

Kleibrink, Jan

Conference Paper

Causal Effects of Educational Mismatch in the Labor Market

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Labour Markets, No. E17-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Kleibrink, Jan (2013) : Causal Effects of Educational Mismatch in the Labor Market, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Labour Markets, No. E17-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/79839>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Causal Effects of Educational Mismatch in the Labor Market

Jan Kleibrink*

April 2013

Abstract

This paper analyzes the effect of educational match on wages in Germany, using data from the German Socio-Economic Panel. Educational mismatch has been discussed extensively, mostly by applying OLS wage regressions which are prone to an unobserved heterogeneity bias. This problem is approached by using an FE and an IV model. In a further step, ability is directly included in the regressions by using data from the International Adult Literacy Survey allowing for an explicit control of otherwise unobserved abilities. Results show that unobserved heterogeneity does not explain the wage differences between years of over-/ undereducation and years of required education. This rejects the hypothesis that mismatched workers compensate for heterogeneity in innate abilities. Results hint at a structural problem in the German educational system as skill demand and supply are not in a long-term equilibrium.

Keywords: Wages; Educational Mismatch

JEL classifications: I14; I21; J31

*Ruhr Universität Bochum, RUB Research School. Room GC 2/148, Universitätsstrasse 150, 44801 Bochum, Germany. Telephone: +49 234 32 25728. Email: jan.kleibrink@rub.de – Acknowledgment: The author thanks Prof. John P. Haisken-DeNew and Dr. Mathias Sinning for highly useful comments and suggestions.

1 Introduction

Overeducation is an extensively discussed topic in economic literature (for an overview, see Hartog, 2000; Sicherman, 1991). The consequences of overeducation were analyzed in different settings including wage regressions (Sicherman, 1991; Groeneveld and Hartog, 2004), job satisfaction (Battu et al., 1999) or general life satisfaction (Haisken-DeNew and Kleibrink, 2013). Although the topic has been analyzed in a vast amount of studies and over a long period of time, several conceptual issues are still to be sorted out. While the operationalization of over- and undereducation has received a lot of attention (see e.g. Verdugo and Verdugo, 1989; Rubb, 2003), the issue of causality has only been tackled over the last years (see e.g. Bauer, 2002; Korpi and TÅehlin, 2009). However, understanding the causality in the relationship between educational match and wages is a necessary prerequisite for a profound understanding of the underlying mechanisms. This study addresses the problem of unobserved heterogeneity, which is still a common concern in this field of literature. The discussion on the explanation of the effects of mismatches has not come to a consensus yet. This paper contributes to this discussion.

The standard way of operationalizing the overeducation framework compares the educational attainment of an individual to workers in the same occupation (classical OMU framework). This framework was introduced by Duncan and Hoffman (1981) and has become widely accepted in the mismatch literature (e.g. Rubb, 2003). The OMU model explicitly differentiates between all three possible educational matches: overeducation (O), educational match (M) and undereducation (U), all of these states measured in years of education. OLS results, which are remarkably stable over nearly all studies in this field, show that overeducation has a positive significant effect on wages. Hence, overeducated individuals earn more than less educated ones who are in the same occupation. However, the positive overeducation effect is significantly smaller than the effect of matched years of education. Hence, overeducated individuals are disadvantaged in terms of wages compared to others with the same education who are in a better match. Undereducation is either negative significant or insignificant (a summary is found in Hartog, 2000). Fewer consensus is found regarding a different problem: Unobserved heterogeneity. Using overeducation in classical OLS wage regressions relies on the assumption that equally educated individuals have the same innate ability and thereby productivity (given other controls). However, this is a very strong assumption and has been criticized in mismatch studies lately (Bauer, 2002; Korpi and TÅehlin, 2009). There are three ways of solving this problem: Fixed effects regressions (Bauer, 2002) the use of instruments (Korpi and TÅehlin, 2009) and the direct inclusion of ability controls (Korpi and TÅehlin, 2009). Each of these

approaches has advantages and weaknesses (a discussion of these approaches can be found in Section 3). Therefore, concentrating on one of them alone involves the danger of interpreting results that are heavily driven by the model assumptions. We compare the results of three different methods, ensuring a broad picture of the effects of educational mismatch. This offers a broad and robust base for our interpretation.

The missing consensus regarding the method of choice is accompanied by a problem regarding the interpretation of results. A prominent approach is the human capital hypothesis (Hartog and Oosterbeek, 1988), which was tested for the German labor market by Bauer (2002). The idea is that mismatched individuals are in a bad match regarding their formal education because they compensate for heterogeneity in their innate ability. When controlling for unobserved heterogeneity in the estimations, the coefficients of all three states should come close to each other to prove this hypothesis. Bauer (2002) finds that this compensation hypothesis cannot be completely rejected in Germany.

In the first part of the analysis, we replicate the OLS wage regressions with our sample. These standard OLS regressions reveal the findings that are well known from literature: There is a wage benefit for overeducated workers as compared to those having the same job but less education. The benefit is not as big as the one for matched years of education. Undereducated workers suffer a significant wage penalty. However, there might be a bias due to unobserved heterogeneity. FE and IV regressions show that unobserved heterogeneity does not explain wage differences between years of required education and years of over-/undereducation as found by Bauer (2002) using the same data set. Using an IV approach on SOEP data and modeling innate ability explicitly by using data from the IALS, we show that differences between educational match and mismatch even become bigger and only the years of required education matter while years of mismatch do not have any wage effects. This neglects the explanation that mismatched workers compensate for skill shortages and leads to an explanation of a structural problem in the German educational system as a huge amount of workers cannot find jobs fitting their educational level.

This analysis contributes to the economic literature in several dimensions: By applying three different strategies of dealing with unobserved heterogeneity, it offers the most comprehensive study on the causal effects of overeducation in Germany. The results contrast former results for the German labor market, offering a new line of argumentation. This is of huge political relevance as it hints at a large-scale problem in the allocation of individuals in the educational sector and in the labor market.

The paper is organized as follows: Section 2 integrates this study into the existing

literature, Section 3 explains the empirical methods applied to derive causal effects. Section 4 introduces the dataset used and explains the key variables. In Section 5 the empirical results are explained and discussed, a conclusion can be found in Section 6.

2 Literature

Educational mismatch has received lots of attention in the economic literature, mainly from an empirical point of view. Although already analyzed from various different angles, there are still gaps to be filled.

Duncan and Hoffman (1981) started the modern empirical mismatch literature by introducing a framework in which individual education consists of three parts: education required for a job, overeducation and undereducation. This was the starting point of the OMU theory. A well-matched worker has exactly the years of education required in his job (M), overeducated workers attained additional years of education which are not needed for their current working life (O). Undereducated workers received less education than required to do their jobs (U). By decomposing attained education into these three parts, it is possible to analyze if education is paid off in the labor market in general or if it matters whether this education is used productively. The OMU framework has become the standard approach in the overeducation literature, as can be seen in Sicherman (1991); Rumberger (1987); Alba-Ramirez (1993); Bauer (2002); Hartog and Oosterbeek (1988); Korpi and TÄhlin (2009). While these studies differ regarding their definition of overeducation, the datasets applied and the countries and time periods studied, they all share a common finding, which has become a stylized fact in the mismatch literature. Required education is positive and significant, the same is true for overeducation but the coefficient is significantly smaller. Undereducation is normally negative and significant but this finding is not as robust as the other two (for an extensive meta-analyzes of the literature, see Hartog, 2000; Rubb, 2003). The classical finding of mismatch studies has mainly been found using OLS wage regressions. However, more recent papers have started using different regression techniques to tackle a possible problem of earlier studies: unobserved heterogeneity. Studies by Bauer (2002) and Korpi and TÄhlin (2009) assume that unobserved heterogeneity biases OLS results and discuss this in the context of the human capital theory. This theory assumes that educational mismatch is not a result of a structural mismatch of skill demand and supply in the labor market but mismatched workers compensate for ability not captured by the educational attainment. Overeducated workers lack ability and compensate for this by getting more

education than they actually need to perform their job. The opposite is true for undereducated workers. As they have a higher innate ability than others, they can get better jobs without having the proper educational attainment. According to Bauer (2002), this assumption is proven right when the coefficients of the three components of education become more equal when controlling for unobserved heterogeneity. Using data from the German SOEP, Bauer (2002) finds that this hypothesis cannot be generally rejected for Germany as the coefficients become similar in size using panel models. Korpi and TÅěhlin (2009) do not only apply these panel models but also an IV model and a direct inclusion of skills. They reject the human capital theory but are concerned with a weak instrument problem.

This study leans on the empirical strategy by Korpi and TÅěhlin (2009). Using a the SOEP, which runs over a longer period of time, we can get more robust results from panel models and thereby replicate the results found by Bauer (2002). We then extend this study by following the strategy of Korpi and TÅěhlin (2009) and apply an IV approach and include skill measures directly. We use different instrument variables and thereby avoid the weak instrument problem our priors have to deal with. In a further step, we apply data from the International Adult Literacy Survey (IALS), a dataset explicitly designed to model skills. By using this dataset, we can directly include ability controls in the OMU regressions and thereby avoid unobserved heterogeneity. Using this strategy we contrast the results by Bauer (2002) and offer a different explanation for the existence of overeducation in the German labor market.

3 Estimation Method

Different estimation methods are applied in this analysis to obtain causal effects. The starting point for wage analyzes are Mincer wage regressions (Mincer, 1970). The classical wage regressions are extended by the variables for required education, overeducation and undereducation. All of these variables are measured in years. This gives the wage regression for the classical OMU framework (e.g. Hartog, 2000):

$$\ln(w_{it}) = F(x_{it}\beta, O_{it}\gamma, M_{it}\zeta, U_{it}\eta) = \beta_0 + x_{it}\beta + O_{it}\gamma + M_{it}\zeta + U_{it}\eta + \varepsilon_{it}$$

with the logarithm of hourly wage $\ln(w_{it})$ as explained variable. The explanatory variables are required education/ educational match (M), overeducation (O) and undereducation (U) in years. x_{it} is a matrix of further controls including the age and

its second polynomial, a dummy for the birth cohort, tenure, the number of children and marital status, nights spent in hospital as control for individual health, a dummy for fulltime employment as well as industry and year dummies. ε_{it} is the error term.

The question arising here is how to operationalize overeducation. Basically, there are three different options: (1) An objective approach relying on an expert valuation; (2) a subjective approach relying on worker's self assessment; (3) an empirical approach. The first option is mainly used in US studies (e.g. Rumberger, 1987; McGoldrick and Robst, 1996)). This can be done as the Dictionary of Occupational Titles (DOT) exists for the US labor market, providing the necessary information. The DOT includes information on occupations and the necessary education for these occupations. This assessment is made by labor market experts. A source like this does not exist for the German labor market. However, the unavailability is not the only reason not to use the expert method. Kiker et al. (1997) point out that this measure can only work if updated regularly to account for technological change as well as hiring standards. If not updated regularly, this measure tends to misclassify respondents over time, a problem growing the longer the last update is ago.

The subjective approach (e.g. Sicherman, 1991; Sloane et al., 1999) has the advantage over the expert valuation that it is updated automatically with each wave of a panel dataset. Respondents are asked for the qualification necessary to do their job and the answer is compared to the actually attained education to evaluate whether a person is in an educational match or not. However, this definition gives rise to other problems. The main criticism of this approach is that it remains unclear which benchmark is used by respondents. They could either use the qualification necessary to actually perform the job, while they could also answer according to hiring standards (Bauer, 2002). Whichever answer they give, it requires respondents really to know about the standards in their occupational field.

This study applies the empirical method of measuring overeducation. This method was applied by e.g. Verdugo and Verdugo (1989); Kiker et al. (1997); Bauer (2002). Verdugo and Verdugo (1989) introduced this measure by using the mean value for each occupation. Overeducated workers are those individuals whose education exceeds the mean value plus one standard deviation, undereducated are those whose education lies below the mean value minus one standard deviation. This approach offers several advantages over the ones discussed before: Firstly, it is naturally updated regularly, similar to the subjective approach. Secondly, it does not have the weakness of the subjective approach as it does not rely on an individual evaluation but on the distribution observed in the labor market. However, the use of a range of one standard deviation was criticized as arbitrary choice (Bauer, 2002) and the

method was shown to be prone to outliers (Kiker et al., 1997). Kiker et al. (1997) also used the empirical method but instead of relying on the mean value, they used the modal value within an occupation. This approach keeps the advantages of the approach by Verdugo and Verdugo (1989) without being prone to its abovementioned weaknesses. Within the course of this analysis, this modal value approach is applied.

The discussion about the operationalization of educational mismatch is important but it is only a first step because applying the measure in classical OLS wage regressions fails to derive causal effects. This is due to the problem of unobserved heterogeneity.

Applying OLS regressions leans on the assumption that, controlling for other covariates, workers only differ by their educational match. This assumption is highly doubtful. As argued in the more recent literature on mismatch (Bauer, 2002; Korpi and TÄhlin, 2009), unobserved influences like intelligence, productivity and motivation are important factors when analyzing mismatch. As these factors cannot be observed, OLS results are biased.

This analysis uses three different approaches to tackle the problem of unobserved heterogeneity. Bauer (2002) points out that using panel data, it is possible to estimate fixed effects regressions to control for unobserved influences. This approach is reproduced in this study. The wage regressions presented above remains the same, however, using the panel nature of the SOEP, individual fixed effects are controlled for. While this strategy controls for unobserved heterogeneity, it has a different problem. Only individuals changing their educational match within the observation period can be observed. Individuals who do not change their educational fit over time are not regarded as their educational match information is time-invariant.

This means that many mismatch observations cannot be regarded within the analysis. To account for this problem, an IV approach is used. Within the mismatch context, this was done by Korpi and TÄhlin (2009), who also point to the difficulty of using an IV approach in the OMU framework: All three education variables, overeducation, undereducation and the required education must be instrumented. Hence, at least three instruments have to be found that fulfill the criteria of instrumental variables; they have to have a significant effect on the instrumented variable (relevance) while they must not affect the outcome variable over a different channel (validity). Korpi and TÄhlin (2009) apply four instruments in their analysis, all of them related to the respondents' youth: the number of siblings, place of residence, economic problems and family disruption. While the authors argue that these instruments are valid, they fail to fulfill the relevance criterion.

Applying instrumental variable approaches to account for unobserved heterogeneity

in education/ wage studies is not a new idea. Angrist and Krueger (1991) use information on the quarter of birth on US census data. Card (1993) uses the proximity of the place of residence to the next college, a very famous approach, which was later shown as failing to fulfill the strength criterion (for a discussion, see Harmon et al., 2003). As already pointed out, unlike studies on the effects of education in general, this analysis need three different instruments to cover all three possible matches. The first, also used by Korpi and TÅehlin (2009), is in line with a study by Butcher and Case (1994), using the presence of siblings. We basically follow their reasoning including the number of siblings as instrument for the educational match. The number of siblings is negatively correlated with the years of education. This is due to a split of parental support and expectations on several children. Parents can be assumed to lay a strong focus on the educational career of single children, while this focus shifts, the higher the number of siblings. This might cause a lack of parental support during education. Our data support this as the number of siblings is negatively correlated with overeducation and the achieved education, while it is positively correlated with the case of undereducation. The second instrument applied is also in line with studies stressing the family background like Harmon and Walker (2000). In line with Korpi and TÅehlin (2009), an indicator for family disruption is applied. We take the number of years living with the biological parents until the age of 16 as instrument, following a similar reasoning as for sibship size. The more time spent living with the biological parents, the higher the support and achievements in the educational career. For the third instrument, we lean on approaches stressing macro changes as exogenous variations. Often used are schooling reforms (e.g. Harmon and Walker, 1995, 1999; Pons and Gonzalo, 2002)). As our data cover a long time and observations from all age groups, schooling reforms cannot be used. Instead, we apply labor market conditions at the respondents' age of 15. This is a time in which individuals decide (1) to stay in the academic track, (2) to leave for the labor market aiming at a more applied vocational education or (3) not to obtain any further education. A high unemployment rate at this time is likely to influence individuals to stay in the schooling system as an outside option to entering the labor market, which does not offer good opportunities at that time. Our data support this hypothesis. While these three variables all have an influence on the educational decision and thereby on the educational match, they do not have other wage effects. As all of these go back to the time of education, other effects on the current hourly wages are hardly found theoretically. This assumption is supported by the fact that each of these instruments in line with IV approaches already used in education- wage studies. We are confident that the instruments applied do not only fulfill the strength criterion, which can be tested statistically, but also the validity criterion. We are aware of

possible weaknesses of IV estimations like a lower precision of coefficients compared to other econometric techniques. Still, it is a further step into the direction of a comprehensive picture of the effects of educational mismatch which enables researchers to come to an interpretation not driven by the assumptions of a single model.

As both FE and IV regressions are not free from weaknesses, we further back our results by a third method. While the former ones aim at avoiding unobserved heterogeneity, this methods aims at making it observable by controlling for ability directly. In the OMU framework, Korpi and TÅhlin (2009) make an approach to do this by including measures of health and verbal ability to capture this dimension, however they do not find significant changes to their results without the controls. We are confronted with the problem that a large scale ability measure is not included in the SOEP. We therefore use data from the International Adult Literacy Survey (IALS) by the OECD. This dataset includes information on work-related reading, writing and math skills. These measure ability of individuals from a more applied perspective than the educational attainment. Including these as ability controls is a way of avoiding unobserved heterogeneity with regards to ability differences.

Applying the FE and IV approaches, as well as the model explicitly controlling for ability accounts for possible weaknesses of each of the models. Using a broad econometric strategy, results can be compared to find a general pattern and thereby contribute to the discussion about the mechanisms behind the wage effects of educational mismatch in the labor market.

4 Data

The data used in the first part of this analysis are from the German Socio-Economic Panel (SOEP), one of the longest running representative panel-datasets in Europe. Established in 1984, it covers more than 20,000 individuals per year and is representative of the German population (Wagner et al., 2007)¹.

The sample is restricted to working individuals between 18 and 65 and covers the years 1991 - 2011. We exclude individuals who lived in Eastern Germany before the German reunification as well as immigrants who came to Germany after their 10th birthday. These restrictions ensure that respondents were educated in the same educational system and results are not driven by a different perception of educational titles from other countries. This proceeding is in line with Bauer (2002), who also

¹All data were extracted using the Stata add-on PanelWhiz, written by Prof. Dr. John P. Haisken-DeNew (Haisken-DeNew and Hahn, 2010).

uses SOEP data for his analysis of educational mismatch.

Table 1: DESCRIPTIVE STATISTICS - CONTROL VARIABLES

	Men		Women	
	Mean	Std.Dev.	Mean	Std.Dev.
Hourly Income	9.81	(6.31)	7.29	(4.85)
Age	40.35	(11.56)	39.49	(11.44)
Cohort	4.76	(1.23)	4.90	(1.20)
Nights in Hospital	0.75	(4.94)	0.76	(4.56)
Children	0.73	(0.95)	0.61	(0.87)
Fulltime	0.94	(0.23)	0.52	(0.50)
Tenure	12.17	(10.81)	9.16	(9.14)
Married	0.67	(0.47)	0.62	(0.48)
N	55415		45529	

Note: Authors' calculations based on SOEP.

Table 1 presents descriptive statistics for the variables used in the wage regressions. The first column presents statistics for men, the second for women. Average hourly wages for males are 9.81 Euros, for females 7.29 Euros. The mean age is around 40 for men, 39.5 for females. Men in our sample have 0.72 children, women 0.6. This can be explained by a lower labor market participation rate for mothers in Germany. The health control shows nearly equal values for men and women. Labor market specific controls show that tenure is about three years higher for men, at about 12 years. Men are most likely fulltime employed (94%), while a little more than half of the women in our sample work fulltime.

Table 2: DESCRIPTIVE STATISTICS - OVEREDUCATION

	Men		Women	
	Mean	Std.Dev.	Mean	Std.Dev.
OE Mode	2.16	(2.83)	1.59	(2.40)
Educ Mode	14.59	(2.27)	14.48	(1.98)
UE Mode	0.43	(1.04)	0.58	(1.06)
N	55415		45529	

Note: Authors' calculations based on SOEP.

In Table 2, descriptive statistics for the main interest variables in the OMU framework can be found. To derive results which show the situation in Germany as precise as possible, we avoid using a standard education variable in the SOEP, which offers mapped information on the years of education.² This variable assigns each individual the years of education typically necessary to obtain the highest achieved degree. For example, leaving education after the A level means having 13 years of education, a vocational education means further 1.5 - 2 years, a university education 5

²A detailed documentation of the data properties is offered by the data provider, the DIW (see <http://www.diw.de>). A documentation of the mapping of years of education is offered in the SOEP documentation by Anger (2011).

further years. While this is a useful tool for many applications, it does not fully meet the requirements of this study. Especially for higher educated individuals, this leads to measurement problems. For this variable, the standard period of studies is used. This, however, is likely to be extended, the higher the educational degree. For a vocational education, which follows a 2-pillar strategy in Germany with school participation and an applied education in the labor market, there are different possibilities regarding the length of education. Degrees from universities and polytechnics are in many cases not achieved within the standard period. This can lead to an underestimation of the attained years of education and thereby to an incorrect measure of over- and undereducation. Therefore, we use the spell data in the SOEP. Education spells within the sample period are directly observed. For education spells before entering the sample, we use data from the biography questionnaire. Respondents entering the SOEP fill in a questionnaire stating their occupation (e.g. *in education, employed, unemployed*) for each year since their 15th birthday. This information is used to derive the overall years of education respondents have really spent in education³.

5 Results

Table 3: Wage Regressions - Full Sample

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.019*** (0.001)	0.047*** (0.004)	-0.049 (0.047)
Educ Mode	0.075*** (0.001)	0.053*** (0.004)	0.107*** (0.020)
UE Mode	-0.046*** (0.001)	-0.058*** (0.005)	-0.383*** (0.118)
Constant	-1.430*** (0.041)	-1.934*** (0.074)	-1.555*** (0.240)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	100944	100944	100944

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.

Table 3 shows the results of the wage regressions for the pooled sample. The first column shows the results from OLS wage regressions using the OMU framework. The classical OMU finding that there are positive returns to overeducation, which are lower than the positive returns to required education, and negative returns to undereducation is remarkably stable across countries and datasets (Hartog, 2000). It is therefore not surprising that this result can also be found here. The OLS regressions show positive returns to overeducation of 2% per year of overeducation. The returns

³Robustness checks using the mapped years of education variable shows that most results are qualitatively but not always quantitatively comparable. For tables, see Appendix.

to the years of required education are about 7.5% and significantly higher. The negative effect of undereducation is nearly 5% per year. All coefficients are highly significant and in line with previous literature on this topic.

As previously discussed, OLS results are likely biased due to unobserved heterogeneity. This problem is tackled by applying linear fixed effects (Column 2) and IV (Column 3) regressions. Applying fixed effects changes the results significantly. The coefficient for the required education within an occupation becomes smaller (0.053). The coefficients for years of over- and undereducation are bigger than in the OLS regressions. The wage benefit for a year of overeducation is 4.7% and thereby nearly the same as the one for years of required education. Hence, the difference between these two influences, which is commonly found in the OLS literature, nearly vanishes here. The coefficient for years of undereducation is also larger than in the OLS case showing a wage penalty of more than 7% per year. All coefficients remain highly significant. These findings are in line with the fixed effects results by Bauer (2002), who uses the same dataset for a different observation period. These results hint at the validity of the assumption that there is compensation taking place, at least for overeducated workers. Controlling for unobserved heterogeneity in the panel context lets the coefficients for years of required education and overeducation become closer to each other, as expected in this theory. However, the FE method to control for unobserved factors has some weaknesses in this context. Only individuals changing their fit are observed as all non-changers are time-invariant and therefore not regarded in the panel model. Changing the fit means (1) changing the job, (2) an overall shift of the requirements in the job or (3) going back to education. All of these cases are rather special ones, while many cases in which individuals do not change their match over the sample period cannot be observed. The concentration on those changers is problematic as there might be other factors underlying these changes. Therefore, a further model is used to tackle unobserved heterogeneity.

The third column of Table 3 shows the results for the IV regressions. Here, the difference to the OLS results is even bigger. Returns to years of overeducation are negative but insignificant. So, the pattern clearly differs from the OLS and FE findings. The returns to years of required education are higher than in the previous regressions (0.107). The coefficient for years of undereducation is much bigger than in the former regressions and still highly significant. The effect is noticeably big, which can be explained by two arguments. The first one is the problem that IV regressions are less precise than OLS regressions (Wooldridge, 2000). This leads to larger confidence intervals and makes the point estimates less meaningful. The

other one is the relevance of the instruments. While the first-stage regressions⁴ and Angrist-Pischke multivariate F test of excluded instruments (Angrist and Pischke, 2008) suggest that the instruments are strong for all three education components, the value for the undereducation case is much smaller (an F-value of 20) than for the other two match variables. There is a significant correlation of the instruments with the endogenous variable to be seen but this is not as strong as in the overeducation and required education cases. Therefore, we are careful with a interpretation of the magnitude of the undereducation point estimate.

The results of the FE as well as the IV regressions point at a bias in the OLS results. However, the direction of the bias is not unambiguous. While the FE results suggest that OLS results are biased downwards, IV results suggest that the overeducation coefficient is overestimated. This leads to very different conclusions. While the FE findings back the theory of human capital compensation, the IV results hint at a structural problem in the German labor market. Only years of education really required in a job are refunded, while any additional education is not. This means that education which goes over and above the required education is not productive in the labor market, otherwise employers would pay for it. At the same time, overeducation is a very widespread phenomenon. There is a lot of education attained which then remains unproductive in the labor market. This means a huge waste of resources in the educational sector. As the two explanations differ gravely, we shed further light by splitting the sample to see whether there are different patterns for men and women and add an additional strategy.

In the following, the sample is split in a male and a female sample to see if the effects of overeducation are different.

Table 4: Wage Regressions - Male Sample

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.019*** (0.001)	0.041*** (0.005)	-0.002 (0.035)
Educ Mode	0.066*** (0.001)	0.046*** (0.005)	0.119*** (0.017)
UE Mode	-0.043*** (0.002)	-0.045*** (0.006)	-0.164 (0.119)
Constant	-1.461*** (0.053)	-1.928*** (0.093)	-2.018*** (0.214)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	55415	55415	55415

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.

Table 4 shows the results for the male sample. The pattern found for the pooled

⁴The results of the first-stage regressions can be found in the *Appendix*.

sample is mainly reproduced here. The OLS results in column 1 reveal the classical pattern of OMU studies. In the FE regression, the coefficients for required education and overeducation become close to each other, while the undereducation coefficient remains mainly unchanged. In column 3, the IV results can be seen. Again, the overeducation coefficient becomes negative and insignificant, while the required education effect becomes bigger and remains statistically significant. Undereducation remains negative and also becomes larger, however, it is statistically insignificant. Hence, the findings from the pooled sample are nearly the same as for males.

Table 5: Wage Regressions - Female Sample

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.015*** (0.001)	0.049*** (0.006)	-0.258 (0.185)
Educ Mode	0.083*** (0.001)	0.056*** (0.007)	0.135** (0.057)
UE Mode	-0.047*** (0.002)	-0.072*** (0.008)	-0.686** (0.280)
Constant	-1.542*** (0.062)	-1.861*** (0.123)	-1.813** (0.775)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	45529	45529	45529

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.

Table 5 shows results for female respondents. Once again, the pattern found before is confirmed. In the OLS regressions, required education has a larger effect than overeducation while this changes in the FE regression. The overeducation effect becomes larger, the required education effect becomes smaller and both become closer to each other. While this finding again seems to confirm the compensation hypothesis, the IV results show a very different pattern, rather supporting the theory of a general matching problem.

All in all, the results of the wage regressions show some remarkable findings. The OLS wage regressions reveal the expected results. There are positive returns to overeducation and required education while the latter exceeds the former. The coefficient for years of undereducation is negative. The FE results do not confirm this general finding completely as the difference between overeducation and required education is not found anymore. This is very similar to the findings by Bauer (2002) and support the theory explaining educational mismatch as a form of compensation for other forms of ability. The IV results do not support this hypothesis by showing that the effect of matched years of education becomes larger than in the other regressions while the overeducation effect become insignificant, the point estimate even negative. This backs the theory of an allocation problem as many individuals acquire education which then is of no use in the labor market.

As the FE and IV regressions show different results, we apply a third strategy. Here, we model ability directly and include it in the OMU regressions. Unfortunately, we cannot achieve this by using data from the SOEP because there is no clear ability measure included. Instead, we use data from the International Adult Literacy Survey. This is a joint project of the OECD and Statistics Canada⁵. In 1994, representative samples from European and Northern American countries were interviewed with the aim of getting a comprehensive picture of skills among adults, exceeding the measure of educational attainment. These include numeracy as well as literacy proficiency. Using this dataset, which was explicitly designed to measure skills, it is possible to include ability measures in the OMU framework directly. Unlike the SOEP, the IALS data is a cross-section and not a panel. However, including the ability measures directly, the panel dimension is expendable for this step of the analysis. The dataset does not include a continuous income variable but income quintiles. This changes the econometric approach. We apply three different models to guarantee that our results are not driven by the choice of the model. Firstly, we estimate a linear OLS model with the income as 5-digit variable. As the explained variable is the income and not a classical categorical variable, the assumption of linearity is not supposed to cause problems. However, to ensure this, we estimate an ordered logit as second model. In a third step, we apply an interval regression.⁶ As the econometric framework is different, the data set is much smaller and we do not have the whole set of standard wage-regression control variables, we do not interpret coefficients quantitatively. However, they can show the direction of the findings.

Table 6: DESCRIPTIVE STATISTICS - IALS DATA

	IALS data	
	Mean	Std.Dev.
Female	0.47	(0.50)
OE Mode	1.42	(2.28)
UE Mode	0.75	(1.61)
Educ Mode	10.93	(2.88)
Employer unchanged	0.80	(0.40)
Fulltime	0.68	(0.47)
Math Skills	1.71	(0.64)
Reading Skills	1.49	(0.56)
Writing Skills	1.58	(0.60)
N	1025	

Note: Authors' calculations based on IALS.

Table 6 shows the descriptive statistics of the IALS data.⁷ The sample is evenly

⁵A detailed description of the data can be found in Murray et al. (1998).

⁶The income intervals are generated using the SOEP income quintiles.

⁷We have not split the IALS data further into a male and a female sample. This is due to the significantly lower number of observations as compared to the SOEP, with a total of about 1000 observations of working individuals in the IALS. As the SOEP regressions show, the sample split

distributed by sex, about half of the sample are females. Required education is around 11 years, which is lower than in the SOEP sample. This is due to the construction of the variable, as we cannot use a measure as precise as the one derived from the SOEP spell data. The years of required education are closer to the mapped SOEP variable of years of education. 80% of the respondents have not changed their employer in the last 12 month, which serves as a proxy for tenure. The numeracy and literacy skills are self-assessed on a scale from 1 (*excellent*) to 4 (*poor*). Mean literacy skills are a little better than numeracy skills, with reading skills closest to *excellent*.

Table 7: Wage Regressions without ability Controls

	Income		
	(OLS)	(OLogit)	(Intervall)
OE Mode	0.040** (0.016)	0.073*** (0.028)	0.013** (0.006)
Educ Mode	0.100*** (0.015)	0.185*** (0.027)	0.040*** (0.006)
UE Mode	0.024 (0.028)	0.029 (0.049)	0.008 (0.011)
Constant	1.476*** (0.175)	–	6.457*** (0.070)
cut1	–	0.476 (0.305)	–
cut2	–	2.182*** (0.306)	–
cut3	–	3.403*** (0.314)	–
cut4	–	4.741*** (0.332)	–
lnsigma	–	–	-0.927*** (0.030)
Standard Controls	Yes	Yes	Yes
N	1025	1025	1025

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on IALS data.

Table 7 shows the results of the IALS regressions without including the ability controls. The general pattern of the OMU literature and the previous OLS regressions is reproduced. Overeducation has a positive significant coefficient, required education is also positive significant and bigger. Undereducation does not have a negative coefficient here but it still has the lowest point estimate and is insignificant.

Table 8 presents the results controlling for numeracy and literacy skills. In all three models, the effect of overeducation becomes smaller and loses significance, in the OLS and ordered logit models it stays weakly significant on the 10% level, in the interval regression it becomes statistically insignificant. The coefficients of the required education remain positive and highly significant, so the gap between the coefficients of overeducation and required education widens. This differs from the FE findings but is in line with the IV findings.

The results from the IALS sample show that the classical OMU finding can be reproduced without ability controls. When controlling for ability directly, the results does not change the findings gravely. Robustness checks with a split sample of the IALS data show that the same is true here.

Table 8: Wage Regressions - IALS data

	Income		
	(OLS)	(Ologit)	(Intervall)
OE Mode	0.028* (0.016)	0.050* (0.028)	0.009 (0.006)
Educ Mode	0.087*** (0.016)	0.162*** (0.028)	0.035*** (0.006)
UE Mode	0.032 (0.028)	0.044 (0.049)	0.011 (0.011)
Math Skills	-0.142** (0.068)	-0.275** (0.114)	-0.057** (0.026)
Writing Skills	-0.148 (0.098)	-0.261 (0.162)	-0.061 (0.037)
Reading Skills	0.061 (0.100)	0.091 (0.166)	0.034 (0.038)
Constant	2.046*** (0.242)	-	6.666*** (0.095)
cut1	-	-0.616 (0.412)	-
cut2	-	1.109*** (0.409)	-
cut3	-	2.348*** (0.412)	-
cut4	-	3.706*** (0.422)	-
Insigma	-	-	-0.934*** (0.030)
Standard Controls	Yes	Yes	Yes
N	1025	1025	1025

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on IALS data.

change significantly into the direction of the IV results for the SOEP sample. This means they reject the human capital compensation theory. They rather give a further hint at an allocation problem in the labor market with many individuals spending a long time in education only to accumulate human capital which then is not used productively.

6 Conclusion

This paper analysis the causal effects of educational mismatch on wages in Germany using data from the German Socio-Economic Panel. Educational mismatch is defined as a situation in which individuals have more/ less formal education than the modal value within an occupation (empirical definition of mismatch). OLS results confirm the former findings of this field of literature. As more recent studies on the wage effects of over-/undereducation have found, these results might be biased due to unobserved heterogeneity (Bauer, 2002; Korpi and TÅehlin, 2009). Taking this into account, we apply a fixed effects approach, an IV approach and use data from the International Adult Literacy Survey to model skills in the regression framework directly. Results from the FE regressions confirm the results by Bauer (2002) and hint at the validity of the human capital compensation theory. According to this theory, overeducated workers compensate for lower innate ability. With fixed effects regressions, we cannot observe non-changers in the context of education, thus

we extend the econometric strategy. Using an IV approach, we account for the endogeneity of the education variables without the restriction of time-variance of the educational match. Results reject the compensation hypothesis. Over- and under-education become insignificant and only required education has a positive earnings effect. This finding is backed by regressions using data from the IALS. This data set includes measures of numeracy and literacy skills among adults. These can be used to include ability in the OMU framework. The regressions results go to the same direction as those of the IV regressions, with required education as only positive significant influence on earnings.

This study contributes to the general discussion about the causal effects of educational mismatch and the mechanisms behind it. The results obtained here give rise to doubts about the compensation theory as it is rejected by all models but the panel model. However, there is a new pattern showing in the data. Eliminating unobserved heterogeneity, only the effect of required education remains positive and significant. This hints at a problem in the German educational system and its link to the labor market. Results show that there is hardly a positive causal effect of overeducation to be found, which means that this additional human capital is unproductive. Overeducation is a common feature of the German labor market with more than 50% of employees in the situation of an educational mismatch. When the additional education is mainly unproductive, this is a massive waste of resources. Individuals could enter the labor market earlier instead of spending further years in education. The German educational system, which is mainly financed by public expenditures, could be slimmed down and resources allocated to more productive usages. Results show that individuals structurally overinvest in their education as there are not enough jobs in which this education can be used productively. Overeducation plays an important role in Germany showing that there is no general skill shortage in the labor market but an oversupply. Of course this does not rule out the possibility of a shortage of high-skilled individuals in certain fields but this is not true for the labor market in general. Overall, the allocation does not work perfectly. Over the past decades, skill demand has risen significantly with the technological development. At the moment, it seems as if the labor market is saturated and the educational system systematically produces overskilled individuals.

7 Bibliography

References

- Alba-Ramirez, A. (1993). Mismatch in the spanish labor market: overeducation? *Journal of Human Resources*, 259–278.
- Anger, S. e. a. (2011). Documentation pgen - person-related status and generated variables. SOEP documentation.
- Angrist, J. D. and A. B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics* 106(4), 979–1014.
- Angrist, J. D. and J.-S. Pischke (2008, December). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.
- Battu, H., C. R. Belfield, and P. J. Sloane (1999). Overeducation among graduates: a cohort view. *Education Economics* 7(1), 21–38.
- Bauer, T. K. (2002). Educational mismatch and wages: a panel analysis. *Economics of Education Review* 21(3), 221 – 229.
- Butcher, K. F. and A. Case (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics* 109(3), 531–563.
- Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Technical report, National Bureau of Economic Research.
- Duncan, G. J. and S. D. Hoffman (1981). The incidence and wage effects of overeducation. *Economics of Education Review* 1(1), 75 – 86.
- Groeneveld, S. and J. Hartog (2004). Overeducation, wages and promotions within the firm. *Labour Economics* 11(6), 701 – 714.
- Haisken-DeNew, J. and J. Kleibrink (2013). Walking Wounded - The Causal Welfare Loss of Underemployment through Overeducation. –.
- Haisken-DeNew, J. P. and M. H. Hahn (2010). Panelwhiz: Efficient data extraction of complex panel data sets - an example using the german soep. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften* 130(4), 643–654.
- Harmon, C., H. Oosterbeek, and I. Walker (2003). The returns to education: Microeconomics. *Journal of Economic Surveys* 17(2), 115–156.

- Harmon, C. and I. Walker (1995). Estimates of the economic return to schooling for the united kingdom. *The American Economic Review* 85(5), 1278–1286.
- Harmon, C. and I. Walker (1999). The marginal and average returns to schooling in the UK. *European Economic Review* 43(4-6), 879 – 887.
- Harmon, C. and I. Walker (2000). The returns to the quantity and quality of education: evidence for men in england and wales. *Economica* 67(265), 19–35.
- Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review* 19(2), 131 – 147.
- Hartog, J. and H. Oosterbeek (1988). Education, allocation and earnings in the netherlands: Overschooling? *Economics of Education Review* 7(2), 185–194.
- Kiker, B., M. C. Santos, and M. de Oliveira (1997). Overeducation and undereducation: Evidence for portugal. *Economics of Education Review* 16(2), 111 – 125.
- Korpi, T. and M. TÅěhlin (2009). Educational mismatch, wages, and wage growth: Overeducation in sweden, 1974 - 2000. *Labour Economics* 16(2), 183 – 193.
- McGoldrick, K. and J. Robst (1996). Gender differences in overeducation: A test of the theory of differential overqualification. *The American Economic Review* 86(2), pp. 280–284.
- Mincer, J. (1970). The distribution of labor incomes: A survey with special reference to the human capital approach. *Journal of Economic Literature* 8, 1–26.
- Murray, T. S., I. S. Kirsch, and L. B. Jenkins (1998). *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey*. US Department of Education, Office of Educational Research and Improvement.
- Pons, E. and M. T. Gonzalo (2002). Returns to schooling in spain: how reliable are instrumental variable estimates? *Labour* 16(4), 747–770.
- Rubb, S. (2003). Overeducation in the labor market: a comment and re-analysis of a meta-analysis. *Economics of Education Review* 22(6), 621 – 629.
- Rumberger, R. W. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources* 22(1), 24 – 50.
- Sicherman, N. (1991). "overeducation" in the labor market. *Journal of Labor Economics* 9(2), pp. 101–122.

- Sloane, P. J., H. Battu, and P. T. Seaman (1999). Overeducation, undereducation and the british labour market. *Applied Economics* 31(11), 1437–1453.
- Verdugo, R. R. and N. T. Verdugo (1989). The impact of surplus schooling on earnings. *The Journal of Human Resources* 24(4), 629–643.
- Wagner, G. G., J. R. Frick, and J. Schupp (2007). The german socio-economic panel study (soep) – scope, evolution and enhancements. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften* 127(1), 139–169.
- Wooldridge, J. (2000). Introductory econometrics, a modern approach, 2003. *South-Western, Thomson Learning*.

8 Appendix

Table 9: First Stage Regressions - Pooled Sample

	Education Components		
	(OE)	(Educ.Mode)	(UE)
No. of Siblings	-0.128*** (0.004)	-0.129*** (0.004)	0.037*** (0.002)
Yrs without Bio. Par.	0.021*** (0.005)	0.006 (0.004)	0.002 (0.002)
UE Rate at Age 15	0.063*** (0.004)	0.004 (0.003)	-0.025*** (0.002)
Age	0.333*** (0.006)	0.089*** (0.005)	-0.074*** (0.003)
Age Squarred	-0.004*** (0.000)	-0.001*** (0.000)	0.001*** (0.000)
Cohort	-0.126*** (0.026)	0.046** (0.022)	0.021* (0.011)
Nights in Hospital	-0.004** (0.002)	-0.006*** (0.001)	0.001* (0.001)
Children	-0.147*** (0.010)	0.099*** (0.009)	0.021*** (0.004)
Fulltime	0.107*** (0.024)	0.484*** (0.019)	-0.029*** (0.009)
Tenure	-0.052*** (0.001)	0.002** (0.001)	0.009*** (0.000)
Married	-0.524*** (0.024)	-0.155*** (0.018)	0.070*** (0.008)
Female	-0.740*** (0.020)	0.109*** (0.016)	0.187*** (0.007)
Constant	-3.577*** (0.229)	11.432*** (0.192)	1.645*** (0.100)
Year dummies	Yes	Yes	Yes
N	100944	100944	100944

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 2000 - 2011. First stage regressions of IV wage regressions presented in this analysis.

Table 10: First Stage Regressions - Male Sample

	Education Components		
	(OE)	(Educ.Mode)	(UE)
No. of Siblings	-0.154*** (0.006)	-0.135*** (0.005)	0.033*** (0.003)
Yrs without Bio. Par.	0.020*** (0.007)	-0.000 (0.005)	0.005** (0.003)
UE Rate at Age 15	0.067*** (0.006)	-0.004 (0.004)	-0.020*** (0.002)
Age	0.370*** (0.009)	0.102*** (0.007)	-0.067*** (0.004)
Age Squarred	-0.004*** (0.000)	-0.001*** (0.000)	0.001*** (0.000)
Cohort	-0.060 (0.038)	0.033 (0.031)	0.027* (0.015)
Nights in Hospital	-0.005* (0.003)	-0.012*** (0.002)	0.001 (0.001)
Children	-0.107*** (0.014)	0.110*** (0.012)	0.021*** (0.006)
Fulltime	-1.055*** (0.069)	-0.208*** (0.045)	-0.004 (0.018)
Tenure	-0.060*** (0.002)	-0.018*** (0.001)	0.009*** (0.001)
Married	-0.406*** (0.037)	-0.095*** (0.027)	0.065*** (0.011)
o.Female	0.000 (.)	0.000 (.)	0.000 (.)
Constant	-3.850*** (0.330)	11.776*** (0.271)	1.416*** (0.133)
Year dummies	Yes	Yes	Yes
N	55415	55415	55415

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 2000 - 2011. First stage regressions of IV wage regressions presented in this analysis.

Table 11: First Stage Regressions - Female Sample

	Education Components		
	(OE)	(Educ.Mode)	(UE)
No. of Siblings	-0.086*** (0.006)	-0.113*** (0.005)	0.041*** (0.003)
Yrs without Bio. Par.	0.017** (0.007)	0.009* (0.006)	-0.001 (0.003)
UE Rate at Age 15	0.043*** (0.005)	0.001 (0.004)	-0.029*** (0.003)
Age	0.326*** (0.008)	0.098*** (0.007)	-0.085*** (0.004)
Age Squarred	-0.004*** (0.000)	-0.001*** (0.000)	0.001*** (0.000)
Cohort	-0.198*** (0.035)	0.061** (0.030)	0.011 (0.016)
Nights in Hospital	-0.004* (0.002)	0.001 (0.002)	0.002* (0.001)
Children	-0.187*** (0.015)	0.096*** (0.013)	0.027*** (0.007)
Fulltime	0.263*** (0.027)	0.526*** (0.022)	-0.021* (0.011)
Tenure	-0.044*** (0.001)	0.027*** (0.001)	0.009*** (0.001)
Married	-0.618*** (0.031)	-0.213*** (0.024)	0.080*** (0.012)
o.Female	0.000 (.)	0.000 (.)	0.000 (.)
Constant	-3.463*** (0.306)	11.515*** (0.265)	2.123*** (0.152)
Year dummies	Yes	Yes	Yes
N	45529	45529	45529

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 2000 - 2011. First stage regressions of IV wage regressions presented in this analysis.

Table 12: Wage Regressions - Full Sample - Mapped Years of Educ.

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.049*** (0.001)	0.055*** (0.004)	-0.061 (0.451)
Educ Mode	0.079*** (0.001)	0.062*** (0.004)	0.057 (0.100)
UE Mode	-0.055*** (0.001)	-0.061*** (0.004)	-0.655*** (0.245)
Constant	-1.214*** (0.040)	-1.983*** (0.069)	-0.160 (0.976)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	100074	100074	100074

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.

Table 13: Wage Regressions - Male Sample - Mapped Years of Educ.

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.054*** (0.002)	0.052*** (0.005)	-0.005 (0.083)
Educ Mode	0.072*** (0.001)	0.058*** (0.004)	0.093*** (0.022)
UE Mode	-0.048*** (0.001)	-0.057*** (0.004)	-0.217*** (0.064)
Constant	-1.292*** (0.052)	-1.986*** (0.089)	-1.322*** (0.261)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	54721	54721	54721

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.

Table 14: Wage Regressions - Female Sample - Mapped Years of Educ.

	Log. Hourly Income		
	(OLS)	(FE)	(IV)
OE Mode	0.038*** (0.002)	0.061*** (0.006)	1.063 (1.966)
Educ Mode	0.084*** (0.001)	0.068*** (0.006)	-0.577 (1.061)
UE Mode	-0.067*** (0.002)	-0.069*** (0.006)	-2.260 (3.659)
Constant	-1.239*** (0.061)	-1.935*** (0.109)	8.353 (15.328)
Year dummies	Yes	Yes	Yes
Standard Controls	Yes	Yes	Yes
N	45353	45353	45353

Note: Robust standard errors in parentheses. *, ** and *** denote significance level of 10%, 5% and 1% respectively. Estimations based on SOEP data 1991 - 2011. Excluded instruments are the unemployment Rate at the age of 15; Number of siblings; Years spent without natural parents during childhood.