeat the entire analysis excluding individuals with any form of university education. In this restricted sample, only 3.9% of all observations are affected by topcoding (3% after weighting). Overall, the results do not change much and key differences are pointed out in the analysis above. Results from this sample are reported in appendix F.

Appendix E: Summary Statistics

Table A1: Summary Statistics for Main Regression

	Full Sample		Low Education Sample	
	Mean	SD	Mean	SD
Daily wage	4.12	0.77	4.08	0.77
Total number of employees	326.48	1309.04	308.89	1237.32
<i>Business volume, categorical</i>				
0 to 72,000	0.22%	4.69%	0.21%	4.58%
72,000 to 120,000	0.24%	4.89%	0.24%	4.89%
120,000 to 166,200	0.34%	5.82%	0.35%	5.91%
166,200 to 245,400	0.72%	8.45%	0.75%	8.63%
245,400 to 332,300	1.30%	11.33%	1.34%	11.50%
332,300 to 490,000	1.93%	13.76%	1.97%	13.90%
490,000 to 715,800	2.78%	16.44%	2.84%	16.61%
715,800 to 1,227,100	5.18%	22.16%	5.29%	22.38%
1,227,100 to 3,163,900	11.28%	31.63%	11.48%	31.88%
More than 3,163,900	57.85%	49.38%	57.67%	49.41%
Missing	18.16%	38.55%	17.87%	38.31%
<i>Business volume per employee, categorical</i>				
0 to 21,300	3.31%	17.89%	3.40%	18.12%
21,300 to 30,700	3.67%	18.80%	3.77%	19.05%
30,700 to 39,800	4.36%	20.42%	4.39%	20.49%
39,800 to 50,000	6.13%	23.99%	6.09%	23.91%
50,000 to 59,700	4.96%	21.71%	4.93%	21.65%
59,700 to 71,600	5.47%	22.74%	5.46%	22.72%
71,600 to 92,900	8.10%	27.28%	8.16%	27.38%
92,900 to 128,600	11.67%	32.11%	11.73%	32.18%
128,600 to 230,100	16.79%	37.38%	16.86%	37.44%
More than 230,100	17.36%	37.88%	17.34%	37.86%
Missing	18.16%	38.55%	17.87%	38.31%
Fraction of employees working part time	18.10%	21.24%	18.12%	21.38%
<i>Investment per employee, categorical</i>				
0 to 500	19.09%	39.30%	19.37%	39.52%
2,100 to 3,000	6.88%	25.31%	6.94%	25.41%
3,000 to 4,200	12.00%	32.50%	11.99%	32.48%
4,200 to 6,100	12.54%	33.12%	12.47%	33.04%
6,100 to 10,000	13.35%	34.01%	13.33%	33.99%
10,000 to 18,600	13.78%	34.47%	13.74%	34.43%
more than 18,600	15.53%	36.22%	15.45%	36.14%
Missing	6.82%	25.21%	6.71%	25.02%
Fraction of female employees	40.26%	28.05%	40.19%	28.33%
DHS employment growth index	0.02	0.17	0.02	0.17
More employees than previous year	60.89%	48.80%	60.79%	48.82%
<i>Wanted to hire people, but did not</i>				
Successfully hired or did not want to	86.77%	33.88%	86.62%	34.04%
Wanted to hire people, but did not	8.64%	28.10%	8.73%	28.23%
Missing	4.59%	20.93%	4.65%	21.06%
<i>Expected business volume relative to last year</i>				
Same	41.54%	49.28%	41.52%	49.28%
Increasing	31.70%	46.53%	31.50%	46.45%
Decreasing	19.86%	39.89%	20.13%	40.10%
Missing	6.90%	25.35%	6.85%	25.26%
Total number of new employees	12.32	47.25	11.52	43.25
Firm was hiring in current year	69.99%	45.83%	69.46%	46.06%
Total number of employees that left	14.07	61.99	13.26	58.22
Employees have left in current year	75.14%	43.22%	74.83%	43.40%

Number of days in current establishment	3034.97	2688.82	3075.94	2708.80
Age at end of year	40.40	11.65	40.35	11.78
Part time job	20.82%	40.60%	21.30%	40.94%
Daily wage topcoded	4.69%	21.14%	3.01%	17.09%
<i>Year</i>				
1993	6.25%	24.21%	6.35%	24.39%
1994	6.25%	24.21%	6.36%	24.40%
1995	6.25%	24.21%	6.35%	24.39%
1996	6.25%	24.21%	6.36%	24.40%
1997	6.25%	24.21%	6.34%	24.37%
1998	6.25%	24.21%	6.27%	24.24%
1999	6.25%	24.21%	6.28%	24.26%
2000	6.25%	24.21%	6.26%	24.22%
2001	6.25%	24.21%	6.22%	24.15%
2002	6.25%	24.21%	6.23%	24.17%
2003	6.25%	24.21%	6.23%	24.17%
2004	6.25%	24.21%	6.16%	24.04%
2005	6.25%	24.21%	6.16%	24.04%
2006	6.25%	24.21%	6.14%	24.01%
2007	6.25%	24.21%	6.13%	23.99%
2008	6.25%	24.21%	6.14%	24.01%
Experience in years	14.11	8.41	14.24	8.44

Note: Weighted statistics calculated from the IAB LIAB MM 9308.

Table A2: Summary Statistics for Time Invariant Firm Characteristics

	Mean	SD		Mean	SD
<i>Industry</i>			<i>Legal Form</i>		
Agriculture and forestry	1.57%	12.43%	Individually-owned firm	34.30%	47.47%
Mining, quarrying and electricity	0.32%	5.65%	Partnership	7.63%	26.55%
Food products	2.29%	14.96%	Limited liability company	42.01%	49.36%
Clothing and textile	0.72%	8.45%	Company limited by shares	2.72%	16.27%
Paper and printing	1.08%	10.34%	Public corporation	6.28%	24.26%
Wood Products, furniture, jewelry, toys	1.53%	12.27%	Other legal form	5.28%	22.36%
Chemical industry	0.36%	5.99%	Missing/don't know	1.78%	13.22%
Rubber/plastic	0.57%	7.53%	<i>Main/Exclusive Ownership</i>		
Non-metallic mineral products	0.63%	7.91%	Eastern German property	10.99%	31.28%
Basic metals, steel, light metal recycling	3.88%	19.31%	Western German property	51.87%	49.97%
Machinery	2.23%	14.77%	Foreign property	2.33%	15.09%
Motor vehicles:			Public property	2.65%	16.06%
Production/sales/repair/fuel	5.01%	21.82%	No principal shareholder	2.86%	16.67%
Other transport equipment	/	/	Unknown	2.25%	14.83%
Electrical equipment	1.61%	12.59%	Missing	27.05%	44.42%
Precision and optical equipment	1.02%	10.05%	<i>Year founded (only after 1990)</i>		
Main construction trade	5.52%	22.84%	Founded before 1990	40.65%	49.12%
Building installation/completion	5.43%	22.66%	1990	2.63%	16.00%
Sales: retail and wholesale	16.60%	37.21%	1991	3.38%	18.07%
Transportation	5.79%	23.36%	1992	2.34%	15.12%
Communication	0.27%	5.19%	1993	2.50%	15.61%
Credit and financial intermediation	1.20%	10.89%	1994	2.84%	16.61%
Insurance	0.98%	9.85%	1995	2.89%	16.75%
Computer and related activities	1.42%	11.83%	1996	2.02%	14.07%
Research and development	0.41%	6.39%	1997	2.17%	14.57%
Legal consulting, advertising	4.60%	20.95%	1998	2.12%	14.41%
Real estate	1.78%	13.22%	1999	2.10%	14.34%
Renting, business activities	5.82%	23.41%	2000	1.84%	13.44%
Hotel and restaurant	6.27%	24.24%	2001	1.39%	11.71%
Education/teaching	2.26%	14.86%	2002	1.27%	11.20%
Human health, veterinary and social work	9.47%	29.28%	2003	1.40%	11.75%
Sanitation	0.53%	7.26%	2004	1.24%	11.07%
Recreation, culture, sports	1.51%	12.20%	2005	1.19%	10.84%
Other services	2.64%	16.03%	2006	1.01%	10.00%
Organizations, lobbying	2.31%	15.02%	2007	/	/
Public administration and social security	2.14%	14.47%	2008	/	/
<i>State</i>			Missing	24.43%	42.97%
Schleswig-Holstein	4.85%	21.48%	<i>Establishment/department is...</i>		
Hamburg	3.22%	17.65%	Independent company w/o other places of business	75.65%	42.92%
Lower Saxony	10.51%	30.67%	Business/office/branch	15.56%	36.25%
Bremen	1.44%	11.91%	Head office	5.35%	22.50%
North Rhine-Westphalia	17.77%	38.23%	Middle-level authority	1.59%	12.51%
Hesse	8.04%	27.19%	Missing	1.85%	13.48%
Rhineland-Palatinate	5.63%	23.05%	<i>Company pays for job training/courses</i>		
Baden-Wuerttemberg	13.24%	33.89%	No	39.80%	48.95%
Bavaria	15.16%	35.86%	Yes	51.61%	49.97%
Saarland	2.09%	14.30%	Missing	8.59%	28.02%
Berlin	3.91%	19.38%	<i>Has Worker's Council</i>		
Brandenburg	2.69%	16.18%	No	79.98%	40.01%
Mecklenburg-Western Pomerania	2.19%	14.64%	Yes	18.37%	38.72%

Saxony	4.29%	20.26%	Missing	1.65%	12.74%
Saxony-Anhalt	2.50%	15.61%	<i>Collective wage agreement</i>		
Thuringia	2.46%	15.49%	Industry-wide wage agreement	43.05%	49.51%
<i>Owner working in Company</i>			Company agreement	4.56%	20.86%
No	/	/	No collective agreement	44.65%	49.71%
Yes	73.81%	43.97%	Missing	7.74%	26.72%
Missing	/	/			

Note: Weighted statistics calculated from the IAB LIAB MM 9308. If one or more cells contained too few observations, at least two cell frequencies could not be disclosed (to prevent calculation from totals). This is indicated by /.

Table A3: Summary Statistics for Time Invariant Individual Characteristics

	Full Sample		Low Educ. Sample	
	Mean	SD	Mean	SD
Female	42.72%	49.47%	43.50%	49.58%
<i>Nationality, grouped</i>				
Germany	92.17%	26.86%	91.95%	27.21%
Turkey	2.25%	14.83%	2.39%	15.27%
Italy	0.89%	9.39%	0.93%	9.60%
Yugoslavia, Serbia and Montenegro	0.83%	9.07%	0.88%	9.34%
Greece	0.37%	6.07%	0.39%	6.23%
France	0.26%	5.09%	0.25%	4.99%
Poland	0.27%	5.19%	0.27%	5.19%
Austria	0.29%	5.38%	0.28%	5.28%
Croatia	0.19%	4.35%	0.20%	4.47%
Portugal	0.21%	4.58%	0.22%	4.69%
Spain	0.16%	4.00%	0.16%	4.00%
Netherlands, Luxembourg	0.14%	3.74%	0.13%	3.60%
Russia, Belarus, Former Soviet Union	0.11%	3.31%	0.10%	3.16%
Bosnia and Herzegovina	0.12%	3.46%	0.12%	3.46%
Great Britain, Ireland and Northern Ireland	0.14%	3.74%	0.13%	3.60%
Romania	0.08%	2.83%	0.08%	2.83%
Czech Republic, Slovakia, Former Czechoslovakia	0.09%	3.00%	0.09%	3.00%
Ukraine, Moldova	0.06%	2.45%	0.06%	2.45%
Hungary	0.06%	2.45%	0.06%	2.45%
Albania	0.04%	2.00%	0.04%	2.00%
Belgium	0.03%	1.73%	0.03%	1.73%
Macedonia	0.03%	1.73%	0.03%	1.73%
Switzerland	0.03%	1.73%	0.03%	1.73%
Bulgaria	0.03%	1.73%	0.03%	1.73%
Slovenia	0.02%	1.41%	0.02%	1.41%
Denmark, Sweden	0.04%	2.00%	0.03%	1.73%
Finland	0.02%	1.41%	0.01%	1.00%
Estonia, Latvia, Lithuania	0.02%	1.41%	0.02%	1.41%
Europe (other)	0.02%	1.41%	0.02%	1.41%
Ethiopia	0.01%	1.00%	0.01%	1.00%
Ghana	0.03%	1.73%	0.03%	1.73%
Morocco	0.08%	2.83%	0.09%	3.00%
Tunisia	0.04%	2.00%	0.04%	2.00%
Africa (other)	0.12%	3.46%	0.12%	3.46%
USA, Canada	0.10%	3.16%	0.09%	3.00%
America (other)	0.06%	2.45%	0.06%	2.45%
Afghanistan	0.03%	1.73%	0.03%	1.73%
Sri Lanka	0.05%	2.24%	0.05%	2.24%
Vietnam	0.04%	2.00%	0.04%	2.00%
India	0.03%	1.73%	0.03%	1.73%
Iraq	0.04%	2.00%	0.04%	2.00%
Iran	0.06%	2.45%	0.06%	2.45%
Lebanon	0.03%	1.73%	0.03%	1.73%
Philippines	0.03%	1.73%	0.03%	1.73%
Thailand	0.03%	1.73%	0.03%	1.73%
China, incl. Tibet	0.03%	1.73%	0.03%	1.73%
Asia (other)	0.18%	4.24%	0.17%	4.12%
Oceania	0.01%	1.00%	0.01%	1.00%
Missing	0.04%	2.00%	0.04%	2.00%
<i>School education and vocational training</i>				
Secondary school w/o completed vocational training	15.52%	36.21%	16.65%	37.25%
Secondary school with completed vocational training	58.22%	49.32%	62.47%	48.42%
Upper secondary school (general/subject-specific aptitude)	1.66%	12.78%	1.78%	13.22%

for higher education) w/o completed vocational training				
Upper secondary school (general/subject-specific aptitude				
for higher education) with completed vocational training	3.79%	19.10%	4.06%	19.74%
Completion of a university of applied sciences	2.78%	16.44%		
College / university degree	4.02%	19.64%		
Missing	14.01%	34.71%	15.04%	35.75%
<i>Year of first employment</i>				
1975 or earlier	23.78%	42.57%	24.58%	43.06%
1976	3.13%	17.41%	3.15%	17.47%
1977	2.53%	15.70%	2.49%	15.58%
1978	2.41%	15.34%	2.37%	15.21%
1979	2.56%	15.79%	2.50%	15.61%
1980	2.58%	15.85%	2.54%	15.73%
1981	2.45%	15.46%	2.41%	15.34%
1982	2.26%	14.86%	2.22%	14.73%
1983	2.29%	14.96%	2.25%	14.83%
1984	2.52%	15.67%	2.46%	15.49%
1985	2.59%	15.88%	2.52%	15.67%
1986	2.84%	16.61%	2.76%	16.38%
1987	2.84%	16.61%	2.76%	16.38%
1988	3.03%	17.14%	2.93%	16.86%
1989	3.66%	18.78%	3.58%	18.58%
1990	4.88%	21.54%	4.80%	21.38%
1991	2.11%	14.37%	2.18%	14.60%
1992	1.78%	13.22%	1.84%	13.44%
1993	2.26%	14.86%	2.26%	14.86%
1994	2.17%	14.57%	2.17%	14.57%
1995	2.22%	14.73%	2.23%	14.77%
1996	2.07%	14.24%	2.05%	14.17%
1997	2.19%	14.64%	2.12%	14.41%
1998	2.18%	14.60%	2.09%	14.30%
1999	5.19%	22.18%	5.14%	22.08%
2000	2.87%	16.70%	2.85%	16.64%
2001	2.19%	14.64%	2.20%	14.67%
2002	1.61%	12.59%	1.63%	12.66%
2003	1.28%	11.24%	1.31%	11.37%
2004	0.99%	9.90%	1.01%	10.00%
2005	0.84%	9.13%	0.86%	9.23%
2006	0.78%	8.80%	0.80%	8.91%
2007	0.64%	7.97%	0.65%	8.04%
2008	0.30%	5.47%	0.30%	5.47%
Age at first employment	24.09	7.64	23.98	7.72

Note: Weighted statistics calculated from the IAB LIAB MM 9308.

Table A4: Summary Statistics for Time Invariant Match Characteristics

	Full Sample		Low Educ. Sample	
	Mean	SD	Mean	SD
Part time job (at beginning of match)	22.98%	42.07%	23.60%	42.46%
<i>Employment Status 8 days before current match</i>				
No previous record	17.75%	38.21%	18.18%	38.57%
Unknown, previous spell not benefits	16.12%	36.77%	15.70%	36.38%
Unknown, previous spell was benefit spell	3.93%	19.43%	3.96%	19.50%
Benefit receipt	19.77%	39.83%	20.26%	40.19%
Employment at other firm	40.22%	49.03%	39.60%	48.91%
Apprentice/trainee at other firm	2.19%	14.64%	2.28%	14.93%
Missing	0.02%	1.41%	0.02%	1.41%
<i>Employment status 8 days before current match, condensed</i>				
No previous record	17.75%	38.21%	18.18%	38.57%
Benefits/gap	39.82%	48.95%	39.92%	48.97%
Employment at other firm	40.22%	49.03%	39.60%	48.91%
Apprentice/trainee at other firm	2.19%	14.64%	2.28%	14.93%
Missing	0.02%	1.41%	0.02%	1.41%
Number of days in emp. status 8 days before current match	1117.21	1635.52	1109.38	1650.03
<i>Year match started</i>				
1993	18.35%	38.71%	18.69%	38.98%
1994	5.15%	22.10%	5.21%	22.22%
1995	3.11%	17.36%	3.12%	17.39%
1996	3.59%	18.60%	3.61%	18.65%
1997	2.56%	15.79%	2.57%	15.82%
1998	5.21%	22.22%	5.20%	22.20%
1999	5.95%	23.66%	5.98%	23.71%
2000	9.14%	28.82%	9.23%	28.94%
2001	7.41%	26.19%	7.36%	26.11%
2002	6.06%	23.86%	6.02%	23.79%
2003	5.26%	22.32%	5.23%	22.26%
2004	5.59%	22.97%	5.50%	22.80%
2005	5.47%	22.74%	5.40%	22.60%
2006	5.47%	22.74%	5.34%	22.48%
2007	5.66%	23.11%	5.60%	22.99%
2008	6.02%	23.79%	5.95%	23.66%
Days of benefit receipt up to beginning of current match	234.37	488.28	243.15	498.10
Days since first employment at beginning of current match	3127.01	3112.95	3117.67	3140.79
Age at beginning of current match	37.43	12.11	37.36	12.27
<i>Match count</i>				
1	93.31%	24.98%	93.93%	23.88%
2	6.36%	24.40%	5.80%	23.37%
3	0.32%	5.65%	0.26%	5.09%
4	0.01%	1.00%	0.01%	1.00%
5	/	/	/	/
6	/	/	/	/
Female	42.24%	49.39%	43.12%	49.52%

Note: Weighted statistics calculated from the IAB LIAB MM 9308. If one or more cells contained too few observations, at least two cell frequencies could not be disclosed (to prevent calculation from totals). This is indicated by /.

Appendix F: Results

Table A5: Regressions of log of Daily Wage on Time Variant Characteristics, All Coefficients

	Full Sample, Match	Full Sample, TWFE	Reduced Sample, Match	Reduced Sample, TWFE
Total number of employees	0.000004* (0.000001)	0.000004** (0.000001)	0.000003* (0.000001)	0.000003* (0.000001)
<i>Business volume, categorical</i>				
72,000 to 120,000	0.0196 (0.0158)	0.0204 (0.017)	0.0042 (0.0115)	0.0016 (0.0118)
120,000 to 166,200	0.0221 (0.0157)	0.0288 (0.0171)	0.0049 (0.0127)	0.0064 (0.013)
166,200 to 245,400	-0.0036 (0.0147)	-0.0002 (0.0162)	-0.0277* (0.0108)	-0.0319** (0.0114)
245,400 to 332,300	0.0063 (0.0148)	0.0094 (0.0163)	-0.0175 (0.0107)	-0.0196 (0.0111)
332,300 to 490,000	-0.0023 (0.0149)	0.0009 (0.0163)	-0.0301** (0.0107)	-0.0316** (0.0111)
490,000 to 715,800	0.0053 (0.0141)	0.0091 (0.0156)	-0.0158 (0.0093)	-0.0177 (0.0097)
715,800 to 1,227,100	0.0116 (0.014)	0.0141 (0.0155)	-0.0083 (0.0091)	-0.0111 (0.0095)
1,227,100 to 3,163,900	0.0131 (0.0139)	0.0149 (0.0154)	-0.0067 (0.009)	-0.0098 (0.0093)
more than 3,163,900	0.0124 (0.0139)	0.0116 (0.0154)	-0.0053 (0.0089)	-0.009 (0.0093)
Missing	0.019 (0.0138)	0.0235 (0.0152)	-0.0033 (0.0087)	-0.0056 (0.0091)
<i>Business volume per employee, categorical</i>				
21,300 to 30,700	0.0054* (0.0022)	0.0055* (0.0026)	0.0067** (0.0021)	0.0068** (0.0023)
30,700 to 39,800	0.0026 (0.0023)	0.0024 (0.0028)	0.0041 (0.0023)	0.0049 (0.0025)
39,800 to 50,000	0.0038 (0.0024)	0.0051 (0.0028)	0.0052* (0.0022)	0.006* (0.0025)
50,000 to 59,700	0.0016 (0.0024)	0.0004 (0.0029)	0.0006 (0.0023)	0.0009 (0.0025)
59,700 to 71,600	0.0019 (0.0025)	0.0021 (0.003)	0.0009 (0.0024)	0.0016 (0.0026)
71,600 to 92,900	-0.0011 (0.0025)	0.0011 (0.0029)	-0.0041 (0.0024)	-0.0032 (0.0027)
92,900 to 128,600	0.0006 (0.0024)	0.0031 (0.0029)	-0.0029 (0.0023)	-0.0025 (0.0026)
128,600 to 230,100	0.0027 (0.0024)	0.0059* (0.0029)	-0.0008 (0.0023)	-0.0004 (0.0026)
more than 230,100	0.0105*** (0.0024)	0.0143*** (0.0029)	0.007** (0.0023)	0.0076** (0.0026)
Fraction of employees working part time	-0.0034 (0.004)	-0.0011 (0.0044)	-0.0029 (0.0041)	-0.0002 (0.0043)
<i>Investment per employee, categorical</i>				
2,100 to 3,000	0.002 (0.0012)	0.0016 (0.0013)	0.001 (0.0013)	0.001 (0.0013)
3,000 to 4,200	0.0017 (0.0009)	0.0025** (0.0009)	0.0007 (0.0009)	0.0012 (0.0009)
4,200 to 6,100	0.0031*** (0.001)	0.0038*** (0.001)	0.0027** (0.001)	0.003** (0.001)
6,100 to 10,000	0.0041*** (0.0009)	0.0048*** (0.0009)	0.0034*** (0.0009)	0.0039*** (0.0009)

10,000 to 18,600	0.0057*** (0.001)	0.0066*** (0.001)	0.0048*** (0.001)	0.0051*** (0.001)
more than 18,600	0.0058*** (0.0009)	0.0068*** (0.001)	0.0055*** (0.0009)	0.006*** (0.001)
Missing	0.0039*** (0.001)	0.0047*** (0.001)	0.0036*** (0.001)	0.0043*** (0.001)
Fraction of female employees	-0.0056 (0.0036)	-0.0031 (0.0039)	-0.0103** (0.0035)	-0.0091* (0.0036)
DHS employment growth index	0.0082*** (0.002)	0.0098*** (0.002)	0.0083*** (0.002)	0.0093*** (0.0021)
More employees than previous year	0.0037*** (0.0004)	0.0042*** (0.0005)	0.0037*** (0.0005)	0.0036*** (0.0005)
<i>Wanted to hire people, but did not</i>				
Wanted to hire people, but did not	-0.0002 (0.0011)	0.0005 (0.0012)	0.0001 (0.0012)	0.0004 (0.0012)
Missing	-0.0002 (0.0022)	0.0008 (0.0023)	0.001 (0.0022)	0.0015 (0.0023)
<i>Expected business volume relative to last year</i>				
Increasing	0.0013** (0.0005)	0.0015** (0.0005)	0.0012* (0.0005)	0.0013* (0.0005)
Decreasing	-0.0044*** (0.0005)	-0.0048*** (0.0006)	-0.0045*** (0.0006)	-0.0047*** (0.0006)
Missing	-0.0009 (0.001)	-0.0013 (0.001)	-0.0007 (0.0011)	-0.0008 (0.0011)
Total number of new employees	0.000013* (0.000006)	0.000012* (0.000005)	0.00002*** (0.000005)	0.000019*** (0.000005)
Firm was hiring in current year	0.0048*** (0.0006)	0.0057*** (0.0006)	0.0046*** (0.0006)	0.005*** (0.0006)
Total number of employees that left	-0.000005 (0.000003)	-0.000007* (0.000003)	-0.000005 (0.000003)	-0.000006 (0.000003)
Employees have left in current year	0.0021*** (0.0006)	0.0028*** (0.0006)	0.002** (0.0006)	0.0022*** (0.0006)
Number of days in current establishment	0.000027*** (0.000005)	0.000018*** (0.000001)	0.000042*** (0.000006)	0.000021*** (0.000001)
Age at end of year, squared	-0.002209*** (0.000128)	-0.003944*** (0.000139)	-0.00285*** (0.000123)	-0.004645*** (0.000132)
Age at end of year^3	0.000039*** (0.000002)	0.000063*** (0.000002)	0.000048*** (0.000002)	0.000075*** (0.000002)
Age at end of year^4	-0.00000027*** (0.00000001)	-0.00000039*** (0.00000001)	-0.00000032*** (0.00000001)	-0.00000047*** (0.00000001)
Part time job	-0.3395*** (0.0046)	-0.4446*** (0.0059)	-0.3227*** (0.0042)	-0.3922*** (0.0046)
Daily wage topcoded	0.0185*** (0.0012)	0.0231*** (0.0013)	0.0247*** (0.0012)	0.032*** (0.0013)
<i>Year</i>				
1994	-0.0351*** (0.0044)	0.0757*** (0.0044)	-0.0303*** (0.0044)	0.0636*** (0.0043)
1995	-0.0562*** (0.0082)	0.1658*** (0.0082)	-0.0429*** (0.0083)	0.1446*** (0.008)
1996	-0.0969*** (0.0122)	0.2365*** (0.0122)	-0.0777*** (0.0123)	0.204*** (0.0119)
1997	-0.1403*** (0.0162)	0.3022*** (0.0162)	-0.1148*** (0.0163)	0.2599*** (0.0158)
1998	-0.183*** (0.0202)	0.369*** (0.0202)	-0.1505*** (0.0203)	0.3173*** (0.0197)
1999	-0.2154*** (0.0243)	0.4417*** (0.0242)	-0.176*** (0.0244)	0.382*** (0.0237)
2000	-0.2521***	0.5178***	-0.2063***	0.4464***

	(0.0283)	(0.0283)	(0.0284)	(0.0276)
2001	-0.2868***	0.5946***	-0.2347***	0.5114***
	(0.0324)	(0.0323)	(0.0325)	(0.0316)
2002	-0.3231***	0.6707***	-0.2643***	0.576***
	(0.0364)	(0.0363)	(0.0365)	(0.0355)
2003	-0.3506***	0.7543***	-0.2893***	0.6444***
	(0.0404)	(0.0404)	(0.0405)	(0.0394)
2004	-0.3986***	0.8177***	-0.3307***	0.6965***
	(0.0445)	(0.0444)	(0.0446)	(0.0433)
2005	-0.4413***	0.8871***	-0.3673***	0.7535***
	(0.0485)	(0.0485)	(0.0487)	(0.0473)
2006	-0.4837***	0.9556***	-0.4028***	0.8112***
	(0.0525)	(0.0525)	(0.0527)	(0.0512)
2007	-0.5172***	1.034***	-0.4293***	0.8781***
	(0.0566)	(0.0566)	(0.0567)	(0.0551)
2008	-0.5461***	1.1169***	-0.4525***	0.9483***
	(0.0607)	(0.0606)	(0.0609)	(0.0592)
Experience in years	0.1579***	0.1171***	0.1525***	0.1193***
	(0.0026)	(0.0014)	(0.003)	(0.0014)
Experience in years^2	-0.003233***	-0.004723***	-0.002067***	-0.00233***
	(0.000109)	(0.000152)	(0.000056)	(0.000059)
Experience in years^3	0.000097***	0.00016***	0.000057***	0.000069***
	(0.000004)	(0.000006)	(0.000002)	(0.000002)
Experience in years^4	-0.000001***	-0.000002***	-0.00000055***	-0.00000072***
	(0.00000005)	(0.00000008)	(0.00000003)	(0.00000003)
Number of observations	9792405	9792405	8693593	8693593
Number of individuals	3068373	3068373	2726651	2726651
Number of firms	24323	24323	23348	23348
Number of matches	3413921		2996587	
Total Sum of Squares	4104476	4104476	3588938	3588938
Residual sum of squares	175455	244647	151418	195440
R2	0.9573	0.9404	0.9578	0.9455
F-stat of all coefficients	1042.7	1037.6	979.3	1001.5
p-value	0	0	0	0
F-stat of all fixed effects	19.1	15.2	19.2	16.4
p-value	0	0	0	0
F-stat individual FE		11.7		12.4
p-value		0		0
F-stat firm FE		52.2		51.3
p-value		0		0

Omitted Categories: *Business volume, categorical: 0 to 72,000; Business volume per employee, categorical: 0 to 21,300; Investment per employee, categorical: 0 to 500; Wanted to hire people, but did not: successfully hired or did not want to; Expected business volume relative to last year: same; Year: 1993*

Notes: Standard errors are clustered at the firm and individual level. Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

Table A6: Regressions of Firm Fixed Effects on Time Invariant Firm Characteristics, All Coefficients

<i>Industry</i>	Full Sample, Match	Full Sample, TWFE	Reduced Sample, Match	Reduced Sample, TWFE
Agriculture and forestry	0.0097 (0.1517)	0.023 (0.1482)	-0.0178 (0.1427)	-0.0186 (0.143)
Mining, quarrying and electricity	-0.0222 (0.1826)	-0.0342 (0.1815)	-0.0444 (0.2213)	-0.0518 (0.2182)
Food products	0.0111 (0.083)	0.0005 (0.0827)	-0.02 (0.0813)	-0.0335 (0.0824)
Clothing and textile	0.0923 (0.1003)	0.0581 (0.0996)	-0.015 (0.1051)	-0.0203 (0.1033)
Paper and printing	0.1463 (0.1497)	0.1269 (0.1483)	0.0751 (0.1531)	0.0694 (0.1545)
Wood products, furniture, jewellery and toys	0.353*** (0.0914)	0.3296*** (0.0925)	0.3384*** (0.0951)	0.3285*** (0.0971)
Chemical industry	0.2447* (0.1203)	0.2355* (0.1073)	0.2226 (0.1395)	0.2382 (0.1223)
Rubber/plastic	0.2319** (0.0757)	0.2242** (0.072)	0.1788** (0.0683)	0.1766** (0.0678)
Non-metallic mineral products	0.2292* (0.0967)	0.2386* (0.0955)	0.2078* (0.0938)	0.2189* (0.0921)
Basic metals, steel, light metal	0.2922*** (0.0633)	0.2799*** (0.0606)	0.2444*** (0.0638)	0.2463*** (0.0627)
Recycling	-0.6061 (0.8601)	-0.6308 (0.8529)	-0.6119 (0.8785)	-0.6555 (0.8869)
Machinery	0.3259*** (0.0892)	0.2903** (0.0884)	0.3036** (0.0924)	0.2811** (0.0922)
Motor vehicles: production/sales/repair/fuel	-0.0523 (0.0753)	-0.0459 (0.0744)	-0.0614 (0.0704)	-0.0521 (0.0716)
Other transport equipment	0.2208 (0.1267)	0.182 (0.1282)	0.1701 (0.1212)	0.1473 (0.1212)
Electrical equipment	0.1582 (0.0913)	0.1857 (0.0955)	0.0964 (0.0932)	0.1217 (0.0958)
Precision and optical equipment	0.1238 (0.0948)	0.1145 (0.0944)	0.112 (0.0947)	0.0992 (0.0963)
Main construction trade	0.2746*** (0.0668)	0.2828*** (0.0656)	0.241*** (0.0683)	0.2514*** (0.0679)
Building installation/completion	0.2129*** (0.0643)	0.2182*** (0.0629)	0.1888** (0.062)	0.2021** (0.0623)
Transportation	0.0196 (0.074)	-0.0077 (0.0724)	-0.0017 (0.0744)	-0.026 (0.0739)
Communication	-0.2788 (0.2347)	-0.2588 (0.2263)	-0.2873 (0.2201)	-0.2823 (0.2184)
Credit and financial intermediation	0.0569 (0.0974)	0.0585 (0.0928)	0.0972 (0.1087)	0.1043 (0.1055)
Insurance	0.001 (0.1713)	-0.0476 (0.1684)	-0.0059 (0.1704)	-0.0456 (0.1693)
Computer and related activities	0.3042* (0.1259)	0.3454** (0.1125)	0.2972* (0.1422)	0.3314* (0.1317)
Research and development	0.219** (0.0814)	0.2032* (0.0808)	0.2881** (0.098)	0.2671** (0.0995)
Legal consulting, advertising	-0.0672 (0.1026)	-0.0626 (0.1011)	-0.1207 (0.1166)	-0.1145 (0.1171)
Real estate	-0.211 (0.1292)	-0.1993 (0.123)	-0.2738 (0.1506)	-0.2701 (0.146)
Renting, business activities	0.0377 (0.0702)	0.0197 (0.068)	-0.0408 (0.0769)	-0.0619 (0.0753)

Hotel and restaurant	-0.3776*** (0.0794)	-0.3595*** (0.0771)	-0.395*** (0.0794)	-0.3912*** (0.0785)
Education/teaching	0.0995 (0.0987)	0.0811 (0.092)	0.0988 (0.1012)	0.0912 (0.0976)
Human health, veterinary and social work	0.0285 (0.0652)	-0.0026 (0.0639)	-0.0264 (0.0655)	-0.0511 (0.0653)
Sanitation	0.1615 (0.149)	0.1295 (0.1394)	0.0359 (0.114)	0.0058 (0.1112)
Recreation, culture, sports	-0.4495** (0.1677)	-0.451** (0.1604)	-0.5665*** (0.1683)	-0.5782*** (0.1665)
Other services	-0.1998* (0.1)	-0.1856 (0.1002)	-0.2485* (0.1035)	-0.2425* (0.1054)
Organizations, lobbying, etc.	-0.395** (0.1442)	-0.4172** (0.138)	-0.3177* (0.1532)	-0.3513* (0.1502)
Public administration and social security	0.0378 (0.0903)	0.0195 (0.087)	0.0152 (0.0932)	0.0101 (0.0924)
<i>State</i>				
Schleswig-Holstein	0.149* (0.0675)	0.1147 (0.0664)	0.1704** (0.0658)	0.1416* (0.0655)
Hamburg	0.1895** (0.0712)	0.178** (0.0688)	0.238*** (0.0718)	0.2336** (0.0713)
Lower Saxony	-0.0497 (0.0619)	-0.0524 (0.0597)	-0.0351 (0.0634)	-0.0369 (0.0623)
Bremen	0.2206*** (0.0632)	0.2132*** (0.062)	0.1776** (0.0677)	0.1825** (0.0685)
Hesse	0.0594 (0.0604)	0.0603 (0.0584)	0.0747 (0.0592)	0.0763 (0.0589)
Rhineland-Palatinate	-0.1463* (0.068)	-0.1623* (0.0685)	-0.0919 (0.0695)	-0.1106 (0.0705)
Baden-Wuerttemberg	0.0046 (0.054)	0.0037 (0.0524)	0.0325 (0.0557)	0.0311 (0.0552)
Bavaria	0.1849*** (0.0547)	0.1706** (0.0534)	0.1601** (0.0563)	0.1568** (0.0562)
Saarland	0.0867 (0.0581)	0.055 (0.0549)	0.1081 (0.0576)	0.0843 (0.0562)
Berlin	0.058 (0.0644)	0.0528 (0.063)	0.1185 (0.066)	0.1158 (0.0651)
Brandenburg	-0.0266 (0.0809)	-0.0053 (0.0774)	-0.0566 (0.0841)	-0.0311 (0.0816)
Mecklenburg-Western Pomerania	0.091 (0.0699)	0.1104 (0.0673)	0.0955 (0.0722)	0.1178 (0.0706)
Saxony	0.0005 (0.0788)	0.0344 (0.0763)	-0.0351 (0.0796)	-0.0012 (0.0779)
Saxony-Anhalt	-0.0365 (0.0818)	-0.0033 (0.0785)	-0.0969 (0.087)	-0.0626 (0.0846)
Thuringia	0.0183 (0.0819)	0.0484 (0.079)	-0.0191 (0.0839)	0.0198 (0.0821)
<i>Legal Form</i>				
Individually-owned firm	-0.3585*** (0.0399)	-0.3462*** (0.0389)	-0.3453*** (0.0405)	-0.3427*** (0.0403)
Partnership	-0.1603*** (0.0479)	-0.1479** (0.047)	-0.1735*** (0.0466)	-0.1624*** (0.0467)
Company limited by shares	0.0728 (0.0619)	0.0734 (0.0641)	0.0785 (0.0663)	0.0847 (0.0688)
Public corporation	-0.058 (0.0795)	-0.0635 (0.076)	-0.0412 (0.0797)	-0.0532 (0.0782)
Other legal form	-0.1069 (0.0733)	-0.101 (0.0709)	-0.137 (0.0742)	-0.1381 (0.0737)

Missing/don't know	0.0281 (0.2088)	-0.0125 (0.2011)	0.0068 (0.2251)	-0.0313 (0.2162)
<i>Main/Exclusive Ownership</i>				
Eastern German property	0.1201* (0.0573)	0.1185* (0.0561)	0.1106 (0.0592)	0.115* (0.0579)
Foreign property	0.2526*** (0.0756)	0.2415** (0.0746)	0.244** (0.0822)	0.2363** (0.082)
Public property	0.0829 (0.0675)	0.0951 (0.0643)	0.0697 (0.0698)	0.0805 (0.0693)
No principal shareholder	0.1496 (0.0881)	0.1461 (0.083)	0.1283 (0.0791)	0.1381 (0.0763)
Unknown	-0.0961 (0.0985)	-0.0976 (0.0983)	-0.0652 (0.1053)	-0.0669 (0.1064)
Missing	0.3172*** (0.088)	0.3807*** (0.0781)	0.3054*** (0.0921)	0.3725*** (0.0855)
<i>Year founded (only after 1990)</i>				
1990	0.0111 (0.0715)	0.0061 (0.0686)	-0.0022 (0.0726)	-0.0067 (0.0711)
1991	0.049 (0.0951)	0.0421 (0.0882)	0.0478 (0.0992)	0.0389 (0.0935)
1992	-0.1807 (0.1104)	-0.1885 (0.1096)	-0.1615 (0.1151)	-0.1759 (0.1152)
1993	-0.0171 (0.0772)	-0.0352 (0.0742)	-0.033 (0.0764)	-0.0477 (0.0754)
1994	-0.0531 (0.1086)	-0.0516 (0.1056)	-0.0436 (0.1102)	-0.0445 (0.1082)
1995	0.018 (0.0865)	0.0042 (0.0861)	-0.0064 (0.0858)	-0.014 (0.0869)
1996	-0.055 (0.0816)	-0.0564 (0.0762)	-0.0257 (0.0875)	-0.0332 (0.0843)
1997	-0.145 (0.0945)	-0.1476 (0.0869)	-0.1792 (0.0982)	-0.1891* (0.0932)
1998	-0.0483 (0.1126)	-0.0556 (0.1087)	-0.1158 (0.1125)	-0.1226 (0.1093)
1999	-0.0718 (0.0927)	-0.1115 (0.0898)	-0.0365 (0.0891)	-0.0723 (0.0883)
2000	0.1873 (0.1023)	0.1291 (0.1014)	0.1945 (0.1072)	0.1491 (0.1079)
2001	-0.2087 (0.1475)	-0.2301 (0.1432)	-0.1473 (0.1592)	-0.1783 (0.1572)
2002	-0.0678 (0.1128)	-0.102 (0.1117)	0.0169 (0.1112)	-0.0221 (0.1109)
2003	-0.0607 (0.1189)	-0.1207 (0.1123)	-0.0099 (0.1207)	-0.0661 (0.1151)
2004	-0.0577 (0.1302)	-0.1243 (0.1292)	0.048 (0.1156)	-0.0155 (0.1121)
2005	0.0726 (0.1008)	-0.0112 (0.0975)	0.1659 (0.1046)	0.0829 (0.1032)
2006	0.3205** (0.1172)	0.2202 (0.1155)	0.3881** (0.118)	0.2903* (0.1178)
2007	0.238 (0.1364)	0.1263 (0.1349)	0.3013* (0.1381)	0.1975 (0.1341)
2008	-0.1132 (0.0928)	-0.2126* (0.089)	-0.0995 (0.0942)	-0.2126* (0.0922)
Missing	-0.0309 (0.0907)	-0.0165 (0.0816)	-0.0823 (0.0946)	-0.0574 (0.0888)
<i>Establishment/Department is...</i>				
Place of business/office/branch	0.0365	0.0436	0.0306	0.0344

	(0.0444)	(0.0424)	(0.0469)	(0.0459)
Head office	0.0713	0.0655	0.0537	0.0572
	(0.0561)	(0.0537)	(0.0573)	(0.0563)
Middle-level authority	0.0597	0.0667	0.0159	0.0143
	(0.0829)	(0.0775)	(0.079)	(0.0799)
Missing	0.0412	0.1051	0.1107	0.1502
	(0.1655)	(0.1531)	(0.1867)	(0.1764)
<i>Company pays for job training/courses</i>				
Yes	-0.119***	-0.1113**	-0.1187***	-0.1077**
	(0.0355)	(0.0345)	(0.0359)	(0.0356)
Missing	-0.1222*	-0.1007	-0.1208*	-0.1087
	(0.0574)	(0.0562)	(0.0583)	(0.0583)
<i>Has Worker's Council</i>				
Yes	0.1215***	0.1385***	0.1149**	0.1313***
	(0.0359)	(0.0347)	(0.0381)	(0.0376)
Missing	-0.1565	-0.1277	-0.1649	-0.1466
	(0.1848)	(0.174)	(0.1985)	(0.1904)
<i>Collective Wage Agreement</i>				
Company agreement	-0.0448	-0.0388	-0.0594	-0.0505
	(0.0726)	(0.0701)	(0.0759)	(0.0748)
No collective agreement	-0.1009**	-0.1187***	-0.1074**	-0.1261***
	(0.0346)	(0.0338)	(0.0351)	(0.0349)
Missing	-0.0453	-0.0317	-0.0542	-0.0325
	(0.074)	(0.0728)	(0.0777)	(0.0778)
<i>Owner working in Company</i>				
No	0.06	0.0538	0.0644	0.0606
	(0.0391)	(0.0375)	(0.0406)	(0.0398)
Missing	-0.3195*	-0.3009*	-0.0562	-0.0688
	(0.1469)	(0.1293)	(0.0985)	(0.0962)
Constant	4.0106***	5.344***	4.3716***	5.6099***
	(0.0581)	(0.0566)	(0.0593)	(0.0588)
Number of Observations	24291	24291	23318	23318
R-squared	0.1619	0.1814	0.1567	0.1760

Omitted Categories: *Industry:* Sales: Retail and Wholesale; *State:* North Rhine-Westphalia; *Legal Form:* Limited liability company; *Main/Exclusive Ownership:* Western German property; *Year founded:* Founded before 1990; *Establishment/Department is...:* independent company/organisation w/o other places of business; *Company pays for job training/courses:* No; *Has Worker's Council:* No; *Collective Wage Agreement:* Industry-wide wage agreement; *Owner working in Company:* Yes;

Notes: Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

Table A7: Regressions of Individual FE on Time Invariant Firm Characteristics, All Coefficients

	Full Sample, Match	Full Sample, TWFE	Reduced Sample, Match	Reduced Sample, TWFE
Female	-0.0559*** (0.0032)	-0.1035*** (0.0029)	-0.0514*** (0.0034)	-0.0797*** (0.0031)
<i>Nationality, grouped</i>				
Turkey	-0.0288*** (0.0072)	-0.0073 (0.0067)	-0.0334*** (0.0073)	-0.0213** (0.0068)
Italy	-0.0128 (0.0123)	-0.0192 (0.0108)	-0.0244 (0.0128)	-0.0318** (0.0113)
Yugoslavia, Serbia and Montenegro	-0.0465*** (0.0126)	-0.0183 (0.0118)	-0.0566*** (0.0127)	-0.0425*** (0.0122)
Greece	0.0152 (0.0187)	0.0097 (0.0171)	0.0086 (0.0199)	0.0022 (0.0187)
France	0.0855*** (0.0256)	0.0798*** (0.0208)	0.0967** (0.0309)	0.0865** (0.0265)
Poland	0.1499*** (0.0293)	0.1187*** (0.0254)	0.1539*** (0.0304)	0.1308*** (0.0279)
Austria	0.0062 (0.0279)	0.0301 (0.0261)	0.0061 (0.033)	0.0134 (0.0309)
Croatia	0.0279 (0.0208)	0.0415* (0.0184)	0.0118 (0.022)	0.0231 (0.0201)
Portugal	-0.0852*** (0.0229)	-0.0354 (0.0217)	-0.1062*** (0.0238)	-0.0716** (0.0228)
Spain	-0.1055*** (0.025)	-0.0778*** (0.0213)	-0.1202*** (0.0278)	-0.1023*** (0.0249)
Netherlands, Luxembourg	0.218** (0.0341)	0.2009*** (0.0298)	0.2042*** (0.0377)	0.1873*** (0.0346)
Russia, Belarus, Former Soviet Union	0.1182*** (0.0336)	0.0955** (0.0311)	0.1167** (0.036)	0.1004** (0.0349)
Bosnia and Herzegovina	0.0722 (0.041)	0.0501 (0.0328)	0.0692 (0.0435)	0.0498 (0.0374)
Great Britain, Ireland and Northern Ireland	0.0759* (0.0328)	0.0788* (0.0309)	0.1212*** (0.0284)	0.1164*** (0.026)
Romania	0.1272** (0.0449)	0.1034* (0.0414)	0.091 (0.0475)	0.0697 (0.0448)
Czech Republic, Slovakia, Former Czechoslovakia	0.2572*** (0.0702)	0.2676*** (0.0785)	0.2795*** (0.0799)	0.2852** (0.088)
Ukraine, Moldova	0.1916 (0.1666)	0.1719 (0.1622)	0.1703 (0.1486)	0.1554 (0.1425)
Hungary	-0.0843 (0.167)	-0.1087 (0.1742)	-0.0532 (0.1284)	-0.0719 (0.1387)
Albania	0.2473** (0.0916)	0.2286* (0.0917)	0.1797** (0.0557)	0.169** (0.0523)
Belgium	0.1845*** (0.0387)	0.1981*** (0.0337)	0.1896*** (0.0468)	0.1946*** (0.0409)
Macedonia	0.0632 (0.0771)	0.0456 (0.0739)	0.0454 (0.0806)	0.0369 (0.0776)
Switzerland	0.1497*** (0.0413)	0.1257*** (0.0349)	0.1325** (0.0507)	0.1139** (0.0435)
Bulgaria	0.0117 (0.1148)	0.0523 (0.1197)	-0.2161** (0.0792)	-0.196* (0.0765)
Slovenia	-0.0439 (0.0681)	-0.0243 (0.0603)	-0.0263 (0.0582)	-0.0213 (0.054)
Denmark, Sweden	0.232*** (0.0592)	0.2214*** (0.0454)	0.2315*** (0.0687)	0.2196*** (0.0547)
Finland	0.0323 (0.0895)	0.0075 (0.0669)	0.0794 (0.0966)	0.0309 (0.0734)

Estonia, Latvia, Lithuania	0.0977 (0.0683)	0.1153 (0.0626)	0.0797 (0.0787)	0.1001 (0.0715)
Europe (other)	0.5538*** (0.131)	0.3933** (0.1387)	0.5417*** (0.1446)	0.4083** (0.1343)
Ethiopia	0.043 (0.0689)	0.0284 (0.0737)	-0.001 (0.073)	-0.0112 (0.0754)
Ghana	0.0987 (0.0538)	0.0615 (0.0563)	0.0945* (0.0447)	0.0749 (0.0501)
Morocco	-0.0627* (0.0298)	-0.0562* (0.0281)	-0.0909** (0.029)	-0.0849** (0.0274)
Tunisia	0.0908 (0.0503)	0.0432 (0.0421)	0.0847 (0.0551)	0.0521 (0.0474)
Africa (other)	0.1093** (0.0343)	0.0857** (0.0322)	0.0914** (0.0349)	0.075* (0.0331)
USA, Canada	0.1646*** (0.0336)	0.1415*** (0.0321)	0.14*** (0.0357)	0.1071*** (0.0301)
America (other)	-0.0066 (0.0605)	-0.0235 (0.0572)	-0.0192 (0.076)	-0.0333 (0.0734)
Afghanistan	0.1216 (0.0726)	0.131* (0.0585)	0.0636 (0.07)	0.0743 (0.0582)
Sri Lanka	0.2474*** (0.0476)	0.2144*** (0.0457)	0.1069*** (0.0293)	0.0946*** (0.0275)
Vietnam	-0.0471 (0.06)	-0.052 (0.0622)	-0.0366 (0.0584)	-0.0392 (0.0592)
India	0.0325 (0.0837)	0.0355 (0.0875)	0.0141 (0.0786)	0.0231 (0.0822)
Iraq	0.3026*** (0.0668)	0.2805*** (0.0646)	0.2552*** (0.0632)	0.2463*** (0.0633)
Iran	0.1635** (0.0569)	0.1134* (0.0511)	0.1383* (0.0679)	0.1049 (0.0623)
Lebanon	0.0698 (0.0759)	0.0195 (0.0611)	0.0811 (0.082)	0.0434 (0.0685)
Philippines	-0.0187 (0.0545)	-0.0015 (0.0538)	-0.0475 (0.0587)	-0.0334 (0.0572)
Thailand	-0.1378 (0.1041)	-0.0994 (0.1005)	-0.163 (0.105)	-0.1431 (0.1047)
China, incl. Tibet	-0.1812* (0.0779)	-0.1856** (0.0684)	-0.3155** (0.1146)	-0.3139** (0.1029)
Asia (other)	0.1114** (0.0405)	0.0963* (0.0386)	0.0842* (0.0429)	0.0735 (0.0421)
Oceania	0.2028** (0.0661)	0.1453* (0.0595)	0.1891* (0.0766)	0.1519* (0.0713)
Missing	-0.0722 (0.1032)	-0.0951 (0.0739)	-0.0829 (0.1092)	-0.1017 (0.0883)
<i>School education and vocational training</i>				
Secondary / intermediate school w/o completed vocational training	-0.2325*** (0.0042)	-0.2605*** (0.0038)	-0.2233*** (0.0042)	-0.2434*** (0.0039)
Upper secondary school w/o completed vocational training	-0.3391*** (0.0152)	-0.3914*** (0.0136)	-0.3507*** (0.0167)	-0.3909*** (0.0158)
Upper secondary school with completed vocational training	0.2338*** (0.007)	0.1849*** (0.0065)	0.2385*** (0.007)	0.2019*** (0.0067)
Completion of a university of applied sciences	0.3861*** (0.007)	0.3147*** (0.0064)		
College / university degree	0.5251*** (0.006)	0.4205*** (0.0053)		
Missing	-0.0248*** (0.0061)	-0.1286*** (0.0056)	-0.0143* (0.0062)	-0.0903*** (0.0058)
<i>Year of first employment</i>				

1976	0.1695*** (0.0096)	-0.0001 (0.0078)	0.1761*** (0.0107)	0.0425*** (0.0092)
1977	0.2175*** (0.0096)	-0.0577*** (0.0081)	0.2127*** (0.0104)	-0.0067 (0.0092)
1978	0.2692*** (0.0122)	-0.1032*** (0.0105)	0.2501*** (0.0133)	-0.0486*** (0.0119)
1979	0.2881*** (0.0111)	-0.1924*** (0.0094)	0.2506*** (0.0121)	-0.1355*** (0.0108)
1980	0.3088*** (0.0111)	-0.2679*** (0.0097)	0.2658*** (0.012)	-0.2011*** (0.011)
1981	0.3446*** (0.0093)	-0.3277*** (0.0083)	0.288*** (0.0097)	-0.2553*** (0.0088)
1982	0.3539*** (0.0099)	-0.4104*** (0.009)	0.2737*** (0.0106)	-0.3467*** (0.0098)
1983	0.3846*** (0.0103)	-0.4804*** (0.0093)	0.2946*** (0.0113)	-0.4088*** (0.0105)
1984	0.4202*** (0.0105)	-0.5414*** (0.0092)	0.3155*** (0.0123)	-0.4678*** (0.0114)
1985	0.4444*** (0.01)	-0.615*** (0.0094)	0.3359*** (0.0109)	-0.529*** (0.0104)
1986	0.4523*** (0.0097)	-0.7006*** (0.0091)	0.3266*** (0.0106)	-0.6165*** (0.0101)
1987	0.5063*** (0.0086)	-0.7474*** (0.0077)	0.369*** (0.0093)	-0.657*** (0.0086)
1988	0.5408*** (0.0087)	-0.8142*** (0.0081)	0.3885*** (0.0091)	-0.721*** (0.0086)
1989	0.579*** (0.0077)	-0.8784*** (0.0071)	0.4203*** (0.0082)	-0.7748*** (0.0078)
1990	0.6343*** (0.0072)	-0.9326*** (0.0066)	0.4754*** (0.0075)	-0.8104*** (0.007)
1991	0.6729*** (0.0095)	-1.0128*** (0.0088)	0.503*** (0.0098)	-0.8778*** (0.0093)
1992	0.6705*** (0.0112)	-1.0868*** (0.0105)	0.4827*** (0.0117)	-0.9605*** (0.0112)
1993	0.7112*** (0.01)	-1.134*** (0.0094)	0.5086*** (0.0102)	-1.0113*** (0.0098)
1994	0.7472*** (0.0088)	-1.1981*** (0.0081)	0.5335*** (0.0094)	-1.0692*** (0.0087)
1995	0.772*** (0.0095)	-1.2728*** (0.009)	0.5552*** (0.0097)	-1.1314*** (0.0093)
1996	0.8177*** (0.0094)	-1.3224*** (0.0089)	0.5755*** (0.0099)	-1.1892*** (0.0096)
1997	0.8261*** (0.0108)	-1.4073*** (0.0103)	0.5841*** (0.011)	-1.2603*** (0.0107)
1998	0.8605*** (0.0099)	-1.4658*** (0.0093)	0.605*** (0.0108)	-1.3163*** (0.0104)
1999	0.6538*** (0.0087)	-1.7615*** (0.0083)	0.3744*** (0.0093)	-1.6194*** (0.009)
2000	0.7696*** (0.011)	-1.7423*** (0.0105)	0.4832*** (0.0117)	-1.592*** (0.0113)
2001	0.7744*** (0.0106)	-1.829*** (0.01)	0.4731*** (0.0106)	-1.6797*** (0.0102)
2002	0.7907*** (0.0127)	-1.9012*** (0.0121)	0.4934*** (0.0128)	-1.7337*** (0.0123)
2003	0.7585*** (0.0155)	-2.0238*** (0.0151)	0.4462*** (0.0154)	-1.8594*** (0.0151)
2004	0.7291*** (0.0148)	-2.1341*** (0.0141)	0.3967*** (0.0155)	-1.9774*** (0.0149)

2005	0.7176*** (0.0166)	-2.2329*** (0.0159)	0.3691*** (0.0177)	-2.0809*** (0.0171)
2006	0.7364*** (0.0187)	-2.3*** (0.0171)	0.3755*** (0.0197)	-2.146*** (0.0185)
2007	0.7236*** (0.02)	-2.4004*** (0.0185)	0.3523*** (0.0219)	-2.2461*** (0.0205)
2008	0.755*** (0.0274)	-2.4621*** (0.0261)	0.3536*** (0.0276)	-2.3267*** (0.0264)
Age at first employment	0.0455*** (0.0003)	0.1075*** (0.0002)	0.0644*** (0.0003)	0.1208*** (0.0003)
Constant	-1.4123*** (0.0078)	-1.7287*** (0.0071)	-1.652*** (0.0081)	-2.0611*** (0.0076)
Number of observations	3062118	3062118	2720888	2720888
R-squared	0.2509	0.7443	0.27981496	0.737457

Omitted Categories: *Nationality, grouped:* Germany; *School education and vocational training:* Secondary/intermediate school with completed vocational training; *Year of first employment:* 1975 or earlier;

Notes: Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.

Table A8: Regression of Match and Firm Fixed Effects on Pre-Match Characteristics, All Coefficients

	Match Effect	Firm Effect
Part time job (at beginning of match)	-0.1187*** (0.0047)	-0.6257*** (0.0151)
<i>Employment status 8 days before current match</i>		
No previous record	-0.04*** (0.0066)	-0.0796*** (0.0176)
Previous spell was benefits or gap	-0.0012 (0.0039)	-0.057*** (0.0107)
Apprentice/trainee at other firm	0.0074 (0.0155)	-0.0164 (0.0488)
Number of Days in Labor Market Status 8 Days before current match (main effect)	-0.000001 (0.000001)	0.000009*** (0.000002)
...if previous spell was benefits or gap (interaction)	0.000001 (0.000003)	0.000003 (0.000008)
...if previous spell was training (interaction)	0.000016 (0.000019)	0.000009 (0.000058)
<i>Year Match Started</i>		
1994	0.0024 (0.0059)	0.0315 (0.0172)
1995	0.0115 (0.0079)	-0.0106 (0.0243)
1996	0.0145* (0.0071)	0.0104 (0.0173)
1997	0.0126 (0.0081)	0.0105 (0.0236)
1998	0.0217** (0.0067)	-0.015 (0.0168)
1999	-0.0181 (0.0104)	-0.1885*** (0.029)
2000	-0.0035 (0.0071)	-0.0955*** (0.0189)
2001	0.0138* (0.0063)	-0.0263 (0.0163)
2002	0.0189** (0.0059)	-0.0035 (0.0166)
2003	0.0243*** (0.0069)	-0.1059*** (0.0213)
2004	0.0223*** (0.0059)	-0.0391 (0.0267)
2005	0.026*** (0.0062)	-0.0119 (0.0189)
2006	0.0353*** (0.0063)	-0.0048 (0.0194)
2007	0.0436*** (0.006)	0.0327 (0.0185)
2008	0.0561*** (0.0065)	0.025 (0.0188)
Female	-0.0529 (0.0433)	-0.376*** (0.1055)
Number of days of benefit receipt up to beginning of current match	0.000007* (0.000003)	-0.000082*** (0.000009)
Years since first employment at beginning of current match	0.0024** (0.0009)	-0.0046* (0.0022)
...interacted with female dummy	-0.0024 (0.0014)	-0.0132*** (0.0036)
Years since first employment squared	-0.000129*** (0.00003)	-0.000058 (0.000077)

...interacted with female dummy	0.000082 (0.000048)	0.000457*** (0.000131)
Age at beginning of current match	0.0024 (0.0015)	0.0364*** (0.0038)
...interacted with female dummy	0.0043 (0.0025)	0.0221*** (0.0059)
Age squared	-0.00002 (0.000019)	-0.000387*** (0.000048)
<i>Match Count</i>		
2	-0.0063 (0.0043)	0.0612*** (0.0117)
...interacted with female dummy	0.0071 (0.0065)	0.0027 (0.0207)
3	-0.0229** (0.0087)	0.0958*** (0.0236)
...interacted with female dummy	0.0208 (0.0142)	-0.0253 (0.052)
4	-0.0576* (0.0252)	0.1722*** (0.0401)
...interacted with female dummy	0.0807 (0.0562)	0.1501 (0.1395)
5	0.1193*** (0.0188)	-0.05 (0.0573)
...interacted with female dummy	-0.0815 (0.0757)	0.1983 (0.1426)
6	-0.2742*** (0.0055)	0.2965*** (0.0263)
...interacted with female dummy	0.0918*** (0.0086)	-0.3977*** (0.0315)
Constant	-0.0565* (0.0258)	3.7472*** (0.0691)
Number of observations	665080	665080
R-squared	0.0343	0.2189

Omitted Categories: *Emp. Status 8 days before current match:* Employment at other Firm; *Year Match Started:* 1993; *Match Count:* 1

Note: Estimates of Match and Firm Fixed Effects are from main regression on full sample. Significance levels: *: Significant at 5%; **: Significant at 1%; ***: Significant at 0.1%.