Pergamon

Economics of Education Review 21 (2002) 221–229

Economics of Education Review

www.elsevier.com/locate/econedurev

Educational mismatch and wages: a panel analysis

Thomas K. Bauer *

IZA, P.O. Box 7240, 53072 Bonn, Germany

Received 5 December 1999; accepted 4 December 2000

Abstract

This paper contributes to the literature considering the wage effects of educational mismatch. It uses a large German panel data set for the period 1984–1998 and stresses the importance of controlling for unobserved heterogeneity when analyzing the labor market effects of over- and undereducation. Using pooled OLS, the estimation results confirm those found in the existing literature. The estimated differences between adequately and inadequately educated workers become smaller or disappear totally, when controlling for unobserved heterogeneity.

JEL classification: I21; J31

Keywords: Overeducation; Undereducation; Human capital model; Job competition model

1. Introduction

In recent years a growing number of studies have examined the incidence and labor market effects of educational mismatch.[1] This literature differentiates between the attained schooling level of an individual and the level of education required for the job an individual performs. Using these two concepts of education, measures of over- and undereducation are derived to study the consequences of educational mismatch. Among the issues discussed in this literature are the effects of inadequate education on job and career mobility (Sicherman, 1991; Hersch, 1991; Robst, 1995a; Alba-Ramirez, 1993; Tsang, Rumberger, & Levin, 1991; Sloane, Battu, & Seaman, 1999; Büchel & Mertens, 2000) and on job satisfaction (Hersch, 1991; Tsang et al., 1991; Büchel, 1998).

Most of the existing empirical studies analyze the wage effects of inadequate schooling.[2] They augment the standard Mincerian earnings function by estimating the returns to required years of education and the returns to years of educational mismatch instead of just estimating the returns to attained years of schooling. A remarkable feature of these empirical studies has been the consistency and robustness of their findings, which seem to hold across different time periods and different countries. In general, they find that (see Hartog, 2000):

1. the returns to actual years of schooling are lower than the returns to required years of schooling on a job;
2. the returns to surplus schooling are positive, but smaller than those to required education;
3. the returns to years of undereducation are negative. This penalty, however, is smaller than the returns to required education and usually smaller than the returns to overeducation;
4. the estimated returns to overeducation are always sig-

* Tel.: +49-228-3894-529; fax: +49-228-3894-510.
E-mail address: bauer@iza.org (T.K. Bauer).

[1] See for example the special issue on "Overschooling" in the *Economics of Education Review*, Vol. 19, 2000.

[2] See Groot (1993), Kiker, Santos, and Oliveira (1997), and Hartog (2000) for a survey of the wage effects of inadequate education.

0272-7757/02/$ - see front matter
PII: S0272-7757(01)00004-8

nificantly different from zero, which is not always the case for the returns to undereducation; and

5. even though the estimated returns to overeducation/undereducation are affected by the measure of required education, the above conclusions are not sensitive to changes in the measure of required schooling.

A potential problem of the existing studies however lies in the data sets that they have used, since most employ only cross-section data. It is possible that the estimation results of these studies are biased due to unobserved heterogeneity of individuals.[3] Controlling for unobserved heterogeneity might be important if the probability of educational mismatch is correlated with innate ability. Using a panel data set — the German Socioeconomic Panel (GSOEP) — this paper investigates the extent and the wage effects of educational mismatch in Germany using two different measures of educational mismatch. Applying panel estimation techniques, the paper analyzes whether previous results concerning the wage effects of inadequate schooling remain when controlling for unobserved heterogeneity. Section 2 describes the econometric approach used in the analysis. A brief description of the data set together with a discussion of possible measures of over- and undereducation is given in Section 3. The estimation results are presented in Section 4. Section 5 concludes.

2. Empirical approach

Two basic specifications have been used in the literature to study the wage effects of inadequate schooling. Following Verdugo and Verdugo (1989), the first model can be written as:

$$\ln Y_{it}=\alpha_0+\alpha_1 E_{it}+\alpha_2 OV_{it}+\alpha_3 UN_{it}+X_{it}\gamma+\varepsilon_{it} \quad (1)$$

where $\ln Y_{it}$ denotes the log of gross real hourly wages of individual i in year t; E_{it} describes years of education actually attained; OV_{it} and UN_{it} are dummy variables which take the value 1 if an individual is overeducated or undereducated, respectively, and 0 if an individual is correctly matched; X_{it} is a vector containing other explanatory variables with the corresponding vector of coefficients γ; and ε_{it} is an error term.

According to Eq. (1), mismatched workers are compared to similar workers with the same level of schooling who work in jobs that just require their attained level of schooling. In the case where productivity and wages are determined by the actual level of education, the coefficients α_2 and α_3 are expected to be zero. If wages are determined by the education level required to perform a job, any number of years of schooling exceeding the required amount would be unproductive and the reward to these additional years would be zero. In this case an overeducated worker would earn less than a similar worker with the same level of schooling who is adequately allocated to a job and α_2 would be negative. Similarly, an undereducated worker would earn more than an adequately allocated worker with the same education and α_3 would be positive.

The second model, which is due to Duncan and Hoffman (1981),[4] decomposes actual years of schooling (E_{it}) into required years of schooling (E^{r}_{it}), years of overschooling (E^{o}_{it}), and years of underschooling (E^{u}_{it}) using the following definition:

$$E_{it}=E^{r}_{it}+E^{o}_{it}-E^{u}_{it} \quad (2)$$

where

$$E^{o}_{it}=\begin{cases}E_{it}-E^{r}_{it}, \text{ if } E_{it}>E^{r}_{it}\\ 0, \text{ otherwise}\end{cases}, \text{ and } E^{u}_{it}=\begin{cases}E^{r}_{it}-E_{it}, \text{ if } E_{it}<E^{r}_{it}\\ 0, \text{ otherwise}\end{cases} \quad (3)$$

Using this definition the second specification can be written as:

$$\ln Y_{it}=\beta_0+\beta_1 E^{r}_{it}+\beta_2 E^{o}_{it}+\beta_3 E^{u}_{it}+X_{it}\gamma+\varepsilon_{it} \quad (4)$$

where β_1 is the return to years of required education; β_2 is the return to an additional year of schooling beyond those required, and β_3 is the return to a year of schooling below the schooling requirement. Differently to the Verdugo–Verdugo model, β_2 and β_3 have to be interpreted relative to workers in the same occupation who are correctly matched.

As Hartog and Oosterbeek (1988) note, two competing theoretical models, the human capital theory and the job competition theory as derived by Thurow (1975), are nested in specification (4). According to the human capital theory, earnings are not affected by the requirements of a particular job, since job level and related variables are subsumed in the age–earnings profile of a Mincer-type wage regression (Mincer, 1974). The human capital model emerges from Eq. (4) if $\beta_1=\beta_2=-\beta_3$. In this case, Eq. (4) reduces to the standard human capital earnings equation:

$$\ln Y_{it}=\gamma_0+\gamma_1 E_{it}+\varepsilon_{it} \quad (5)$$

[3] Some studies of the wage effects of educational mismatch use data with a panel character (see for example Battu, Belfield, & Sloane, 1999; Daly, Büchel, & Duncan, 2000; Dolton & Vignoles, 2000). None of these studies, however, tried to employ the panel character of their data to control for unobserved heterogeneity.

[4] See Groot (1993) for a survey of the empirical evidence using the Duncan–Hoffman specification.

In contrast to the human capital theory, the job competition theory is a demand-side theory, where marginal productivity is taken as a fixed characteristic of a particular job and is not related to the characteristics of the worker. Therefore, earnings are related to the job rather than to the worker. The job competition model emerges from Eq. (4) if $\beta_2=\beta_3=0$.

So far, existing empirical studies on the wage effects of inadequate schooling are based on cross-section data, raising the possibility that the results of these studies are biased due to unobserved heterogeneity. Controlling for unobserved heterogeneity might be important if individuals with lower innate ability need more education to attain a job for which they are formally overeducated. Using data from two cohorts of UK graduates, Chevalier (2000) shows for example that at least 30–40% of the estimated income differential between overeducated and correctly matched graduates can be explained by controlling for unobserved ability. Robst (1995b) shows that there is a negative relationship between college quality and the probability of being overeducated and a positive relationship between college quality and the probability of leaving overeducation status. Following this argument, one would expect that the estimated coefficients on the overeducation dummy in Eq. (1) will become more positive when controlling for unobserved heterogeneity, since unobserved ability and the probability of being overeducated are negatively correlated. Applying the same argument, the coefficient on the undereducation dummy should become less positive, since unobserved ability and the probability of being undereducated are positively correlated.

Two main results of the existing empirical studies estimating Eq. (4) are that the returns to surplus schooling are positive ($\beta_2>0$) but smaller than the returns to required education ($\beta_1>\beta_2$) and that the returns to years of undereducation are negative ($\beta_3<0$), but that this penalty is smaller than the returns to required education ($\beta_1>-\beta_3$). Note that if unobserved heterogeneity plays an important role, the estimated returns to the years of over- and undereducation are biased against the human capital model. Assume that overeducated workers have less and undereducated workers more unobserved abilities than correctly matched workers in the same occupation. In this case, the estimated returns to years of overeducation will be underestimated and the returns to years of deficit education will be overestimated, when not controlling for unobserved heterogeneity. Hence, it could be expected that the absolute value of the returns to years of required education, overeducation and undereducation become more similar to each other when controlling for unobserved heterogeneity. To investigate the role of unobserved heterogeneity in the analysis of the wage effects of educational mismatch, Eqs. (1), (4) and (5) are estimated using pooled OLS, random and fixed effects panel estimators.[5]

3. Data and measurement of required education

Educational mismatch is usually measured as the difference between the educational attainment of an individual and the skill requirements of the job an individual performs, i.e. a worker is considered to be overeducated (undereducated) if he has more (less) education than required for his or her job and adequately educated if his or her education just meets the requirements. In order to classify a worker to be inadequately educated one has to estimate the amount of schooling required for a particular job. Three different methods have been discussed in the literature to determine the amount of required education.[6] First, several studies have used an exogenous definition of schooling requirements which is based on information obtained from the evaluation of occupations by professional job analysts. The second method uses the self-assessment of workers to determine the education required for a job. The following empirical analysis makes use of the third method, which derives required education in a particular occupation from realized job matches.

Two different measures can be derived from this method. The first measure, which follows from Verdugo and Verdugo (1989), defines required schooling as a one-standard-deviation range around the mean level of schooling within an occupation. Workers are considered to be adequately educated if their actual education falls within this range, overeducated if their actual education is greater than one standard deviation above the mean for the specific occupation, and undereducated if their actual education is more than one standard deviation below the mean education. The main criticism of this measure has been its subjectivity, since there is no rationale behind the arbitrary choice of one standard deviation. As Kiker et al. (1997) note, this measure is more sensitive to technological change and changes in workplace organization than others, potentially suggesting misleading conclusions about the development of inadequate schooling over time.

Kiker et al. (1997) suggest using the modal value instead of the mean level of education within a given occupation to measure required schooling. Hence, workers are considered to be adequately educated if their actual education equals the mode value of education within their occupation, overeducated if attained education is above the modal value of education, and undereducated if attained education falls below the mode value. The main advantages of this measure over the one

[5] See Hsiao (1986) and Baltagi (1995) for surveys of panel data estimation techniques.

[6] A detailed discussion of the advantages and disadvantages of the different measures is given by Sicherman (1991), Kiker et al. (1997), Büchel (1998) and Hartog (2000).

proposed by Verdugo and Verdugo (1989) is that it is less sensitive to outliers and technological change.

The sample used in this study is drawn from the German Socio-Economic Panel (GSOEP), a panel data set for the years 1984–1998 consisting of about 13,500 individuals and 7000 households living in Germany. All full-time employed, prime-aged males and females of German nationality were selected. Foreigners and East Germans were excluded in order to study only those individuals who received their education in a similar schooling system. In addition, self-employed, civil servants, and those currently in education or training were not considered, leaving a sample of 26902 person–year observations of 4344 individuals.

A major criticism of the mean and the mode measure of required education is that such measures might result in unreliable estimates because of small cell sizes in some occupations. To mitigate this problem, the two measures of required education are calculated for every year of the panel using occupations disaggregated on a 2-digit level rather than a 3-digit level. Furthermore, occupations with less than 10 observations in a year are excluded from the analysis.[7] After eliminating all observations with missing values, a final sample of 13364 person–year-observations of 1824 males and 5273 person–year-observations of 922 females remain. The vector of control variables, X_{it}, includes experience and experience squared, tenure and tenure squared, a dummy for marital status, three dummies for firm size, two dummies for region, five dummies for industries and year dummies.[8]

Table 1 provides means as well as the respective overall, between-group and within-group standard deviations for the most important variables. According to the mean index, the majority of German employees are adequately educated. About 12% of males are estimated to be overeducated and 10% to be undereducated. Among females, 10% are overeducated and 15% are undereducated. Overeducated males have about 1.3 years and overeducated females about 1.2 years more schooling than required for the job they perform. The respective numbers for undereducated males and females are 0.7 and 0.6 years, respectively. Note that these numbers are comparable to those found in other studies using the mean index (see Hartog, 2000). However, the numbers differ from those found by Daly et al. (2000) and Büchel (1998), who use a subjective measure of required education to determine the incidence of educational mismatch in Germany. Using the 1984 wave of the German Socioeconomic Panel, Daly et al. (2000) conclude that about 14% of German males are overeducated with about 2.2 years of surplus education and 21% of females are overeducated with about 2.6 years of overeducation. Compared to the numbers obtained in this study, Daly et al. (2000) report relatively low numbers of undereducated persons. Based on their results 6.9% of German males and 7.4% of German females are undereducated. However, conditional on being undereducated the average years of deficit education obtained by Daly et al. (2000) are very similar to those reported in Table 1. The huge difference of the incidence of undereducation in Germany between Daly et al. (2000) and this study could partly be explained by the tendency of the mean index to find symmetry in the incidence of under- and overeducation (see Hartog, 2000).

The mode index gives a slightly different picture of the incidence of educational mismatch in Germany than the mean index. Both the share of employees being over- or undereducated as well as the average years of over- and undereducation are higher for the mode index. The difference between the two measures reflects the different computation of the two indices. Required education for the mean index is defined as a range around the mean years of schooling in an occupation and in the mode index as the most frequent level of schooling. It could be expected that in the mean index, individuals have a higher probability of being classified as adequately educated than in the mode index (Kiker et al., 1997).

A major question arising when estimating fixed effects panel models is whether the schooling variables have enough variation to identify the effects of educational mismatch on wages. In the sample used for the analysis, a change in the mismatch status could be observed for 5.3% of the observations when using the mean index and for 16.2% of the observations when using the mode index. Table 1 further shows that required years schooling and average years of undereducation have a much lower within-group variation for the mean index when compared to the mode index. The small numbers of a change in the mismatch status and the low within-group variation for the variables calculated on the basis of the mean index cast some doubts on whether the wage effects of educational mismatch could be identified in the fixed effects model.

[7] All estimations reported below have also been performed using required education calculated on the basis of 3-digit level occupational codes. This procedure reduces the sample to 8964 person–year observations of 1470 males and 3951 person–year-observations of 754 females. Note that the main conclusions reported below are not sensitive to a change in the aggregation level of occupations. The main conclusions also do not change when individuals working in the public sector are excluded from the sample. The estimation results using required education derived from the 3-digit level occupational codes and when excluding individuals working in the public sector are available on request.

[8] The estimation results for these additional explanatory variables will not be discussed. A full set of the estimation results is available on request.

Table 1
Summary statistics[a]

	Male	Female
Log (real hourly wage)	2.994	2.752
	(0.41) [0.34] {0.28}	(0.42) [0.36] {0.25}
Years of schooling	11.354	11.136
	(2.07) [2.17] {0.25}	(1.82) [1.91] {0.26}
Mean index:		
Years of required schooling	11.284	11.337
	(1.36) [1.32] {0.44}	(1.02) [1.03] {0.36}
Years of overeducation (for those overeducated)	1.301	1.202
	(1.27) [1.10] {0.51}	(1.10) [1.05] {0.43}
Years of undereducation (for those undereducated)	0.702	0.622
	(0.58) [0.51] {0.25}	(0.54) [0.55] {0.23}
Overeducated	0.123	0.107
Undereducated	0.104	0.156
Mode index:		
Years of required schooling	11.249	11.301
	(1.89) [1.65] {1.02}	(1.72) [1.56] {0.98}
Years of overeducation (for those overeducated)	1.738	1.733
	(1.74) [1.79] {0.41}	(1.54) [1.51] {0.42}
Years of undereducation (for those undereducated)	2.085	1.844
	(1.85) [1.89] {0.80}	(1.64) [1.66] {0.77}
Overeducated	0.308	0.299
Undereducated	0.206	0.370
Number of individuals	1824	922
Number of observations	13364	5273

[a] (...): Overall standard deviation; [...] between standard deviation; {...} within standard deviation. Source: German Socioeconomic Panel (SOEP) 1984–1998, own calculations.

4. The returns to over- and undereducation

4.1. Verdugo–Verdugo model

Table 2 reports the estimation results for Eq. (1). For the mean index, the results for the pooled OLS suggest that overeducated male workers earn 10.6% less and undereducated male workers 8% more than male workers with the same amount of education who are working in occupations which fully utilize their educational level. Undereducated females do not earn significantly more than correctly matched females with the same education. Overeducated females earn 15.1% less than similar females in occupations for which they are adequately educated.

Using the mode index, the wage gap between correctly matched and overeducated males drops to 1.2%. However, this difference is not statistically different from zero. The earnings advantage of undereducated males drops to 7.3%. Similarly, the earnings disadvantage of overeducated females drops from 15.1% to 3.3% when using the mode index. In contrast to males, however, this disadvantage is still statistically significant. Undereducated females earn about 3.9% more than correctly matched females with the same level of schooling. Overall, the estimation results using pooled OLS are in line with previous studies using the same specification, i.e. overeducated workers earn less and undereducated workers earn more than equally educated workers who are employed in occupations that fully utilize their education.

Using panel estimation techniques changes the above picture dramatically. Lagrange multiplier tests for random effects (Breusch & Pagan, 1980) favors the random effects model over the pooled OLS, and Hausman specifications tests (Hausman, 1978) show that the random effects model can be rejected against the fixed effects model for all specifications. Note however, that the estimated coefficients for the fixed effects model when using the mean measure should be interpreted with care, since only a few individuals in our sample change their matching status.

For both the random effects and the fixed effects model, the estimated coefficients of the educational mismatch dummies change in the expected direction. In most cases the absolute values of the estimated coefficients on the dummies indicating educational mismatch are significantly lower when unobserved characteristics are accounted for. Referring to the fixed-effects estimates for the mode index, the wage difference between overed-

Table 2
Estimated earnings functions: Verdugo and Verdugo model[a]

	Pooled OLS		Random effects		Fixed effects	
	Mean	Mode	Mean	Mode	Mean	Mode
Male						
Years of schooling	0.088**	0.079*	0.073**	0.070**	0.093**	0.092**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)
Overeducated	−0.106**	−0.012	−0.042**	−0.016**	−0.028**	−0.017**
	(0.016)	(0.010)	(0.008)	(0.006)	(0.09)	(0.006)
Undereducated	0.080**	0.073**	0.029**	0.020**	0.010	0.005
	(0.015)	(0.011)	(0.008)	(0.006)	(0.010)	(0.006)
R^2	0.68	0.68	0.67	0.67	0.12	0.12
Female						
Years of schooling	0.083**	0.074**	0.062**	0.059**	0.084**	0.082**
	(0.006)	(0.006)	(0.004)	(0.004)	(0.007)	(0.007)
Overeducated	−0.151**	−0.033**	−0.022*	−0.008	0.006	−0.004
	(0.025)	(0.016)	(0.012)	(0.008)	(0.013)	(0.008)
Undereducated	0.026	0.039**	−0.003	−0.011	−0.009	−0.027**
	(0.021)	(0.016)	(0.011)	(0.007)	(0.12)	(0.008)
R^2	0.70	0.70	0.67	0.68	0.14	0.13

[a] Standard errors in parentheses. * Statistically significant at least at the 10% level. ** Statistically significant at least at the 5% level. All regressions include the variables experience and experience squared, tenure and tenure squared, a dummy for marital status, three dummies for firm size, two dummies for region, five dummies for industry and year dummies. Male sample: 13364 observations of 1824 individuals; female sample: 5237 observations of 922 individuals. For the random effects and the fixed effects model R^2 refers to the overall R^2. Breusch and Pagan (1980) Lagrangian multiplier test statistics for random effects have been significant at least at the 1% level for all specifications. Hausman tests on the hypothesis that the coefficients in the random effects model and the fixed effects model are the same could be rejected for all specifications at least at the 1% level. Source: German Socioeconomic Panel (SOEP) 1984–1998, own calculations.

ucated and correctly matched male workers drop to 1.7% and the positive wage differential between undereducated male workers and correctly educated workers disappears. The results are slightly different for females. Overeducated females do not have an earnings disadvantage compared to adequately matched females with the same level of schooling; undereducated females even show an earnings disadvantage to correctly matched females. Overall, the estimation results for both the random and the fixed effects model show that accounting for unobserved heterogeneity is important when analyzing the wage effects of educational mismatch.

4.2. Duncan and Hoffman model

Table 3 reports the estimation results for Eqs. (4) and (5) together with values of F-tests for the hypothesis of the human capital model that the years of adequate-, over- und undereducation should be rewarded equally, i.e. $\beta_1=\beta_2=-\beta_3$, and for the prediction of the job competition model that only adequate schooling is rewarded, i.e. $\beta_2=\beta_3=0$. The estimation results for the pooled OLS model come to the same conclusions as the existing literature. For both sub-samples and both measures of required education, the returns to required education are higher than the returns to attained years of schooling. The returns to years of surplus education are positive and the returns to years of deficit education are negative. The penalty to a year of deficit education and the additional returns to a year of overschooling are lower than the returns to required education. Note that for the mode index the estimated returns to required schooling and educational mismatch are very similar to those obtained by Daly et al. (2000), who use just the 1984 wave of the German Socioeconomic Panel and derive required education from the self-assessment of workers. When using the mean index the estimated returns are in most cases significantly higher in absolute terms compared to those of Daly et al. (2000).[9]

Similar to the results of the Verdugo and Verdugo model, the conclusions change dramatically, when using panel estimation techniques. For all specifications, Hausman specification tests and Lagrange multiplier tests

[9] For males and females, Daly et al. (2000) estimate a coefficient of 0.09 for years of required education. For males, the estimated coefficient on years of overeducation is 0.049, the estimated coefficient on years of deficit education −0.078. For females these coefficients take the values 0.066 and −0.038, respectively. All coefficients are statistically significantly different from zero.

Table 3
Estimated earnings functions: Duncan and Hoffman model[a]

	Pooled OLS		Random effects		Fixed effects	
	Mean	Mode	Mean	Mode	Mean	Mode
Male						
Years of schooling	0.073**		0.067**		0.090**	
	(0.002)		(0.002)		(0.005)	
R^2	0.68		0.66		0.12	
Years of required schooling	0.107**	0.084**	0.069**	0.071**	0.012**	0.092**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.005)
Years overeducated	0.090**	0.066**	0.062**	0.063**	−0.010	0.086**
	(0.009)	(0.004)	(0.005)	(0.002)	(0.007)	(0.005)
Years undereducated	−0.100**	−0.050**	−0.087**	−0.065**	−0.013	−0.092**
	(0.016)	(0.005)	(0.010)	(0.003)	(0.011)	(0.005)
R^2	0.69	0.69	0.67	0.67	0.23	0.12
H_0: $\beta_1=\beta_2=-\beta_3$	1.82	54.75**	6.27**	40.41**	7.53**	6.11**
H_0: $\beta_2=\beta_3=0$	75.06**	223.91**	251.90**	861.74**	2.41	162.83**
Female						
Years of schooling	0.064**		0.060**		0.086**	
	(0.005)		(0.004)		(0.008)	
R^2	0.69		0.67		0.14	
Years of required schooling	0.125**	0.087**	0.060**	0.064**	−0.005	0.085**
	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)	(0.007)
Years overeducated	0.052*	0.045**	0.045**	0.057**	−0.019*	0.084**
	(0.023)	(0.007)	(0.008)	(0.004)	(0.011)	(0.007)
Years undereducated	−0.115**	−0.070**	−0.063**	−0.062**	0.022	−0.088**
	(0.018)	(0.006)	(0.012)	(0.004)	(0.016)	(0.007)
R^2	0.71	0.70	0.67	0.67	0.25	0.14
H_0: $\beta_1=\beta_2=-\beta_3$	4.31**	24.81**	2.83	5.78*	1.64	0.48
H_0: $\beta_2=\beta_3=0$	26.82**	91.17**	54.88**	285.62**	2.06	81.15**

[a] Standard errors in parentheses. * Statistically significant at least at the 10% level. ** Statistically significant at least at the 5% level. All regressions include the variables experience and experience squared, tenure and tenure squared, a dummy for marital status, three dummies for firm size, two dummies for region, five dummies for industry and year dummies. Male sample: 11462 observations of 1709 individuals; female sample: 4675 observations of 869 individuals. For the random effects and the fixed effects model R^2 refers to the overall R^2. Breusch and Pagan (1980) Lagrangian multiplier test statistics for random effects have been significant at least at the 1% level for all specifications. Hausman tests on the hypothesis that the coefficients in the random effects model and the fixed effects model are the same could be rejected for all specifications at least at the 1% level. Source: German Socioeconomic Panel (SOEP) 1984–1998, own calculations.

indicate again that the random effects model is rejected against the fixed effects model and that the pooled OLS is rejected against the random effects model. Overall, for both measures and both sub-samples the estimated returns to years of required education fall and become very similar to the estimated returns to required education. Referring to the random effects model, the estimated returns to years of surplus schooling drops from 9% to 6.2% for males and the penalty for an additional year of undereducation drops from 10% to 8.7% when using the mean index. For females the returns to years of overeducation stays roughly constant and the penalty for an additional year of deficit education drops from 11.5% to 6.3%. For the mode index, the estimated returns to years of over- and undereducation are roughly similar to those of the pooled OLS estimates for both sub-samples.

As expected, the fixed effects model for the mean index does not lead to reasonable results, which can be explained by the low within-sample variation of the schooling variables. For the mode index, however, controlling for unobserved individual characteristics using fixed effects changes the estimated coefficients in the expected way. For both sub-samples, the estimated returns to years of required education do not change significantly. Furthermore, the returns to required education are not significantly different from the returns to attained years of schooling. For males, the returns to an additional year of surplus schooling increase by 2 percentage points and the penalty for an additional year of deficit education

increases by 4.2 percentage points as compared to the results of the pooled OLS. For the female sub-sample the returns to overschooling increase by 4 percentage points and the negative effect of years of underschooling increases by 2 percentage points. For both sub-samples, the absolute values of the coefficients on years of required education and years of educational mismatch become more similar to each other. The hypothesis that years of required education, years of overeducation and years of deficit education are rewarded equally could no longer be rejected for females. For males, however, the human capital hypothesis could still be rejected at a level of 5% significance.

5. Conclusion

This paper provides an empirical evaluation of the wage effects of over- and undereducation using a German panel data set. In particular, the paper investigates whether controlling for unobserved individual characteristics changes the major conclusions of the literature on the wage effects of educational mismatch, which are based mainly on empirical studies using cross-section data. The empirical analysis of the paper uses two different measures of educational mismatch, drawn from previous literature. The first measure defines required education as a one-standard-deviation range around the mean schooling level in 2-digit occupations. The second measure defines required schooling as the modal value of years of schooling in 2-digit occupations.

The results of the effects of educational mismatch on wages using pooled OLS estimation confirm the main conclusions of the existing empirical literature. Overeducated workers earn less and undereducated workers earn more than workers with the same level of educational attainment but who work in occupations that fully utilize their education. The returns to required education appear to be significantly higher than the returns to attained years of schooling. The returns to years of surplus education are positive and the returns to years of deficit education are negative. Finally, the penalty to a year of deficit education and the additional returns to a year of overschooling are lower than the returns to required education. The estimation results for the pooled OLS imply that both the human capital model, which assumes equal rate of returns to adequate-, over- and underschooling, and the job competition model, which assumes that there are no returns on over- and underschooling, could be strongly rejected.

The estimated effects change dramatically when one controls for unobserved heterogeneity using panel estimation techniques. The earnings differences between inadequately educated workers and equally educated workers who work in occupations for which they are adequately educated becomes at least smaller, and in most cases disappears totally. The rates of return to years of attained education become similar to the rates of return to required schooling and the absolute values of the coefficients on years of required education and years of educational mismatch become more similar to each other. Similar to the results for the pooled OLS, the job competition model could be strongly rejected when using panel estimators. The human capital model could still be rejected for males, but no longer for females. The results of this paper indicate the need to test whether existing results on the labor market effects of educational mismatch for various countries change when the importance of unobserved individual effects is taken into account.

Acknowledgements

The author would like to thank Kurt Brännäs, Don DeVoretz, John Haisken-DeNew, Joop Hartog, Magnus Lofström, Markus Pannenberg, Anja Thalmaier, Rainer Winkelmann, Melanie Ward and two anonymous referees for their helpful comments on earlier drafts of this paper.

References

Alba-Ramirez, A. (1993). Mismatch in the Spanish labor market: overeducation? *Journal of Human Resources, 28* (2), 259–278.

Baltagi, B. H. (1995). *Econometric analysis of panel data.* New York: John Wiley.

Battu, H., Belfield, C., & Sloane, P. (1999). Overeducation among graduates: a cohort view. *Education Economics, 7* (1), 21–38.

Breusch, T., & Pagan, A. (1980). The LM test and its applications to model specification in econometrics. *Review of Economic Studies, 47* (1), 239–254.

Büchel, F. (1998). *Zuviel gelernt? Ausbildungsinadäquate Erwerbstätigkeit in Deutschland.* Bielefeld: W. Bertelsmann Verlag.

Büchel, F., & Mertens, A. (2000). *Overeducation, undereducation, and the theory of career mobility. IZA Discussion Paper No. 195, Bonn.*

Chevalier, A. (2000). Graduate over-education in the UK. In *Third European summer school in labor economics, Buch at the lake of Ammersee, Germany, June.*

Daly, M. C., Büchel, F., & Duncan, G. J. (2000). Premiums and penalties for surplus and deficit education: evidence from the United States and Germany. *Economics of Education Review, 19* (2), 169–178.

Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review, 19* (2), 179–198.

Duncan, G. J., & Hoffman, S. D. (1981). The incidence and wage effects of overeducation. *Economics of Education Review, 1* (1), 75–86.

Groot, W. (1993). Overeducation and the returns to enterprise-

related schooling. *Economics of Education Review*, *12* (4), 299–309.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review*, *19* (2), 131–147.

Hartog, J., & Oosterbeek, H. (1988). Education, allocation and earnings in the Netherlands: overschooling? *Economics of Education Review*, *7* (2), 185–194.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*, 1251–1271.

Hersch, J. (1991). Education match and job match. *Review of Economics and Statistics*, *73* (1), 140–144.

Hsiao, C. (1986). *Analysis of panel data.* New York: Cambridge University Press.

Kiker, B. F., Santos, M. C., & De Oliveira, M. M. (1997). Overeducation and undereducation: evidence for Portugal. *Economics of Education Review*, *16* (2), 111–125.

Mincer, A. (1974). *Schooling, experience and earnings.* New York: NBER Press.

Robst, J. (1995a). Career mobility, job match, and overeducation. *Eastern Economic Journal*, *21* (4), 539–550.

Robst, J. (1995b). College quality and overeducation. *Economics of Education Review*, *14* (3), 221–228.

Sicherman, N. (1991). Overeducation in the labor market. *Journal of Labor Economics*, *9* (2), 101–122.

Sloane, P. J., Battu, H., & Seaman, P. T. (1999). Overeducation, undereducation and the British labour market. *Applied Economics*, *31* (11), 1437–1453.

Thurow, L. C. (1975). *Generating inequality*. New York: Basic Books.

Tsang, M. C., Rumberger, R. W., & Levin, H. M. (1991). The impact of surplus schooling on worker productivity. *Industrial Relations*, *30* (2), 209–228.

Verdugo, R., & Verdugo, N. (1989). The impact of surplus schooling on earnings: some additional findings. *Journal of Human Resources*, *24* (4), 629–643.