INNODRIVE WORKING PAPER SERIES

Worker and Firm Heterogeneity

Identifying the Return to Education in a Wage Equation

Øivind A. Nilsen,
Arvid Raknerud and
Terje Skjerpen

March 2011

INNODRIVE Working Paper No 19.

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n° 214576

Worker and Firm Heterogeneity Identifying the Return to Education in a Wage Equation[1]

Øivind A. Nilsen
Arvid Raknerud
Terje Skjerpen

Abstract: There is a long tradition in economics for estimating wage equations at the individual level. Traditional explanatory variables applied are education length and (potential) experience. With linked employer-employee data we have access to firm-level information in addition to individual level information. However, even when being in such a favourable data situation, there will still be factors that are unobservable. Since we, in contrast to Abowd et al. (1999), also include time-invariant variables, such as years of education (which are time invariant in our sample), and are occupied with return to education, we find it natural to work within a framework with random effects. In this paper we consider panel data from the Norwegian machinery industry (NACE 29) covering the period 1995-2006. The results indicate that taking account of unobserved firm-specific heterogeneity to a minor extent influences the estimates of educational premium in the labour market.

---

[1] This document has been made in conjunction with the project *INNODRIVE –Intangible Capital and Innovations*: *Drivers of Growth and Location in EU*. Grants from the seventh European framework program are highly appreciated (Grant no. 214576).

The affiliations of the first author, Øivind A. Nilsen, are Norwegian School of Economics and Business Administration, and IZA. The affiliation of the two other authors is Statistics Norway.

Correspondence: Terje Skjerpen, Statistics Norway, Research Department, PO Box 8131, N-0033 Oslo, Norway. E-mail: terje.skjerpen@ssb.no

1. Introduction

Access to employer-employee data makes it possible to take account of both individual and firm level information when estimating wage equations. Utilizing only individual data may lead to estimation bias or at least loss of efficiency which may produce misleading policy conclusions. Abowd *et al*. (1999) constitute a seminal contribution with respect to wage modelling using employer-employee data. They represented both unobserved individual- and firm-specific heterogeneity by fixed effects. However, the fixed effects specification has problematic features. Abowd *et al*. (1999) do only include two-dimensional observed covariates. However, when quantifying wage equations one is also interested in the effect of observed individual-specific variables, i.e. one-dimensional variables, or in variables that almost may be regarded as such a variable. An example is the length of education which for most of the workers does not vary. The effect of a change in education is identified in the random effects model, but not in the fixed effects model.[1] Another advantage with the random components model is that it is far more parsimonious with respect to the number of parameters. On the other hand a problem with the random effects model is that when the random effect is correlated with the observed right-hand side variables, biased estimation is a consequence.

Several authors have been occupied with the so-called sorting mechanism, cf. for instance Eckhout and Kircher (2010). Do high-skilled workers end up in high- or low-productivity firms? This issue may also be addressed within a random effects set-up. Utilizing the estimated second order moments of the random effects and the data one may predict all

[1] There may also be intermediate cases in a situation with several covariates when it is possible to identify the effect of unidimensional variables even in the presence of fixed effects. However, this requires an a priori assumption stating that some of the covariates are uncorrelated with the random unobserved individual-specific term.

the individual- and firm-specific random variables. Based on these predicted values one can then look at the correlation between predicted individual- and firm-specific variables.

In this paper we consider different specifications. Among these are a model with both unobserved individual- and firm-specific random effects, which we also allow to be correlated conditional on the outcome of a matching variable.[2] To estimate the parameters in the model we apply the Helmert-transformation on the wage equation in order to sweep out the individual-specific random effects. After having carried out the Helmert-transformation we specify the log-likelihood function utilizing a state space form. To maximize the log-likelihood we propose a quasi-Newton algorithm. We also show that the fixed effects estimators are obtained as limiting cases when the variance of the random effects becomes arbitrary large. We are not aware of any analysis that has utilized the Helmert transformation in a panel data context with employer-employee data so this seems to be a novelty of the current paper.

We also comment on what is the best specification of unobserved heterogeneity in wage equations when one has access to unbalanced employer-employee panel data. What one ultimately seeks for is a test corresponding to the standard Hausman test applied in panel data models where one only addresses one-way unobserved heterogeneity. Also hybrid cases are of some interest. To distinguish between the different specifications that are applied we introduce the following abbreviations: FN, RN, RR, FF and RF, where N stands for no heterogeneity and F and R stand for fixed and random, respectively. The first capital letter states the specification in the individual dimension, whereas the second one states the specification in the firm dimension. The two first models are the standard specifications,

[2] However, the specification may be said to somewhat asymmetric in that whereas we allow for the influences of individual specific observed variables we do not add observed firm-specific variables. Some contributions in the literature which estimate wage equations on employer-employee data have allowed for such effects, for an example cf. the analyses by Lallemand *et al*. (2005), Plasman *et al*. (2007) and Heyman (2007).

assuming fixed or random individual heterogeneity (and no firm heterogeneity), whereas the last three account for unobserved heterogeneity in different ways in both dimensions.

We apply our modelling framework to a sample of individuals working in a traditional manufacturing industry, production of machinery (NACE 29).[3] Panel employer-employee data for Norway for the period 1995−2006 are utilized. The data material used consists of 16,198 observations. We have 2,191 individuals and 751 firms. As observed covariates in the wage-equation we use length of education, a fourth order polynomial in experience, three dummies for type of education, a dummy for gender, five dummies for labour market areas and eleven year dummies. Of the skill-related variables only those involving experience vary both across individuals and time.

For the specifications involving random effects, we are interested in the magnitude of the relative variances. What is the ratio between the variance of the individual random effects and the variance of the genuine error terms (which may be viewed as a kind of signal to noise ratio) in the two cases, that is (i) with random firm-specific effects and (ii) without random firm-specific effects. Another issue is whether the estimated correlation between the individual- and firm-specific effects is zero in the RR model. This latter issue may be addressed by performing a likelihood-ratio test.

Using the RR specification the individual-specific variation turns out to be more important than the firm-specific variation. The estimated variance ratio is about 3. With respect to slope coefficients they differ only modestly between the RR specification and the RN specification. This is not very surprising. If the RR specification is valid, we know that the covariance matrix of the gross error terms, which are made up by the sum of two one-dimensional random terms and a genuine error term, will have a certain structure. Accounting

[3] NACE Rev. 1.

for this structure is necessary for obtaining efficient estimates of the slope parameters of the wage equation, but is not necessary for obtaining consistent estimates of these parameters. Furthermore, the estimate of the correlation coefficient between the individual- and firm-specific random effects is not statistically different from zero. Constraining this parameter to zero does not produce estimated slope parameters that are very different from those obtained in the specification where it is allowed to be estimated as a free parameter. However, as emphasized by Eeckhout and Kircher (2010) it is not straightforward to interpret such an empirical finding.

The rest of the paper is organized in the following way: In Section 2 we outline the modelling framework and comments on the different specifications. The focus of Section 3 is how the Helmert transformation can be utilised to simplify estimation. Section 4 contains a description of the data used in the subsequent empirical analysis. Here, we put some emphasis on describing how frequent job transitions are for the individuals over the years in which they are observed. Such transitions are of substantial importance when it comes to identifying unobserved firm effects, no matter whether they are viewed as fixed or random. In Section 5 we report the results of the empirical analysis. Some concluding remarks are offered in Section 6.

2. Modelling framework

The general wage equation may be written as

(1) $\log(W_{ijt}) = \alpha + X_{1,i}\beta_1 + X_{2,it}\beta_2 + X_{3,it}\beta_3 + X_{4,t}\beta_4 + \mu_i + \nu_j + \eta_{ijt}$,

where $X_{k,it}$ (k = 1,…,4) are (row) vectors of observed variables. The variables in these four vectors are respectively, (i) individual time-invariant attributes, (ii) individual time-varying attributes, (iii) labour market area dummies and (iv) year dummies. More specifically we have

$X_{1,i}$ = {years of schooling, type of education dummies, gender},

$X_{2,it}$ = {powers of experience up to the fourth order}.

The left hand side variable of Eq. (19 has three subscripts. Subscript *i* represents individual, subscript *j* firm and subscript *t* time. The symbol μ_i represents unobserved individual-specific heterogeneity, whereas the symbol ν_j represents unobserved firm-specific heterogeneity. At the end of the equation we have the genuine error term, η_{ijt}. We assume throughout that $E\eta_{ijt} = 0$ and $E\eta_{ijt}^2 = \sigma_{\eta\eta}^2$ $\forall i,\ j,\ t$. As outlined below, this general specification may be specialized in different ways.

In Table 1 we present five different specifications of Eq. (1). The two first specifications are benchmark specifications. In the first specification (FN) we consider fixed individual-specific effects, but neglect unobserved firm heterogeneity. For this model the vector of time-invariant observed variables, $X_{1,i}$, is omitted because of lack of identification. As we will elaborate more on later we have also omitted the vector with year dummies, $X_{4,t}$. The second specification (RN) has random individual-specific effects and no unobserved firm-specific heterogeneity. All the four vectors with variables are retained in this case.

In the third specification (FF) we assume fixed effects in both the individual and firm dimension.[4] As for the FN specification the variable vectors $X_{1,i}$ and $X_{4,t}$ are dropped from the regressions. In the fourth specification (RR) we assume random effects in both the two dimensions. Furthermore, we also allow for correlation between the two types of random effects. In the literature some authors have been occupied with the so-called sorting mechanism, i.e., whether high-productivity workers end up in high-productivity firms. After having estimated the unknown parameters of the RR specification it is possible to predict the

[4] This is the specification considered in the seminal paper by Abowd *et al.* (1999). However, these authors seem to disregard uni-dimensional variables from the outset in their analysis.

values of the two latent components. In a way such predicted values correspond to the estimated fixed effects in the FF specification. So both the fixed and random effects models may be applied for addressing sorting issues.

The fifth specification (RF) constitutes a hybrid case assuming random individual-specific effects and fixed firm-specific effects. Whereas the models FN, RN, FF and RF are estimated in Stata, RR is estimated using a code written in Gauss.[5]

3. The Helmert transformation – a way to "integrate out" the random person effects

In this paper we use the Helmert transformation to integrate out – or sweep out – the random individual effects from models with both individual and firm effects, while preserving orthogonality of the additive error terms. Formally, the Helmert-transformations, $\vec{y}_{i;t}$, of a time series y_{it}, $t = 1,...,T_i$, are defined as $\vec{y}_{i;t} = \sqrt{t/(t+1)}\left(y_{i,t+1} - t^{-1}\sum_{s=1}^{t} y_{i,s} \right)$, $t = 1,...,T-1$ and $\overline{y}_i = T_i^{-1}\sum_{s=1}^{T_i} y_{is}$. It is easy to check that the corresponding Helmert transformed error terms, $\vec{\varepsilon}_{i;t}$, are uncorrelated, given that the orginal error terms, ε_{it}, are uncorrelated and homoscedastic (constant variance over time). Moreover, the individual effects will be swept out from all the transformed variables, except $\overline{y}_i$. Of course, the Helmert transformation is not the only way of sweeping out the individual effects (see e.g. Andrews *et al*., 2008, for a discussion of the within estimator in this context), but it has the huge advantage of preserving the orthogonality of the error terms. To examine the fixed effects estimator in relation to the

[5] The FF model is estimated by utilizing the Stata command felsdvreg, cf. Cornelissen (2008). For an alternative algorithm cf. Guimarães and Portugal (2010).

above framework, we apply the (simplified) notation

$y_{it} = x_{it}\beta + \mu_i + G_{it}v + \varepsilon_{it}$,

where $v = (v_1,...,v_M)'$ is the vector of all the *M* firm-effects and G_{it} is an appropriate selection vector of zeros and ones, such that $G_{it}v = v_{i(t)}$ is the firm effect corresponding to firm $i(t)$, where $i(t)$ is the stochastic index denoting individual *i*'s firm at time *t*: $i(t) \in \{1,...,M\}$. To simplify the notation, we assume that all individuals enter the sample at *t=1*. We allow for unbalanced data, with individual *i* exiting the sample at $t = T_i$.

Let us first consider the fixed effects estimator, which is obtained by minimizing the quadratic form:

$$Q(v,\mu;\beta) = \sum_{i=1}^{N} T_i(\overline{y}_i - \overline{x}_i\beta - \mu_i - \overline{G}_i v)^2 + (\vec{y} - \vec{X}\beta - \vec{G}v)'(\vec{y} - \vec{X}\beta - \vec{G}v),$$

where

$$\overline{y}_i = T_i^{-1}\sum_t y_{it}$$

$$\overline{x}_i = T_i^{-1}\sum_t x_{it}$$

$$\overline{G}_i = T_i^{-1}\sum_t G_{it}.$$

Moreover, using the Helmert transformation defined above, $\vec{y} = (\vec{y}_1',...,\vec{y}_N')'$, with $\vec{y}_i = (\vec{y}_{i;1},...,\vec{y}_{i;T_i-1})'$, and $\vec{G} = (\vec{G}_1',...,\vec{G}_N')'$ with $\vec{G}_i = (\vec{G}_{i;1}',...,\vec{G}_{i;T_i-1}')'$. The first order conditions then become

$$\overline{y}_i - \overline{x}_i\beta - \mu_i - \overline{G}_i v = 0$$

$$\sum_{i=1}^{N} T_i \overline{G}_i'(\overline{y}_i - \overline{x}_i\beta - \mu_i - \overline{G}_i v) + \vec{G}'(\vec{y} - \vec{X}\beta - \vec{G}v) = 0$$

$$\sum_{i=1}^{N} T_i \ \overline{x_i}'(\overline{y_i} - \overline{x_i}\beta - \mu_i - \overline{G_i}v) + \overrightarrow{X}'(\overrightarrow{y} - \overrightarrow{X}\beta - \overrightarrow{G}v) = 0 \,.$$

It is worthwhile noticing that finding the fixed effects estimator by a standard equation solver, is much easier when applied to the system above than to the untransformed system studied in Abowd *et al.* (2002). The reason is that the untransformed system involves *N*+*M*+*q* (where q is the number of elements in β) non-trivial equations, while our approach only requires solving a system of *M*+*q* equations (with *M* much less than *N*), since each μ_i is trivially obtained after having solved the system with respect to (β, v).

Now assume that μ_i is random: $\mu \sim IN(0, \sigma_\varepsilon^{\ 2}\Sigma)$, where $\mu = (\mu_1, ..., \mu_N)'$. Then the first equation in the system must be modified:

$$\overline{y_i} - \overline{x_i}\beta - E(\mu_i \mid y) - \overline{G_i}v = 0$$

where

$$E(\mu \mid y) = \sigma_\varepsilon^{\ 2}(\Sigma^{-1} + I)^{-1}(\overline{y} \ - \overline{x}\,\beta - \overline{G}\ v)$$

with $\overline{y} = (\overline{y_1}, ..., \overline{y_N})$ and $(\overline{x}, \overline{G})$ defined similarly (by stacking $\overline{\mathrm{x_i}}$ and $\overline{G_i}$). Σ may be estimated from

$$Var(\mu \mid y) = \sigma_\varepsilon^{\ 2}(\Sigma^{-1} + I)^{-1} \,.$$

It is clear from the above relations, that the fixed effects estimator is a limiting case of the random effects estimator when $\Sigma^{-1} \to 0$ (which can be interpreted as assuming a “diffuse” prior for the random effects). See Francke *et al.* (2010) for more details about the relation between the fixed effects and random effects estimator (within a state space model framework).

4. Data issues

The initial sample includes 241,904 observations, for 53,665 individuals. The sample covers the period 1995−2006 and is collected for individuals and firms in the Norwegian machinery industry (NACE 29). Totally there are 2,593 firms in the initial sample. There are some individuals whose educational length changes over the sample period. For these individuals we keep only the observations where the educational length is the same as their maximum length. Furthermore, we require that the registered (potential) experience, should increase with one year from year *t* to year *t+1*. We also exclude individuals that change gender. This is data flaw. We include only individuals whose annual earning is between 50,000 and 3,500,000 NOK (fixed prices) per year. It is also required that each individual should have two or more observations after the already mentioned exclusion criteria are applied. After the described data clean the sample includes 201,833 observations, 36,183 individuals and 2,178 firms over the period 1995−2006.

Since we are focusing on models with both individual-specific and firm-specific unobserved effects, we focus on a sample of individuals observed in at least two different firms over their sample period. This is necessary to be able to identify both of the two unobserved effects. We are focusing on this sample also for the pure individual fixed effects and the random effects models, i.e., where we do not control for firm-specific effects. The decision of only to focus on these individuals is based on the fact that we want to avoid that the sample varies from one model to the next, i.e., that variation in the analysed samples could drive the potential differences in the estimated sets of coefficients. There are 9,400 individuals, with a total of 70,509 observations, that change firms over their sample period, and thus help in identifying the fixed firm and fixed individual effects. Finally, from the described data cleaned sample we are drawing randomly observations for 2,191 individuals to speed up estimation. This is especially important for the model in which the Helmert transformation model is applied.

Tables 2 and 3 below provide some information about the unbalanced panel dataset. The individuals are observed from minimum 2 to maximum 12 years. In average there are 7.6 observations per observational unit. We consider only individuals who move from one employer to another at least 1 time during the period they are in the sample. As is seen from Table 2 1,308 individuals are observed for all the 12 years. They account for a little more than a fifth of the observations. In order to identify firm effects it is of vital importance that the individuals change employers. Table 3 shows that 1,860 out of the individuals change employers only one time during the period they are in the data set. None of the individuals change employer more than three times. Only 39 individuals change employer three times during the period they are in the sample.

5. Empirical results

Table 4 contains estimation results of the wage equation under different assumptions with respect to the treatment of unobserved individual- and firm-specific heterogeneity. Five different specifications are considered. In the first specification we are looking at the FN specification, whereas the second specification is the RN specification. These first two wage-equation specifications may be viewed as benchmark models. As the third, fourth and fifth specifications we have the FF, RR and RF specifications, respectively. As is well known the effects of one-dimensional variables are not identified in fixed effects panel data models, which is the reason why some of the cells in Table 4 are left empty. Of the variables related to individual attributes, experience is the only one which has two-dimensional variation. In our data education (EDU) is an individual-specific variable. The experience variable (EXP) is operationalised as potential experience. It is defined as the worker's age minus his number of years of education minus his age when he started at school. Thus, in our situation EXP becomes an individual-specific trend variable. By reshuffling terms one may convert the individual trend variables to a trend that is common to all the individuals and modify the fixed effects accordingly. Such a transformation eludicates that it is hard to separate the effects of the year dummies from the effect of the EXP variable. Hence in light of this, the former group of variables is not included when estimating the models with fixed effects in the individual dimension. When we, on the other hand, are estimating the models with random effects in the individual dimension none of the variables have been excluded from the regressions.

Looking at the estimates in Table 4, one sees that they are mainly of the same sign and not very different in magnitude. In the case without firm heterogeneity one may conduct the standard Hausman test to test the random effects model against the fixed effects model. To this aim one utilizes the deviation in estimated slope coefficients attached to the experience variables and the labour market area variables. The Hausman test statistic, χ^2=16,353.5 (p=0.000), shows that the hypothesis of independence between the unobserved individual specific term and the regressors is firmly rejected. If one takes a closer look at the estimated fourth order polynomial of experience, the functions differ between the two models, cf. Figures 1 and 2. For the RN model the maximum effect of experience is achieved at somewhat less than 30 years, whereas in the FN model the effect of experience on the real hourly wage rate is increasing for any level of the experience variable.[6] The graphs related to RR and RF are qualitatively equal to the one for the RN model and the graph related to the FF model is qualitatively equal to the one of the FN model.

[6] As mentioned earlier in the FN model we have omitted the year dummies in order to avoid singularity problems. Thus the parameters involved in the fourth order polynomial of experience may be interpreted as a type of gross parameters that besides representing the effect of the experience variable also pick up calendar effects. Thus one may question the validity of the Hausman-procedure in this case.

The returns to education are estimated to 7, 6.6 and 6.3 per cent for the RN, RR and RF models, respectively. All the three dummies representing education type are significant at the 5 per cent significance level in the RN and RR models, whereas this is the case for only the two first models in the RF model. Having a "Business and administration" type of education yields a real wage amendment of about 10 per cent according to the RN and RR models and about 6 per cent according to the RF model, as compared to the reference group.[7] The estimates of the coefficients attached to the year dummies (not displayed) show a trending pattern over the years, which may be associated with the effect of productivity growth on real wages.

Let us then turn to the specification with fixed effects in both dimensions (FF). The hypothesis of no firm-effects is heavily rejected. Also in this case we carry out a type of Hausman test by confronting the fixed effects model without firm heterogeneity (FN) with the specification containing fixed effects in both dimensions (FF). The specification neglecting unobserved firm heterogeneity is firmly rejected as χ^2=39.2807 (p=0.000).

6. Concluding remarks

In this paper we have considered different specifications of real wage equations when one has access to employer-employee panel data. Such data enable controlling for both unobserved individual- and firm-specific effects. Earlier contributions in this area, of those Abowd *et al.* (1999) being the most known and cited, have stuck to a specification with both fixed individual- and firm-specific effects. However, a feature of such a specification is that one cannot identify effects on the real wage of one-dimensional variables such as for instance education, type of education and gender. Thus, it might be worthwhile also to consider random effects specifications, which we have done in this paper. Even if we, in our application, have to reject the random effects models when they are formally tested against

[7] The reference category consists of the education types 1–3 and 6–9, cf. Table A1.

the fixed effects models, we find that most of the estimates of the slope parameters that are common to both models are rather equal.

References

Abowd, J.M., Creecy, R.H. and F. Kramarz (2002): Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Mimeo March 2002.

Abowd, J.M., Kramarz, F. and D.N. Margolis (1999): High Wage Workers and High Wage Firms Wages. *Econometrica*, 67(2), 251–333.

Andrews. M.J., Gill, L., Schank, T. and R. Upward (2008): High Wage Workers and Low Wage Firms: Negative Assortative Matching and Limited Mobility Bias? *Journal of the Royal Statistical Society, Series A*, 171, 673–697.

Cornelissen, T. (2008): The Stata Command felsdvreg to fit a linear model with two high-dimensional fixed effects. *Stata Journal*, 8(2), 170–189.

Eeckhout, J. and P. Kircher (2010): Identifying Sorting – in Theory. LSE Research Online.

Francke, M.K., Koopman, S.J. and A.F. De Vos (2010): Likelihood Functions for State Space Models with Diffuse Initial Conditions. *Journal of Time Series Analysis*, 31, 407–414.

Guimarães, P. and P. Portugal (2010): A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects. *Stata Journal*, 10(4), 628–649.

Heyman, F. (2007): Firm Size or Firm Age? The Effect On Wages using Matched Employer-Employee data. *Labour*, 21(2), 237–263.

Lallemand, T., Plasman, R. and F. Rycx (2005): Why do Large Firms Pay Higher Wages? Evidence from Matched Worker-Firm Data. *International Journal of Manpower*, 26(7-8), 705–723.

Plasman, R., Rycx, F. and I. Tojerow (2007): Wage differentials in Belgium: The Role of Worker and Employer Characteristics. *Cahiers Economiques de Bruxelles*, 50(1), 11–40.

Table 1. An overview of different models with respect to the treatment of unobserved individual and firm heterogeneity

Model name	Assumption with respect to unobserved individual heterogeneity	Assumption with respect to unobserved firm heterogeneity	Slope parameter restrictions
FN	$\mu_i = \mu_i^F$	$\nu_j = 0, \forall j$	$\beta_1 = 0, \beta_4 = 0$
RN	$\mu_i = \mu_i^R$	$\nu_j = 0, \forall j$	
FF	$\mu_i = \mu_i^F$	$\nu_j = \nu_j^F$	$\beta_1 = 0, \beta_4 = 0$
RR	$\mu_i = \mu_i^R$	$\nu_j = \nu_j^R$	
RF	$\mu_i = \mu_i^R$	$\nu_j = \nu_j^F$	

Notes: μ_i^F and ν_i^F denote, respectively, a fixed individual-specific and a fixed firm-specific value. μ_i^R and ν_j^R denote, respectively, a random individual-specific and a random firm-specific effect. It is assumed that $E\mu_i^R = 0$, $E\left(\mu_i^R\right)^2 = \sigma_{\mu\mu}^2 \, \forall i$ and that $E\nu_j^R = 0$, $E\left(\nu_j^R\right)^2 = \sigma_{\nu\nu}^2 \, \forall j$. In the RR specification we allow for correlation between the random individual-specific and the random firm-specific effects and denote it by $\rho_{\mu\nu}$.

Table 2. Summary statistics. The unbalancedness of the panel data

Number of years a person is in the sample	Number of persons being in the sample for the indicated number of years	Number of observations
2	107	214
3	177	531
4	195	780
5	211	1,055
6	204	1,224
7	177	1,239
8	170	1,360
9	177	1,593
10	189	1,890
11	276	3,036
12	308	3,696
Sum	2,191	16,618

Table 3. Summary statistics. Individual mobility

Number of years a person is in the sample	Number of persons that have moved the indicated number of time			Sum
	1 move	2 moves	3 moves	
2	107	0	0	107
3	164	13	0	177
4	174	21	0	195
5	188	22	1	211
6	180	20	4	204
7	146	30	1	177
8	137	31	2	170
9	144	26	7	177
10	154	32	3	189
11	230	40	6	276
12	236	57	15	308
Sum	1,860	292	39	2,191

Table 4. Wage estimation results under different assumptions with respect to individual and firm heterogeneity. Standard errors in parentheses

Explanatory variable	Fixed individual effects. No firm effects (FN)	Random individ. effects. No firm effects (RN)	Fixed individual effects. Fixed firm effects (FF)	Random individ. effects. Random firm effects (RR)	Random individ. effects. Fixed firm effects (RF)
Length of education		0.0695***		0.0655***	0.0632***
		(0.0026)		(0.0023)	(0.0026)
Experience	0.0659***	0.0434***	0.0689***	0.0422***	0.0427***
	(0.0050)	(0.0047)	(0.0055)	(0.0044)	(0.0049)
$(\text{Experience}/10)^2$	–0.1613***	–0.1622***	–0.1653***	–0.1578***	–0.1567***
	(0.0387)	(0.0357)	(0.0409)	(0.0331)	(0.0368)
$\text{Experience}^3/10^4$	0.2702**	0.2689**	0.2702**	0.2621***	0.2484**
	(0.1147)	(0.1054)	(0.1198)	(0.0977)	(0.1081)
$\text{Experience}^4/10^6$	–0.1826	–0.1772*	–0.1782	–0.1756*	–0.1547
	(0.1143)	(0.1051)	(0.1185)	(0.0973)	(0.1073)
Dummy for male		0.1793***		0.1837***	0.1859***
		(0.0178)		(0.0157)	(0.0171)
Dummy for edu. type 0		0.1116***		0.1051***	0.0791***
		(0.0273)		(0.0241)	(0.0265)
Dummy for edu. type 4		0.0991***		0.0961***	0.0616**
		(0.0300)		(0.0264)	(0.0287)
Dummy for edu. type 5		0.0509**		0.0459**	0.0217
		(0.0249)		(0.0219)	(0.0240)
$N^{-1}\sum_{i=1}^{N}\left(\mu_i^F\right)^2$	0.4620				
$\sigma^2_{\mu\mu}$		0.2313		0.0401	0.2061
$\sigma^2_{\nu\nu}$				0.0027	
$\rho_{\mu\nu}$				0.1006	
$\sigma^2_{\eta\eta}$	0.2267	0.2263		0.0431	0.2141
Number of observations	16,618	16,618	16,618	16,618	16,618

Notes: The models with random individual effects also include regional variables, dummy variables for years and a constant term. The FF model contains regional variables and the FN model contains regional variables and a constant term. The estimates of the parameters attached to these variables are not reported in this table. The model RF has been estimated in Stata by using the standard procedure for one-dimensional random effects models. The fixed firm effects are represented by firm dummies.
"education type 0" = "General programs",
"education type 4" = "Business and administration", "education type 5" = "Natural Sciences, Vocational and Technical subjects".
*** means significant at the 1 per cent significance level according to the standard normal distribution.
** means significant at the 5 per cent significance level according to the standard normal distribution.
* means significant at the 10 per cent significance level according to the standard normal distribution.

Table A1. Types of education in Norwegian classification of education

Education category	Description
0	General Programs
1	Humanities and Arts
2	Teacher, Training and Pedagogy
3	Social Sciences and Law
4	Business and Administration
5	Natural Sciences, Vocational and Technical subjects
6	Health, Welfare and Sport
7	Primary Industries
8	Transport and Communications, Safety and Secuity
9	Unspecified

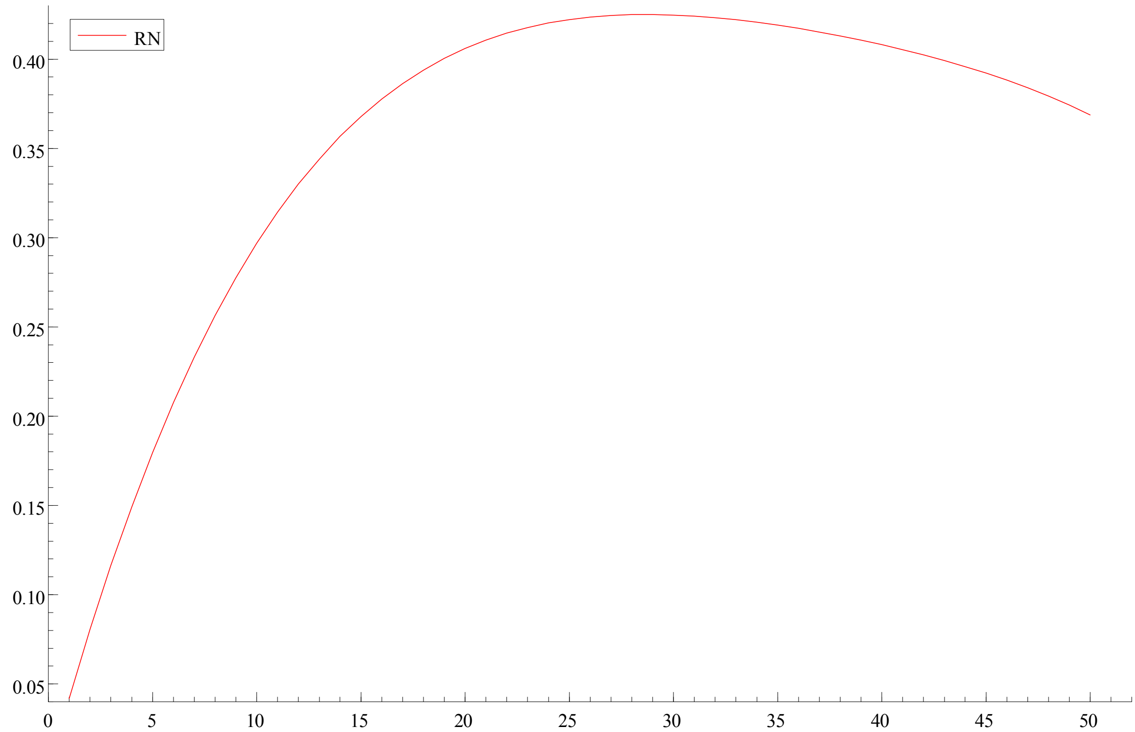

Figure 1. The total effect of the fourth order experience polynomial in the real wage equation. RN

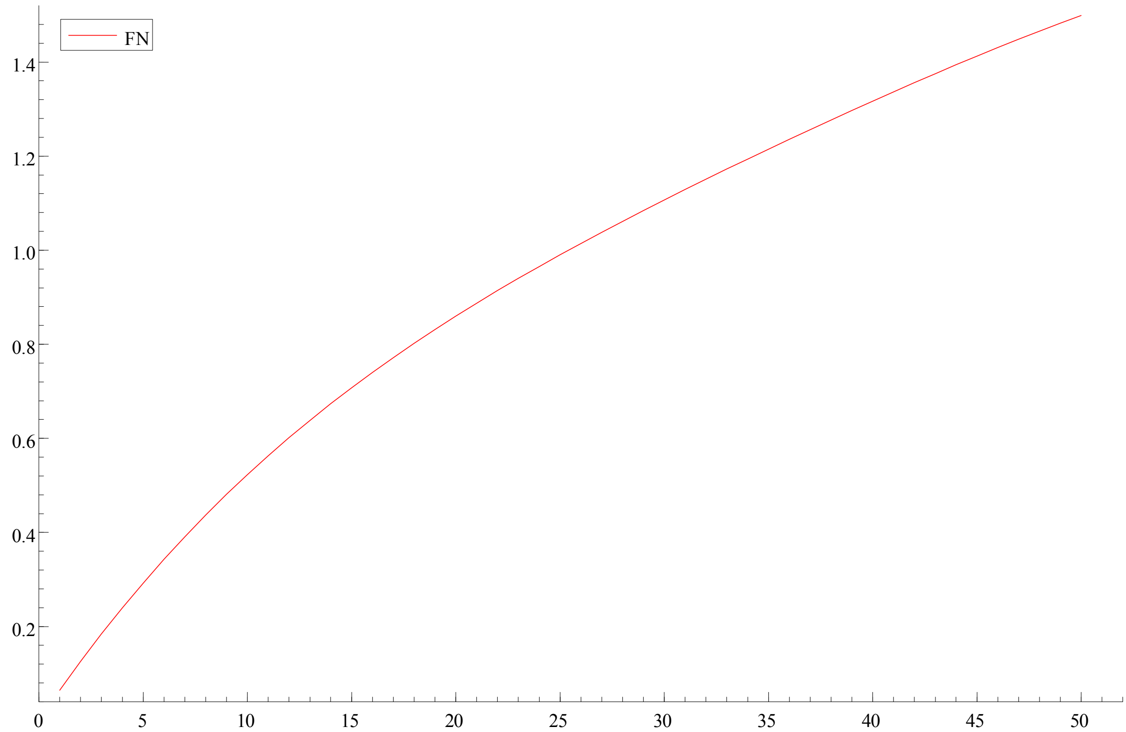

Figure 2. The total effect of the fourth order experience polynomial in the real wage equation. FN