Skill Mismatch and Wage Inequality*

Michael F. Maier†

15th July 2015

Preliminary Version

Abstract

This paper examines the importance of skill mismatch as an explanation for wage inequality across different education groups. The analysis contributes to the view that both the supply and the demand side should be accounted for when investigating wage inequality. The empirical analysis uses a general measure of skill mismatch as well as a specific measure of overeducation among university graduates. The results show that the incidence of skill mismatch is an important explanation factor for wage inequality in Germany across time and across birth cohorts. Especially for the group of university graduates, the difference in wage dispersion among matched and mismatched workers is considerably high. Wage dispersion related to overeducation tend to be less pronounced for younger cohorts of university graduates. This is important from a policy perspective, as the degree of wage dispersion related to overeducation can be considered as a risk in association with the investment in education.

Keywords: wage inequality, mismatch, variance components models, variance function regression

JEL classification: J21, J24, J31

Michael F. Maier, Phone: +49-621-1235-307, E-mail: maier@zew.de

---

*I gratefully acknowledge support from the German Ministry of Education and Research (BMBF) under grant 01PW11001 under the funding framework "The Economics of Science". The views expressed in this article are those of the author and do not necessarily reflect the views of the BMBF. For helpful discussions and comments, I thank Bernd Fitzenberger, Stephan Dlugosz and my colleagues from the department of Labour Markets, Human Resources and Social Policy at ZEW. All opinions and mistakes are my own.

†Centre for European Economic Research (ZEW), Mannheim, P.O. Box 103443, D-68034 Mannheim.

1 Introduction

Many advanced countries have experienced rising wage inequality over the last decades. The empirical literature has developed several explanations for changes in the distribution of wages and the consequential increase in wage inequality (Katz and Autor, 1999; Acemoglu and Autor, 2011). Institutional factors on the one hand and a changing demand for skills on the other hand are the dominant paradigms to explain increasing wage inequality. Abstracting from institutional factors, the increase in the relative demand for high-skilled workers caused the returns to higher post-secondary education to increase relative to the returns to lower educational degrees leading to an increase in overall wage inequality (Gottschalk and Smeeding, 1997; Martins and Pereira, 2004; Dustmann et al., 2009). From that the literature has concluded that the observed increase in the supply of skilled workers is absorbed by an even higher rise in the demand for skilled workers. Another strand of the literature stresses the argument that overall wage inequality is the result of increasing inequality within certain skill or education groups (Juhn et al., 1993; Lemieux, 2006b). This could be either due to the rising heterogeneity of these groups or due to their increasing share among the whole labour force (Machado and Mata, 2005; Lemieux, 2006a).

The literature on within-group respectively residual wage inequality argues that individuals holding the same educational degree might still obtain different returns from their education. This argument refers mainly to supply side factors such as heterogeneity coming from individual characteristics (for example ability, motivation, etc.) which interact with educational attainment and earnings capacity to yield individual-specific returns to education (see for example Card, 1999; Leuven et al., 2004). In contrast to variation in the supply of skills, the impact of differences in the characteristics of demand side variables on returns to education has been analysed. Working in specific industries, occupations or firms might lead to different wages for workers with the same education level (Fitzenberger and Kurz, 2003; Kambourov and Manovskii, 2009; Card et al., 2013).

This paper adds to the literature on wage inequality by examining the prevalence of skill mismatch as a source of heterogeneity within different educational groups. The concept of skill mismatch considers the assignment of skills to specific job tasks, or general requirements of jobs and occupations. Thus, the analysis contributes to the view that supply and demand side factors should be accounted for when investigating wage inequality. Theoretically, the analysis builds on assignment models where heterogeneous workers face different job requirements or job tasks (Sattinger, 1975; Albrecht and Vroman, 2002; Autor and Handel, 2013). The assignment process causes a variety of matches yielding different returns to the same education level. In particular, skill mismatch can be considered as a manifestation of search frictions which results in wage dispersion among workers with the same education level (Gautier and Teulings, 2006; van den Berg and van Vuuren, 2010).

The specific research interest of this paper is to examine whether there is a link at all between skill mismatch and the wage inequality within different education groups and how this relationship evolves over time. In particular, for university graduates the possibility of overeducation can be seen as a risk, since their investment in education would not pay off completely compared to the situation where job requirements are fully met. On the other hand, vocational education comprises more specific skills which bears higher risks not to find an adequate employment compared to an endowment with more general skills (Wolter and Ryan, 2011).

Based on worker self-assessment two types of skill mismatch are considered in this paper. First, mismatch is specified if a worker does not meet the job requirements either with respect to the formal education level or with respect to her training occupation respectively field of study. This captures a broad scale of market frictions. As a second type, overeducation specifically for university graduates is defined as a mismatch, i.e. the job's required formal education level is lower than the actually attained university degree.

The analysis implicitly assumes a sub-optimal assignment of some workers to jobs which do not fit their formal educational attainment or their occupational skills. Technological shocks and changes in the education system might have a general impact on the incidence of such mismatches and on the possibility of subsequent adjustment processes. Some workers are thus systematically more likely to be affected by the impact of skill mismatch on their wage. Furthermore, as the paper focuses on differences between education groups, the substitutability between older and younger workers might be relevant (Card and Lemieux, 2001). Therefore, the analysis considers not only effects over time but also across birth cohorts.

The implications from higher within-wage inequality due to skill mismatch for overall wage inequality are a priori ambiguous. First, higher wage dispersion within an education group could either increase, decrease or have no effect on the mean wage in the group. Second, a change in the mean wage has different implications for overall wage inequality depending at which education level the change occurs. In order to investigate the contribution of skill mismatch on wage inequality, I apply estimation methods which account jointly for the wage inequality between and within different education levels. First, the estimation of variance components generally informs about the importance of mismatch for between- and within-group inequality. Second, variance function regressions reveal the impact of mismatch on both types of wage inequality for each education group.

The results demonstrate a considerable role of general skill mismatch as a contributing factor to wage inequality. The contribution of education to wage inequality is higher among mismatched workers compared to matched workers across year groups and birth cohorts. Differences in wage inequality due to general skill mismatch are predominant among the group of university graduates. The second part of the analysis focuses therefore on the impact of overeducation as a specific measure for skill mismatch on wage inequality among university graduates. The empirical analysis reveals two sources of within-group wage inequality. First, overeducated university graduates receive significant lower returns to education. Second, among mismatched university graduates wage dispersion is significantly higher compared to matched university graduates. The are no considerable differences between matched and mismatched regarding the contribution to changes in wage inequality over time. For younger birth cohorts differences in wage dispersion between matched and overeducated university graduates are lower.

This paper combines the literature on wage inequality with the empirical literature on the incidence and consequences of skill mismatch (see Leuven and Oosterbeek, 2011, for a review). Empirical analyses in that field mainly study the case of overeducation. As a main finding, overeducated individuals receive smaller wage returns compared to individuals whose educational attainment meets the job's requirements. Several studies provide evidence for the mean returns of under- and overeducation in the German context (Daly et al., 2000; Bauer, 2002; Büchel and Mertens, 2004; Kleibrink, 2013). However, the empirical implications of skill mismatch for wage dispersion have been so far rarely investigated. A review by Hartog (2000) points out that the changes in the dispersion of skill supply and skill requirements explain changes in the wage inequality for Portugal between 1982 and 1992. The study by Green and Zhu (2010) shows for the UK an increasing dispersion of returns to higher education related to rising incidence of overeducation during the 1990s and 2000s.

The paper is structured as follows. Section 2 gives a review of the literature on within-group inequality and its results. Section 3 describes the data source, defines the applied measures of skill mismatch and presents the incidence of skill mismatch and its implication on wages. The econometric approach is described in section 4 and estimation results are presented and discussed in Section 5. Conclusions are drawn in Section 6.

2 Theory and evidence on wage inequality

As skills are incorporated in formal education, an important strand of the literature on wage inequality investigates changes in the supply and demand related to different education levels. Wage inequality can either increase because the returns to a certain educational degree increase, or because the wage dispersion within an education group increases. The studies which are reviewed in the following stress the importance of within-group inequality and especially the question why wage inequality within education levels might exist. The review does not incorporate the discussion on labour market institutions such as unions, minimum wages, or the schooling system and their influence on wage inequality.

Much of the increasing wage inequality in the US since the 1990s is due to an increasing dispersion at the upper part of the wage distribution. Autor et al. (2005) show that topend residual wage inequality has been relatively increasing since the late 1980s. In line with this finding, Lemieux (2006a) shows that predominantly the within-group dispersion among highly educated has increased. According to Lemieux (2006b) the increasing dispersion of unobserved skills contributes to rising wage inequality. Unobserved skills are more dispersed among more educated individuals which in turn raises overall wage inequality through higher achievement rates in education. The results are based on a decomposition of the variance term of education in both studies.

In the European context, Budria and Telhado-Pereira (2011) show that for tertiary educated workers wage dispersion is higher and has increased faster since the mid 1980s in several countries. The evidence for Germany is mixed. Only for the lower secondary education category the difference in the returns between the 90th and the 10th quantile is significant. Over time dispersion in the lower secondary education category remained the same while it decreased for upper secondary education and exhibits no trend for tertiary education. Machado and Mata (2001) investigate how an additional year of schooling contributes to the dispersion within educational categories. For Portugal they find that this dispersion has increased between the 1980s and the 1990s. Hartog et al. (2001) confirm this result but suggests that the dispersion of workers with secondary education has grown more substantially than for tertiary education. Machado and Mata (2005) additionally investigate the link between workers' heterogeneity within the same education and a rising overall wage inequality. As more educated workers display higher wage dispersion throughout the period, their increasing share leads to an increase in wage inequality even if the dispersion of returns remains the same over time.

Explanations for the empirical findings on wage inequality can be derived from factors which constitute the supply side or the demand side of the labour market. Looking at the supply side, wage differences among workers with the same formal education can be related to individual factors causing heterogeneity in returns to educations (Card, 1999). As these factors such as ability, motivation or family background are relevant for the worker's productivity, they are inducing individual-specific earning capacities given the educational degree. Leuven et al. (2004), for example, investigate wage differentials resulting from differences in the relative supply of cognitive ability. Taber (2001) relates variations of the college premium to an increasing demand for unobserved skills. Kambourov and Manovskii (2009) explain within-group wage inequality by heterogeneity in occupation-specific human capital. As occupation-specific skills are not fully transferable, workers' mobility across occupations lead to heterogeneity in occupational experience and thus to wage differentials for individuals of the same education-age group.

Apart from supply side factors which are measured by individual characteristics, wage differentials can be a result of differences in the wage setting process on the demand side of the labour market. A study for Germany by Fitzenberger and Kurz (2003) indicates a relationship between human capital and wage differences across industries over the period between the mid 1980s and the mid 1990s. Inter-industry wage differences can be explained by efficiency wage considerations of firms (Krueger and Summers, 1988). Specific human capital might cause some firms to pay wages above the market clearing value to retain workers. Other studies confirm the importance of firm heterogeneity in explaining wage inequality, especially because higher paid workers tend to be sorted to higher paying firms (Abowd et al., 1999; Card et al., 2013).

A recent literature strand studies the task content of jobs or occupations to attain a more subtle view on how technological change affects the demand for skills and thus wage inequality (Acemoglu and Autor, 2011; Firpo et al., 2011; Autor and Handel, 2013). This literature explicitly conceptualises the link between job requirements on the demand side and human capital as a characteristic of the supply side. Wage differentials are a result of an assignment process of skills to job tasks, while skill groups differ in their comparative advantage across tasks and thus receive different returns from performing a task.

Therefore, within-group wage inequality might be the consequence of an assignment process on the labour market which produces a variety of matches between supply and demand (Sattinger, 1975). Albrecht and Vroman (2002), for example, model skill differences across workers and different skill requirements across jobs to explain the increases in wage inequality between low skilled and high skilled workers. The search and matching literature discusses assignment processes of workers to jobs in the specific context of imperfect labour markets due to search and information frictions (Shimer and Smith, 2000).

This analysis builds on the theoretical foundation of the search and matching literature. Their models can be used to illustrate the incidence of skill mismatch and its relation to wage inequality. Workers with the same education level receive different remuneration due to search frictions on the labour market (Gautier and Teulings, 2006; van den Berg and van Vuuren, 2010). In this respect, skill mismatch can be considered as a manifestation of search frictions. Sub-optimal job matches arise due to workers' opportunity costs associated with searching for the perfect job match. Search frictions prevent quick adjustments as workers are not able to change their job immediately. Dolado et al. (2009) consider skill mismatch transitory leading to job-to-job transitions which in turn contribute to equilibrium wage inequality.

3 Incidence for skill mismatch and its implications on wages on the German labour market

The empirical analysis is based on a sample taken from the German Socio-Economic Panel Study (SOEP).[1] The SOEP is a nationally representative survey of German households which

[1] For details on the SOEP, see Wagner et al. (2007).

is conducted annually. The analysis focuses on West Germany over the years 1984 to 2011. The sample is restricted to males who are between 25 and 60 years old within each cross section.[2] The dependent variable is the logarithm of hourly real wage. It is obtained using information on monthly earnings and weekly working hours. To avoid extreme outliers, I exclude individuals that have a wage below the 1st percentile and above the 99th percentile of the wage distribution. The sample is furthermore restricted to dependent workers in order to make sure that the notion of mismatch applies to a sufficiently homogeneous type of employment relationship.[3]

Several papers study the role of a change in the educational composition of the labour force on overall wage inequality (Machado and Mata, 2005; Lemieux, 2006a; Autor et al., 2008; Lindley and Machin, 2011). A considerable expansion of one education group most likely leads to an increase in the heterogeneity within this group, because differences in school quality or differences in ability become more relevant. Apart from other factors, this might per se translate into an increase in wage inequality within this education group (and could in turn simultaneously induce a decrease in wage inequality for a diminishing education group). As this paper focuses primarily on wage inequality within education groups, it is thus important to also consider the magnitude of impact (i.e. the size of the relative share) of each education group on overall wage inequality. For the analysis, I define five education groups which entail different sets of skills relevant for the labour market and which are likewise fairly homogeneous with respect to remuneration of these skills.[4]

Figure 1 depicts the share of education levels for several birth cohorts starting with individuals who were born between 1940 and 1944 until the birth cohort of those born between 1970 and 1979. The bin width measures the size of the cohort indicating the larger cohorts four, five, and six in our sample as the baby boomers of the mid 1950s until mid 1960s in West Germany.

Overall, the labour force has not changed considerably, as consecutive cohorts do not bring in different mixes of education levels. The uppermost segments for example depict the share of university graduates in each cohort. This share has only increased moderately over the consecutive cohorts from almost 20 percent for the oldest cohort to 24 percent for the recent cohort. Some changes can also be found for the workers holding a vocational degree. While the share of those who previously obtained a lower or middle secondary school degree

[2]The analysis excludes females, as it is expected that wage inequality is considerably related to the selective labour supply of this group. Until now, there is no convincing econometric treatment to account for selection at different points in the distribution (see for a discussion Huber and Melly, 2015).

[3]This sample restriction also excludes civil servants.

[4]This definition of education categories for Germany is common in the literature. I distinguish firstly workers who do not have a vocational degree, secondly workers with a vocational degree who hold a degree from lower or middle secondary school, thirdly workers with a vocational degree who hold a degree from upper secondary school, fourthly master craftsman which represent workers with a higher vocational qualification, and fifthly workers with a university degree, comprising graduates from university and applied university.

(vocational) decreases, the share of workers combining a vocational degree with an upper secondary school degree (HS + vocational) increases over cohorts.

This analysis uses two distinct measures of skill mismatch on within-group wage inequality. For both measures I use the direct evidence from the SOEP data which is based on workers' self-assessment. Workers can be mismatched to the skill requirements in two ways. Either they hold a job which requires training in an occupation respectively field of study which they have not been trained for or the job requires a different formal educational level than the worker has attained. The first skill mismatch variable compares the group of workers which match to the skill requirements of their job in both dimensions with the group of workers who do not match to the skill requirements in at least one dimension. Although the characteristics of mismatches in this comparison could be quite different, the measure gives the opportunity to generally examine differences in wage dispersion between matched and mismatched workers in the different education groups.[5]

The empirical literature on skill mismatch mainly studies the case of overeducation. Therefore, the second skill mismatch variable compares a job's required formal education level with the actually obtained education level of the worker. The variable depicts the mismatch only for university graduates to examine a preferably narrow comparison. Thus, the variable defines an university graduate to be overeducated if she works in a job which does not require a university degree. This group is compared to matched university graduates and all other education levels.

According to the literature the incidence of skill mismatch relates to certain determinants which also comprises that individuals may choose to work in an employment where they are mismatched. Generally, the groups of young workers, women, migrants and unmarried workers are more likely to be overeducated (Leuven and Oosterbeek, 2011). Furthermore, empirical findings suggest that less able workers are more likely to become overeducated. For example Chevalier and Lindley (2009) find a negative correlation between the unobserved ability and the incidence of overeducation in the UK, while Büchel and Pollmann-Schult (2004) find for German workers holding a vocational degree a negative effect of school grades on the probability of being overeducated. For university graduates, the incidence of overeducation is found to be varying across field of studies (Berlingieri and Erdsiek, 2012) and by the quality of the university (Robst, 1995; McGuinness, 2003).

Sicherman and Galor (1990) argue that individuals might choose a job where they are overeducated if in this job the probability of promotion is higher. This argument indicates

[5]Generally, all education categories can exhibit both dimensions of skill mismatch, so there should be no systematic accumulation of a specific type of mismatch in a specific education category. With respect to the match in formal education, the measure includes two patterns of mismatch: the case of overqualified workers who hold a higher education degree than required and the case of underqualified workers who hold a lower education degree than required. Per definition, there are no underqualified workers among university graduates and there is no case of overqualification in the group of the least educated workers.

that mismatch is only a temporary state during the career. Empirical studies, however, find overeducation to be rather persistent over the career (Rubb, 2003; Mavromaras and McGuinness, 2012). In particular, if workers are mismatched at the beginning of the career, the probability to change afterwards to adequate employment positions is considerably low (Baert et al., 2013). For Germany, Bauer (2002) reports that only 16 percent change their mismatch status.

Given the evidence on determinants on overeducation the following paragraphs present descriptive statistics to assess how different the groups of matched and mismatched workers are and whether there are changes over time. Table 1 illustrates the incidence for mismatch across year groups and birth cohorts. Column (1) presents the share of general mismatch among all workers. An increasing number of workers are in a job which fits to their learned occupation, respectively their field of study, and their attained education level. This trend is visible across year groups and across birth cohorts. A very tentative interpretation could be that the matching process on the labour market has improved over time. Furthermore, younger birth cohorts could be increasingly educated and trained in alignment with the existing demand on the labour market.

Column (2) depicts the degree of overeducation among the group of university graduates in the sample across year groups and birth cohorts. Different from the incidence of general skill mismatch, the prevalence of mismatch is low, not exceeding a share of about 20 percent. Also the pattern over time and cohorts is different from the case of general skill mismatch, depicting the highest incidence of overeducation for the middle year groups and cohorts. This finding might be related to baby boomer cohorts in our sample (those who were born between 1955 and 1979) who have increasingly attained higher education levels compared to their antecessors. In reaction, this increase in the relative supply of higher educational attainment might have led to an increasing number of university graduates in positions for which they are overeducated.

Table 2 presents the incidence for skill mismatch across age, education and occupation groups. General skill mismatch within different education levels appears quite diverse. Workers with a higher education level are generally less likely to be in a mismatched position. The differences range from a share of 28 percent of mismatched workers among university graduates to a share of 90 percent among those who have no post-secondary education degree.

There are not many differences across age groups with respect to either general skill mismatch or overeducation. The share of general skill mismatch is fairly stable over the life-cycle. The oldest age group depicts with 14 percent the lowest share of overeducated workers which gives a slight indication towards a resolution of mismatch due to overeducation over the life-cycle.[6]

[6]This finding does not change when the share of mismatched workers across age groups is considered separately for each cohort.

Due to shortages in supply and changes in demand, skill mismatch might be differently prevalent in occupations. Table 2 provides the shares of mismatched and overeducated workers across occupation groups at the one-digit level. General skill mismatch is predominant among plant and machine operators as well as among workers in elementary occupations. With regards to the prevalence of overeducation, there is much more heterogeneity across occupation groups. Whereas legislators, senior officials, managers and professionals are mainly matched, overeducation is predominant in service and industry occupations.

Tables 3 and 4 show descriptive evidence for differences in wages with respect to general mismatch and overeducation. With regard to differences across year groups and birth cohorts, mean wages are systematically lower for the group of generally mismatched workers as well as for the group of overeducated workers compared to their matched counterparts. This mean wage difference though does not change over year groups or cohorts for both measures of mismatch.

Table 3 shows, as expected, that mean wages are increasing with the education degree. Interestingly, the mean wage differences between mismatched and matched workers varies across education categories. While among workers with a vocational or a university degree mean wages are lower for the mismatched group, it is higher among workers with an high school degree and vocational degree, mastercraftsman or for workers with no upper secondary degree.

Descriptive evidence on the dispersion of wages measured by the standard deviation reports education as an unambiguous source of differences between matched and mismatched workers. Regarding the group of generally mismatched workers in table 3, for each education category wage dispersion is higher compared to the group of matched workers. Likewise, wage dispersion in table 4 is higher among the group of overeducated workers for each year group and cohort.

Given the variation in the incidence of skill mismatch over time and across cohorts, the question remains to what extent skill mismatch explains variation in the returns to education. For each education group, skill mismatch is expected to contribute to within-group inequality, as mean wages differ between matched and mismatched workers according to the descriptive evidence. Furthermore, wage dispersion is higher within the group of mismatched workers. A higher dispersion in the productivity, due to the fact that a worker's skills does not meet the requirement of job, leads to a higher dispersion in the wage. Likewise, higher search frictions among mismatched workers cause a higher dispersion in reservation wages and thus more variation in actual wages.

The implications from higher within-wage inequality due to skill mismatch on overall wage inequality are conceptually unclear. First, higher wage dispersion within an education group could either increase, decrease or have no effect on the mean wage in the group. Second, the influence of changes in mean wages on overall wage inequality depend on the education level. An increase in the mean wage of workers with the lowest education level would decrease overall wage inequality, ceteris paribus. In contrast, a rising mean wage for the highest education level would increase overall wage inequality, ceteris paribus. Thus, the following analysis investigates the importance of skill mismatch jointly for between-group and within-group inequality and the variation across different education groups. Furthermore, the analysis examines whether the relationship between skill mismatch and returns to education is changing over time and across cohorts and thus gives an explanation for changes in within-group wage inequality.

4 Econometric approach

To assess the impact of skill mismatch on within-group wage inequality, the paper applies two different econometric approaches. First, the analysis explores the contribution of education to overall wage inequality using a variance component model. The goal is to examine how education levels differently contribute to wage inequality for the group of matched and the group of mismatched workers.

The variance component model applied in this paper mainly builds on the analysis of Lemieux (2006b). To capture dispersion in the returns to education for workers with the same degree, a random coefficient model is assumed

$$y_{it} = \alpha_t a_i + \beta_t b_i E_i + \boldsymbol{\gamma_t} c_i \boldsymbol{X}_i + u_{it} \tag{1}$$

where the log wage y_{it} is a linear function of unobserved ability a_i, education E_i, a quadratic experience term $\boldsymbol{X}_i$, and a term u_{it} capturing measurement error. The random coefficient model assumes that person-specific returns b_i and c_i exist besides mean returns to education and experience, β_t and $\boldsymbol{\gamma}_t$. Under the assumption that the random effects a_i, b_i and c_i are uncorrelated and have a mean of one, the expected value of equation 1 conditional on observables yields the common Mincer-type regression:

$$E(y_{it}|E_i, X_i) = \alpha_t + \beta_t E_i + \boldsymbol{\gamma_t X_i}. \tag{2}$$

This equation explains the influence of education on the level of wages assuming that all individuals receive the same returns from education and experience. Taking the conditional variance of wages leads to an equation which models the influence of returns to education and experience on the dispersion of wages:

$$Var(y_{it}|E_i, X_i) = \alpha_t^2\sigma_a^2 + \sigma_b^2\beta_t^2 E_i^2 + \sigma_c^2\boldsymbol{\gamma_t^2 X_i^2}. \tag{3}$$

Equation 3 exhibits the variance component σ_a^2 for unobserved ability, as wells as σ_b^2 and σ_c^2, which represent the person-specific part of the returns to education and experience, respectively. To identify the parameters of equation 3, I follow the proposition by Lemieux (2006b) to jointly estimate the conditional mean and the variance equation by non-linear least squares on the basis of the methods of moments. As the interest of the analysis lies in the differential impacts of skill mismatch, the following equation system is separately estimated for $m = \{0, 1\}$ subgroups defining the group of mismatched and the group of matched workers:

$$y_{itm} = \alpha_{tm} + \boldsymbol{\beta}'_{tm}\boldsymbol{E}_{im} + \boldsymbol{\gamma}'_{tm}\boldsymbol{X}_{im} + u_{itm} \tag{4}$$

$$r_{itm}^2 = \alpha_{tm}^2\sigma_{am}^2 + \sigma_{bm}^2\boldsymbol{\beta}_{tm}^{2'}\boldsymbol{E}_{im}^2 + \sigma_{cm}^2\boldsymbol{\gamma}_{tm}^{2'}\boldsymbol{X}_{im}^2 + v_{itm}. \tag{5}$$

Equation 4 fits the expected mean of wages y_{it}, whereas equation 5 fits the expected variance of wages empirically represented by the squared residual r_{it}^2.

The second goal of the empirical analysis is to examine the dispersion in the returns to different education levels with respect to skill mismatch. The analysis therefore applies as a second econometric approach variance regression functions. Like in Lemieux (2006b), this approach implies to estimate a regression of the first moment and the second moment of log wages as shown in equations 6 and 7:

$$y_{it} = \beta_{0t} + \boldsymbol{\beta}'_{1t}\boldsymbol{E}_i + \beta_{2t}match_i + \boldsymbol{\beta}'_{3t}\boldsymbol{E}_i * match_i + \boldsymbol{\beta}'_{4t}\boldsymbol{X}_i + \varepsilon_{it} \tag{6}$$

$$\rho_{it}^2 = \lambda_{0t} + \boldsymbol{\lambda}'_{1t}\boldsymbol{E}_i + \lambda_{2t}match_i + \boldsymbol{\lambda}'_{3t}\boldsymbol{E}_i * match_i + \boldsymbol{\lambda}'_{4t}\boldsymbol{X}_i + \eta_{it}. \tag{7}$$

Unlike the variance component model, it is possible to assess the impact of each single regressor on the squared wage residual ρ_{it}^2. For this purpose, separate coefficients are independently estimated in the mean and variance regression. The estimation procedure applied in the paper refers to Western and Bloome (2009) who use a maximum likelihood approach. The estimates are obtained in a stepwise procedure. Firstly, the mean regression from equation 6 is fitted which provides the squared residuals ρ_{it}^2. Secondly, equation 7 is estimated by a generalised linear model. This regression model allows to specify the dependent variable as gamma distributed which accounts for the positive right-skewed distribution of the fitted squared residuals.[7] From this regression, the expected value of the dependent variable $\hat{\sigma_i^2}$ is derived. In a third step, the mean regression from equation 6 is re-estimated applying a weighted linear regression with weights $\frac{1}{\hat{\sigma_i^2}}$. The updated residuals are used to evaluate each individual's contribution to the log-likelihood:

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}; y_i) = -0.5[log(\hat{\sigma_i^2}) + (y_i - \hat{y_i})/\hat{\sigma_i^2}] \tag{8}$$

Steps two and three are repeated using the updated estimated parameters for each iteration until convergence is achieved.

5 Estimation Results

5.1 Assessing the importance of skill mismatch for wage inequality

The first part of the empirical analysis examines the importance of education determining wage inequality among matched and mismatched workers. The empirical analysis jointly fits the predicted wage in a mean equation and the squared wage residuals in a variance equation using education and age as explanatory variables according to equations 4 and 5.[8] Thus, the first equation models the wage variation between different education and experience groups, while the second equation explains wage dispersion within those groups. The estimated variance components for education and age in the second regression indicate their relative importance for the residual wage inequality. The estimations are conducted separately for matched workers and for mismatched workers who either hold a job which requires training in an occupation respectively field of study which they have not been trained for or a job which requires a different formal educational level than the worker has attained.

The estimates from the mean regression across year groups in the upper part of table 5 are as expected. Workers with no vocational degree have on average lower wage returns than the comparison group of workers which attained a lower or intermediary school degree and hold a vocational degree[9]. Workers with a high school degree combined with a vocational degree and master craftsmen exhibit similar wage premia compared to the base category. Finally,

[7] The range of the expected value of the fitted squared residuals ρ_{it}^2 is $(0, \infty)$. In order to map this to the range $(-\infty, \infty)$ of the linear predictor of equation 7, I use a log link function.

[8] The predicted wage and the squared wage residuals are derived from a OLS wage regression which uses a flexible specification of all interactions between education levels and the quartic polynomial of age.

[9] This comparison group is chosen because workers with a vocational degree and lower or middle school degree are the predominant education group on the labour market. Furthermore, this group is the most balanced one in terms of the shares of matched and mismatched workers, as can be seen in table 2.

university graduates receive the highest wage returns.

While the estimation results of the mean part of the regression are fairly similar for the groups of matched and mismatched workers across year groups, the estimated variance components indicate some differences between the two groups. The variance component for education among the group of mismatched workers is always significant, which is also the case for matched workers, except for the second year group. Furthermore, the magnitude of the estimates is quite high, indicating that education is an important determinant for wage inequality in both groups. However, for each year group the variance component of education among mismatched workers is higher compared to the group of matched workers. Thus, wage inequality due to education is different with respect to the incidence of general skill mismatch. A time trend for the education variance component is not visible. Only for the second half of the 1990s, wage inequality among matched and mismatched workers is considerably higher than for other periods. This finding might be a result of the joining of the East and West German labour markets after reunification. The variance component with respect to age is also mainly significant across year groups but lower in magnitude. Except for the first year group, the variance component for age is higher among the group of mismatched workers. Overall the results indicate that heterogeneity in the returns to education and experience is an explanatory factor for wage inequality within those groups while the magnitude of the relationship is higher among the group of mismatched workers.

Estimation results across birth cohorts in Table 6 confirm the results across year groups. The contribution of heterogeneity in the returns to education to residual wage inequality is higher among mismatched workers compared to the group of matched workers. This is in particular the case for the second and third birth cohort, where the variance component for education is not significant among matched workers. Only for the most recent birth cohort the impact of variability in the returns to education is higher among matched workers.

As a robustness check, the so far conducted empirical analysis is reviewed by using a linear spline function in years of education at four education groups as an alternative specification to measure educational attainment.[10] Equations 9 to 12 define the variables for the education groups according to the years of education E.

$$E_{lower} = \begin{cases} E, & if\ 0 \leq E \leq 9 \\ 9, & if\ E > 9 \end{cases} \tag{9}$$

[10]The four education groups correspond to the measurement of education categories in the SOEP data. The first education group represents workers which hold at most a lower school degree. The second education group corresponds to workers who attained at most an intermediary school degree and completed afterwards as an apprenticeship degree. The third education group is equivalent to the group of workers who attained a high school degree followed by a completed apprenticeship. The fourth group comprises workers with degrees from a higher technical college or from university.

$$E_{lower+apprenticeship} = \begin{cases} 0 & if\ E \leq 9 \\ E - 9 & if\ 9 < E \leq 11.5 \\ 2.5 & if\ E > 11.5 \end{cases} \tag{10}$$

$$E_{higher+apprenticeship} = \begin{cases} 0 & if\ E \leq 11.5 \\ E - 11.5 & if\ 11.5 < E \leq 14.5 \\ 3 & if\ E > 14.5 \end{cases} \tag{11}$$

$$E_{tertiary} = \begin{cases} 0, & if\ 0 \leq 14.5 \\ E - 14.5, & if\ E > 14.5 \end{cases} \tag{12}$$

The difference in years of education within an education group reflects different pathways which individuals took to attain their highest education degree. The spline function therefore measures differences in returns to an additional year of education comparing different pathways to the same final educational attainment.

Tables 7 and 8 show the results using the linear spline function in years of education across year groups and across birth cohorts, respectively. For example, for mismatched workers in the period between the years 1984 and 1989, an additional year of education yields the highest returns in the second and third education group which represent a lower or intermediary school degree followed by a completed apprenticeship and a high school degree also followed by a completed apprenticeship, respectively. An additional year in tertiary education (fourth education group) implies a lower increase in returns. The evidence of highest returns to years of education for high school attainment and completed apprenticeship applies both for mismatched and matched workers across year groups and cohorts. The increase in the return to a year of apprentices with lower or intermediary school degrees is less pronounced for more recent year groups and cohorts.

Changing the specification with respect to the measurement of educational attainment does not change the importance of the estimated variance component. The heterogeneity in returns to education is still a significant determinant for residual wage inequality and has a higher impact among mismatched workers. The specification of a linear spline function additionally facilitates the interpretation of the terms α_{tm} and σ^2_{am} as the impact of unobserved ability on the level and dispersion of wages.[11] Though the variance component for unobserved ability is mainly significant across year groups and cohorts, the impact of heterogeneity in

[11] Interpreting σ^2_{am} as the unobserved ability component assumes that there is no measurement error in the residual dispersion of wages.

returns to ability on residual wage is comparably less pronounced.

Equation 5 shows that the variance components interact with education and age and their mean returns. Thus, to evaluate the importance of education as a source of wage inequality, I decompose the change of the wage variance between the periods 1984 to 1989 and 2005 to 2010. This allows to distinguish effects of education on between-group and within-group wage inequality and to separate them from effects due to changes in the composition of education and experience. I apply a measure from Lemieux (2006b) where price effects are computed by replacing the respective coefficients in the base period by the coefficients from the end period. This counterfactual tells what the wage variance in the base period would have been if the respective returns had been as the ones in the end period. In contrast, the computed composition effect shows how the wage variance would have been in the base period if education and experience had been as in the end period.

Table 9 shows that the increase in the total wage variance from the base period between the years 1984 to 1989 to the end period between the years 2005 to 2010 is almost two times higher among the group of mismatched workers. However, the counterfactual values for the between-group variance $Var(y_{itm})$ and for the within-group variance $E(r_{itm}^2)$ show a relatively higher increase in within-group inequality which is comparable for the group of matched and mismatched workers. The contribution from changes in the returns to education via the price effects to the changes in between-group and within-group inequality is similarly small for both groups. In particular the changes in within-wage inequality among matched and mismatched workers is due to changes in the composition of education and experience.

5.2 Skill Mismatch among university graduates and wage inequality

So far the analysis indicates that education plays an important role as a source of residual wage inequality predominantly for the group of mismatched workers. The following analysis investigates whether the education levels contribute differently to within-group inequality considering general skill mismatch. For this, mean and variance functions are jointly estimated according to equations 6 and 7. In the following, I concentrate on the results of the variance regressions, as the results from mean regressions resemble the evidence from the first part of the empirical analysis.

Table 10 presents estimated coefficients from variance regressions across year groups. The coefficient of the match variable is significantly negative for each year group. Thus, the group of workers who hold a job for which they do not meet the skill requirements, either with respect to the training occupation respectively field of study or to the formal educational level, face in general a higher degree of wage dispersion. For example, being matched in the years from 1984 to 1989 reduces residual wage dispersion by 20 percent $(1-e^{-0.22} = 0.197)$. Workers who hold a vocational degree together with a lower or middle school degree are used as a comparison group for the interpretation of education coefficients. The results suggest differences across education groups in the contribution of skill mismatch to within-group wage inequality. For workers with a high school degree and a subsequent vocational degree as well as for master craftsmen, there are no differences in wage dispersion between matched and mismatched workers. Workers with no vocational degree at all exhibit different wage dispersion across years groups. In the second half of the 1980s, wage dispersion for mismatched workers was significantly lower, whereas since the year 2000, wage inequality within the group of workers without vocational degree has been significantly higher among matched individuals.

With regard to the university degree, wage dispersion is significantly higher than the wage dispersion for the vocational degree. Across year groups, within-wage inequality is always significantly higher among mismatched university graduates compared to their matched counterparts. Thus, within-wage inequality across years is predominant among university graduates while the contribution of mismatched workers is considerably higher.

A similar picture can be seen across birth cohorts in table 11. The contribution of university degree to the variance of wages is always significantly different from zero. For each birth cohort wage dispersion among mismatched university graduates is higher compared to workers with a vocational degree. Wage dispersion among matched university graduates is significantly lower compared to mismatched university graduates. The difference between both groups is higher for former birth cohorts. Thus, the importance of skill mismatch in determining wage inequality within university education is more important for former cohorts.

The following part of the empirical analysis examines the particular relevance of overeducation for wage inequality among university graduates as the previous results have shown that general skill mismatch is especially relevant for within-group wage inequality among university graduates. For two further reasons the analysis focuses on this particular issue. First, compared to the skill mismatch measure applied in the previous analysis, overeducation is a more precise measure for skill mismatch, relating only to the job requirements regarding the formal education. Second, overeducation among university graduates is economically relevant, because the degree of wage dispersion related to overeducation can be interpreted as an investment risk associated with commencing in tertiary education. The risk is the larger the higher the wage dispersion related to overeducation, assuming that individuals randomly draw matches or mismatches after they have graduated. This is even more relevant if there is no adequate compensation for such risks, as the literature suggests a wage penalty related to overeducation (Leuven and Oosterbeek, 2011).

Table 12 presents mean returns and wage dispersion for the group of overeducated university graduates in contrast to matched university graduates. The effects are compared to the base category of workers with a vocational degree irrespective whether they are matched or mismatched. The upper part of the table depicts the mean returns for overeducated and matched university graduates. Overeducated workers face a wage cut, which more than halves their wages in comparison to matched university graduates. Across year groups, mean returns for overeducated workers remain stable in contrast to rising mean returns among matched university graduates. This widening gap indicates growing wage inequality between matched and mismatched university graduates. Indeed, this evidence relates to previous studies which conjecture falling returns at the lower end of the wage distribution for higher educated workers as a result of overeducation (e. g. Fersterer and Winter-Ebmer, 2003; Martins and Pereira, 2004).

The results from the variance regression in the lower part of Table 12 clearly show that overeducated university graduates exhibit higher wage dispersion than their matched counterparts. For example, the wage dispersion among overeducated university graduates is about 77 percent ($e^{0.57} = 1.768$) higher than among workers holding a vocational degree. Among matched university graduates wage dispersion is for some years not even different compared to the group of workers with a vocational degree. This is in accordance with the positive correlation between overeducation and the dispersion in the return to graduate education found by Green and Zhu (2010) for the UK.

The comparison of variance coefficients across year groups among matched and overeducated workers suggests no clear time trend. Still, for the group of overeducated workers an exceedingly high wage dispersion for the second half of the 1990s can be found. At the same time their mean returns exhibit the lowest value in this period. Very tentatively this could be related to a growing participation of relatively high educated East German workers contributing to higher heterogeneity on the West German labour market.

The results for birth cohorts in Table 13 describe a similar picture. For each cohort, mean returns are lower among overeducated university graduates. Returns for matched university graduates are quite stable, dropping only for the two recent cohorts, whereas among overeducated workers, especially the middle birth cohorts (individuals born between 1950 and 1960), face low returns.

Wage dispersion is higher among overeducated university graduates compared to other education levels. This result holds across year groups. However, the results across birth cohorts for matched and overeducated university graduates show a particular development which suggests the existence of cohort effects. Wage dispersion among overeducated workers decreases for younger cohorts, while wage dispersion among matched workers slightly increases. Compared to workers with a vocational degree, the latter group exhibits lower wage dispersion for the two oldest cohorts but slightly higher wage dispersion for younger cohorts. The decreasing differential in wage dispersion thus indicates a less pronounced contribution of overeducation on within-wage inequality for younger cohorts.

6 Conclusion

This paper explores the relationship between skill mismatch and wage inequality within education groups. Wage inequality is induced by the difference between the groups of matched and mismatched workers with respect to the level and the dispersion of wages. Related to the theoretical literature on search and matching processes, skill mismatch reflects the existence of frictions on the labour market. A measure of general skill mismatch is defined where workers can be mismatched related to the job's requirements either in terms of the required training in an occupation respectively field of study or in terms of the formal educational level. This measure captures a broad scale of frictions.

Variance component models are estimated to measure the contribution of general skill mismatch on wage inequality. The results demonstrate the important role of skill mismatch as a contributing factor to wage inequality. Across year groups and birth cohorts, the contribution of varying returns to education on wage inequality is higher among mismatched workers compared to the group of matched workers. For both groups, changes in their overall wage inequality between the period 1984 to 1989 and the period 2005 to 2010 are mainly ascribed to changes in within-wage inequality. However, price effects due to changes in the returns to educations contribute little to changes in wage inequality. Thus, although mismatch is an important factor for wage inequality, changes over time are mainly caused by composition effects within the groups of mismatched and matched workers.

To specify the contribution of skill mismatch within each education group, variance function regressions are estimated. Differences in wage inequality due to general skill mismatch are predominant among the group of university graduates. Wage dispersion among mismatched university graduates is significantly higher compared to matched university graduates. The analysis focuses thus in a second step on the role of overeducation as a specific measure for skill mismatch on wage inequality among university graduates. The results show that two channels are relevant for the explanation of wage inequality. Overeducated university graduates receive significant lower returns to education and have a significant higher contribution to wage dispersion than matched university graduates. Changes over time and cohorts are less pronounced. Across year groups, the gap in returns to education between matched and mismatched university graduates seems to increase. While relatively stable over year groups, the difference in wage dispersion between matched and mismatched university graduates is lower for younger birth cohorts.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Acemoglu, D. and Autor, D. H. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4B. Elsevier.

Albrecht, J. and Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, 43(1):283–305.

Autor, D. H. and Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31(2):S59–S96.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2005). Rising wage inequality: The role of composition and prices. *NBER Working Paper*, 11628.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in the U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323.

Baert, S., Cockx, B., and Verhaest, D. (2013). Overeducation at the start of the career: Stepping stone or trap? *Labour Economics*, 25:123–140.

Bauer, T. K. (2002). Educational mismatch and wages: A panel analysis. *Economics of Education Review*, 21:221–229.

Büchel, F. and Mertens, A. (2004). Overeducation, undereducation, and the theory of career mobility. *Applied Economics*, 36(8):803–816.

Büchel, F. and Pollmann-Schult, M. (2004). Overeducation and human capital endowments. *International Journal of Manpower*, 25(2):150–166.

Berlingieri, F. and Erdsiek, D. (2012). How relevant is job mismatch for German graduates? *ZEW DP No.12-075*.

Budria, S. and Telhado-Pereira, P. (2011). Educational qualifications and wage inequality: Evidence for Europe. *Revista de Economia Aplicada*, 19:5–34.

Card, D. (1999). The causal effect of education on earnings. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3. Elsevier.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Card, D. and Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2):705–746.

Chevalier, A. and Lindley, J. (2009). Overeducation and the skills of UK graduates. *Journal of Royal Statistical Society Series A*, 172(2):307–337.

Daly, M. C., Büchel, F., and 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.

Dolado, J. J., Jansen, M., and Jimeno, J. F. (2009). On-the-job search in a matching model with heterogeneous jobs and workers. *The Economic Journal*, 119:200–228.

Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the German wage structure. *The Quarterly Journal of Economics*, 124(2):843–881.

Fersterer, J. and Winter-Ebmer, R. (2003). Are Austrian returns to education falling over time? *Labour Economics*, 10:73–89.

Firpo, S., Fortin, N. M., and Lemieux, T. (2011). Occupational tasks and changes in the wage structure. *IZA Discussion Paper*, No. 5542.

Fitzenberger, B. and Kurz, C. (2003). New insights on earnings trends across skill groups and industries in West Germany. *Empirical Economics*, 28:479–514.

Gautier, P. A. and Teulings, C. N. (2006). How large are search frictions? *Journal of the European Economic Association*, 4:1193–1225.

Gottschalk, P. and Smeeding, T. M. (1997). Cross-national comparisons of earnings and income inequality. *Journal of Economic Literature*, 35:633–687.

Green, F. and Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62(4):740–763.

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

Hartog, J., Pereira, P. T., and Vieira, J. A. C. (2001). Changing returns to education in Portugal during the 1980s and 1990s: OLS and quantile regression estimators. *Applied Economics*, 33:1021–1037.

Huber, M. and Melly, B. (2015). A test for the conditional independence assumption in sample selection models. *Journal of Applied Econometrics*, forthcoming.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3):410–42.

Kambourov, G. and Manovskii, I. (2009). Occupational mobility and wage inequality. *Review of Economic Studies*, 76:731–759.

Katz, L. F. and Autor, D. H. (1999). Changes in the wage structure and earnings inequality. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3A. Elsevier.

Kleibrink, J. (2013). Causal effects of educational mismatch in the labour market. *Ruhr Economic Papers No. 421*.

Krueger, A. B. and Summers, L. H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica*, 56(2):259–293.

Lemieux, T. (2006a). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498.

Lemieux, T. (2006b). Postsecondary education and increasing wage inequality. *The American Economic Review*, 96(2):195–199.

Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*. Elsevier.

Leuven, E., Oosterbeek, H., and van Ophem, H. (2004). Explaining international differences in male skill wage differentials by differences in demand and supply of skills. *The Economic Journal*, 114:466–486.

Lindley, J. and Machin, S. (2011). Rising wage inequality and postgraduate education. IZA Discussion Papers 5981, Institute for the Study of Labor (IZA).

Machado, J. and Mata, J. (2001). Earnings functions in Portugal 1982-1994: Evidence from quantile regressions. *Empirical Economics*, 26:115–134.

Machado, J. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20:445–465.

Martins, P. S. and Pereira, P. T. (2004). Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics*, 11(3):355–371.

Mavromaras, K. and McGuinness, S. (2012). Overskilling dynamics and education pathways. *Economics of Education Review*, 31(5):619–628.

McGuinness, S. (2003). University quality and labour market outcomes. *Applied Economics*, 35(18):1943–1955.

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

Rubb, S. (2003). Overeducation: A short or a long run phenomenon for individuals? *Economics of Education Review*, 22(4):389–394.

Sattinger, M. (1975). Comparative advantage and the distribution of earnings and abilities. *Econometrica*, 43:455–468.

Shimer, R. and Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2):343–369.

Sicherman, N. and Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, 98(1):169–192.

Taber, C. R. (2001). The rising college premium in the eighties: Return to college or return to unobserved ability? *Review of Economic Studies*, 68:665–691.

van den Berg, G. J. and van Vuuren, A. (2010). The effect of search frictions on wages. *Labour Economics*, 17:875–885.

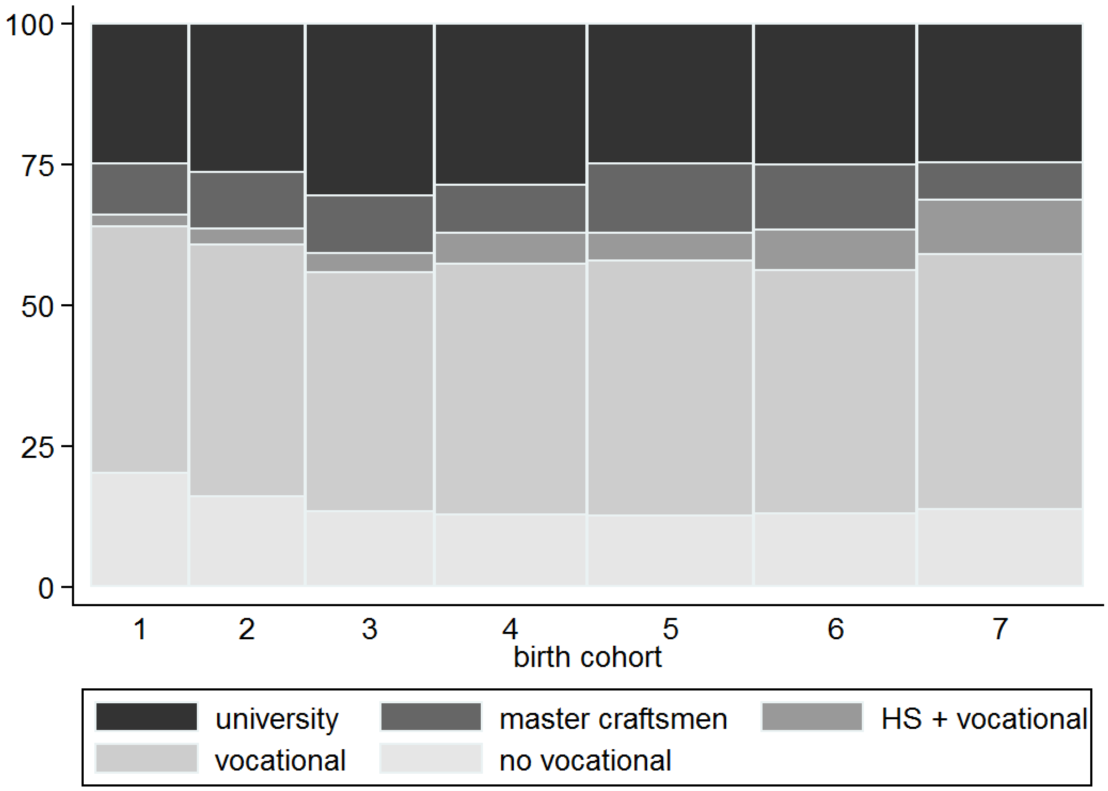
Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – scope, evolution and enhancements. *Schmollers Jahrbuch*, 127 (1):139–169.

Western, B. and Bloome, D. (2009). Variance function regressions for studying inequality. *Sociological Methodology*, 39:293–326.

Wolter, S. C. and Ryan, P. (2011). Apprenticeship. In Hanushek, E. A., Machin, S. J., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume III B. Elsevier.

Figures and Tables

Figure 1: Share of education levels by birth cohort



Note: The figure shows the share of education levels by each birth cohort. Workers in birth cohort 1 are born between 1940 and 1944, in birth cohort 2 between 1945 and 1949, in birth cohort 3 between 1950 and 1954, in birth cohort 4 between 1955 and 1959, in birth cohort 5 between 1960 and 1964, in birth cohort 6 between 1965 and 1969, in birth cohort 7 between 1970 and 1979. The bin width represents the size of the birth cohort.

Table 1: Incidence of mismatch among all workers and incidence of overeducation among university graduates across year groups and cohorts

	share of mismatched workers	share of overeducated university graduates
by year group		
1984-1989	0.54	0.16
1990-1994	0.51	0.20
1995-1999	0.49	0.23
2000-2004	0.43	0.17
2005-2010	0.43	0.16
by cohort		
born 1940-1944	0.53	0.13
born 1945-1949	0.49	0.15
born 1950-1954	0.46	0.17
born 1955-1959	0.45	0.21
born 1960-1964	0.47	0.20
born 1965-1969	0.45	0.18
born 1970-1979	0.42	0.16

Note: The first column depicts the share of mismatched workers by each year group and birth cohort. The particular shares correspond to the sample of all workers which comprises 59308 observations. The second column depicts the share of overeducated university graduates by each year group and birth cohort. The particular shares correspond to the subsample of all university graduates which comprises 15377 observations.

Table 2: Incidence of mismatch among all workers and incidence of overeducation among university graduates across age, education and occupation groups

	share of mismatched workers	share of overeducated university graduates
by age group		
25-30	0.47	0.17
31-40	0.46	0.19
41-50	0.48	0.19
51-60	0.48	0.14
by education degree		
no vocational	0.90	-
vocational	0.48	-
high school + vocational	0.42	-
master craftsmen	0.40	-
university	0.28	0.17
by occupation group		
legislators, senior officials and managers	0.50	0.21
professionals	0.26	0.05
technicians and associate professionals	0.48	0.37
clerks	0.51	0.59
service workers and shop and market salesman	0.52	0.75
skilled agricultural and fishery worker	0.38	0.62
craft and related trade workers	0.37	0.85
plant and machine operators and assembly workers	0.81	0.94
elementary occupations	0.90	0.72

Note: The first column depicts the share of mismatched workers by each age, education and occupation group. The particular shares correspond to the sample of all workers which comprises 59308 observations. The second column depicts the share of overeducated university graduates by each age, education and occupation group. The particular shares correspond to the subsample of all university graduates which comprises 15377 observations.

Table 3: Differences in mean and standard deviation of wages between matched and mismatched workers across year groups, cohorts and education groups

	mismatched		matched	
	mean	sd. dev.	mean	sd. dev.
by year group				
1984-1989	2.59	0.32	2.73	0.34
1990-1994	2.68	0.31	2.80	0.32
1995-1999	2.67	0.34	2.80	0.33
2000-2004	2.77	0.40	2.91	0.39
2005-2010	2.74	0.43	2.90	0.41
by cohort				
born 1940-1944	2.72	0.35	2.93	0.36
born 1945-1949	2.76	0.38	2.91	0.37
born 1950-1954	2.76	0.39	2.93	0.36
born 1955-1959	2.71	0.38	2.88	0.40
born 1960-1964	2.71	0.36	2.85	0.38
born 1965-1969	2.68	0.36	2.81	0.36
born 1970-1979	2.57	0.37	2.69	0.34
by education degree				
no vocational	2.57	0.32	2.55	0.29
vocational	2.66	0.34	2.70	0.31
high school + vocational	2.82	0.38	2.78	0.34
master craftsmen	2.89	0.36	2.81	0.32
university	2.95	0.43	3.11	0.36

Note: The mean and standard deviation are measured in log hourly wages.

Table 4: Differences in mean and standard deviation of wages between matched and overeducated university graduates across year groups and cohorts

	overeducated		matched	
	mean	sd. dev.	mean	sd. dev.
by year group				
1984-1989	2.80	0.39	3.03	0.34
1990-1994	2.86	0.34	3.09	0.29
1995-1999	2.81	0.41	3.05	0.34
2000-2004	2.88	0.44	3.16	0.36
2005-2010	2.84	0.44	3.14	0.38
by cohort				
born 1940-1944	3.05	0.45	3.18	0.30
born 1945-1949	2.97	0.45	3.20	0.32
born 1950-1954	2.83	0.43	3.18	0.32
born 1955-1959	2.82	0.43	3.18	0.35
born 1960-1964	2.86	0.39	3.14	0.35
born 1965-1969	2.81	0.39	3.04	0.36
born 1970-1979	2.69	0.36	2.88	0.41

Note: The mean and standard deviation are measured in log hourly wages.

Table 5: Mean coefficients and variance components among matched and mismatched workers across year groups

	1984-1989	1990-1994	1995-1999	2000-2004	2005-2010
if mismatched:					
constant	1.60***	1.79***	1.72***	1.49***	1.20***
no vocational degree (1)	-0.13***	-0.10***	-0.08***	-0.07***	-0.08***
high school + vocational degree (3)	0.16***	0.08***	0.06***	0.17***	0.17***
master craftsmen (4)	0.13***	0.14***	0.11***	0.15***	0.16***
university degree (5)	0.36***	0.32***	0.29***	0.36***	0.40***
age	0.05***	0.04***	0.04***	0.05***	0.06***
age²/100	-0.05***	-0.04***	-0.04***	-0.05***	-0.05***
variance components:					
constant (σ^2_{a0})	0.03***	0.02***	0.02***	0.04***	0.06***
education (σ^2_{b0})	0.42***	0.44***	1.05***	0.48***	0.46***
age/100 (σ^2_{c0}/100)	0.12	0.19*	0.39***	0.64***	1.00***
if matched:					
constant	1.41***	1.52***	1.55***	1.40***	1.03***
no vocational degree (1)	-0.14***	-0.12***	-0.10***	-0.08***	-0.07***
high school + vocational degree (3)	0.17***	0.10***	0.06***	0.16***	0.17***
master craftsmen (4)	0.13***	0.13***	0.10***	0.15***	0.16***
university degree (5)	0.36***	0.33***	0.28***	0.36***	0.38***
age	0.05***	0.05***	0.05***	0.05***	0.07***
age²/100	-0.05***	-0.05***	-0.05***	-0.05***	-0.06***
variance components:					
constant (σ^2_{a1})	0.03***	0.03***	0.03***	0.04***	0.09***
education (σ^2_{b1})	0.17***	-0.06	0.30***	0.06**	0.07**
age/100 (σ^2_{c1}/100)	0.29***	0.18***	0.00	0.33***	0.08*

Note: The table reports estimates of mean coefficients and variance components from an empirical model which jointly fits the conditional mean of log hourly wages and the conditional variance of log hourly wages via nonlinear least squares. Education is measured by dummy variables using workers with a vocational degree as the comparison group. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 6: Mean coefficients and variance components among matched and mismatched workers across birth cohorts

	born 1940-1944	born 1945-1949	born 1950-1954	born 1955-1959	born 1960-1964	born 1965-1969	born 1970-1979
if mismatched:							
constant	0.32***	1.36***	1.25***	1.16***	1.20***	1.44***	0.87***
no vocational degree (1)	-0.16***	-0.10***	-0.17***	-0.10***	-0.01***	-0.04***	-0.04***
high school + vocational degree (3)	0.36***	0.19***	0.10***	0.15***	0.18***	0.09***	0.15***
master craftsmen (4)	0.11***	0.19***	0.08***	0.14***	0.13***	0.17***	0.13***
university degree (5)	0.39***	0.40***	0.32***	0.37***	0.37***	0.32***	0.29***
age	0.09***	0.05***	0.06***	0.06***	0.07***	0.06***	0.09***
$age^2/100$	-0.08***	-0.04***	-0.05***	-0.06***	-0.07***	-0.07***	-0.12***
variance components:							
constant (σ^2_{a0})	0.23	0.00	0.01*	0.02***	0.02***	0.02***	0.00***
education (σ^2_{b0})	0.35***	0.28***	0.74***	0.66***	0.45***	0.56***	0.27*
age /100 ($\sigma^2_{c0}/100$)	0.26***	1.66***	1.02***	0.80***	0.88***	1.18***	1.05***
if matched:							
constant	0.82***	1.29***	0.99***	1.17***	1.15***	1.17***	1.03***
no vocational degree (1)	-0.15***	-0.10***	-0.14***	-0.08***	-0.01	-0.02***	-0.02**
high school + vocational degree (3)	0.33***	0.20***	0.09***	0.13***	0.17***	0.08***	0.13***
master craftsmen (4)	0.12***	0.20***	0.08***	0.14***	0.13***	0.15***	0.12***
university degree (5)	0.40***	0.40***	0.32***	0.38***	0.37***	0.30***	0.29***
age	0.07***	0.05***	0.07***	0.06***	0.07***	0.07***	0.08***
$age^2/100$	-0.06***	-0.04***	-0.06***	-0.06***	-0.07***	-0.08***	-0.09***
variance components:							
constant (σ^2_{a1})	0.07***	0.01*	0.05***	0.03***	0.03***	0.03***	0.02
education (σ^2_{b1})	-0.11***	-0.07	0.06	0.12***	0.12***	0.02	0.70***
age /100($\sigma^2_{c1}/100$)	0.33***	1.00***	0.30***	0.54***	0.55***	0.76***	0.68***

Note: The table reports estimates of mean coefficients and variance components from an empirical model which jointly fits the conditional mean of log hourly wages and the conditional variance of log hourly wages via nonlinear least squares. Education is measured by dummy variables using workers with a vocational degree as the comparison group. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 7: Mean coefficients and variance components among matched and mismatched workers across year groups using a linear spline function in education

	1984-1989	1990-1994	1995-1999	2000-2004	2005-2010
if mismatched:					
constant	1.28***	1.48***	1.47***	1.23***	1.06***
E_{lower}	0.01***	0.01***	0.01***	0.01*	-0.00
$E_{lower+apprenticeship}$	0.07***	0.06***	0.04***	0.04***	0.04***
$E_{higher+apprenticeship}$	0.06***	0.05***	0.05***	0.07***	0.08***
$E_{tertiary}$	0.03***	0.04***	0.04***	0.04***	0.04***
age	0.05***	0.05***	0.04***	0.05***	0.06***
age^2/100	-0.05***	-0.05***	-0.04***	-0.05***	-0.06***
variance components:					
constant (σ^2_{a0})	0.02***	0.02***	0.01*	0.05***	0.08***
education (σ^2_{b0})	1.17***	1.11***	2.41***	0.80***	0.74***
age/100 (σ^2_{c0}/100)	0.18**	0.19**	0.41***	0.57***	0.46***
if matched:					
constant	0.92***	1.14***	1.18***	1.13***	0.50***
E_{lower}	0.04***	0.03***	0.03***	0.02	0.05***
$E_{lower+apprenticeship}$	0.05***	0.04***	0.04***	0.03***	0.03***
$E_{higher+apprenticeship}$	0.08***	0.06***	0.05***	0.08***	0.09***
$E_{tertiary}$	0.03***	0.04***	0.03***	0.03***	0.03***
age	0.06***	0.05***	0.05***	0.06***	0.07***
age^2/100	-0.06***	-0.05***	-0.04***	-0.05***	-0.06***
variance components:					
constant (σ^2_{a1})	0.02	0.05***	0.06	0.06***	0.28***
education (σ^2_{b1})	0.33***	-0.10	0.65***	0.15**	0.13***
age/100 (σ^2_{c1}/100)	0.26***	0.17***	0.35	0.31***	0.08**

Note: The table reports estimated variance components from an empirical model which jointly fits the conditional mean of log hourly wages and the conditional variance of log hourly wages via nonlinear least squares. Education is measured by a linear spline function in years of education. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 8: Mean coefficients and variance components among matched and mismatched workers across birth cohorts using a linear spline function in education

	born 1940-1944	born 1945-1949	born 1950-1954	born 1955-1959	born 1960-1964	born 1965-1969	born 1970-1979
if mismatched:							
constant	0.08**	1.24***	1.00***	1.09***	1.14***	1.36***	0.98***
E_{lower}	0.01**	0.00	0.01	-0.01	0.00	0.00	0.01
$E_{lower+apprenticeship}$	0.09***	0.06***	0.08***	0.05***	0.01***	0.02***	0.01***
$E_{higher+apprenticeship}$	0.09***	0.08***	0.05***	0.07***	0.08***	0.05***	0.06***
$E_{tertiary}$	0.02***	0.03***	0.04***	0.03***	0.03***	0.05***	0.04***
age	0.09***	0.05***	0.06***	0.07***	0.07***	0.06***	0.08***
$age^2/100$	-0.08***	-0.04***	-0.06***	-0.07***	-0.08***	-0.07***	-0.09***
variance components:							
constant (σ^2_{a0})	1.95***	0.00	0.00	0.02**	0.02***	0.02***	0.02
education (σ^2_{b0})	0.77***	0.73***	1.57***	1.03***	0.83***	1.11***	0.76***
age /100 ($\sigma^2_{c0}/100$)	0.23***	1.63***	0.84***	0.76***	0.84***	1.10***	1.37***
if matched:							
constant	0.46***	1.23***	0.55***	0.85***	1.10***	0.99***	1.18***
E_{lower}	0.02***	0.01	0.05***	0.03***	0.00	0.00	0.03
$E_{lower+apprenticeship}$	0.07***	0.06***	0.03***	0.01***	0.01***	0.01***	0.01***
$E_{higher+apprenticeship}$	0.11***	0.09***	0.07***	0.08***	0.08***	0.05***	0.05***
$E_{tertiary}$	0.00	0.02***	0.02***	0.03***	0.03***	0.04***	0.03***
age	0.07***	0.05***	0.07***	0.06***	0.07***	0.08***	0.05***
$age^2/100$	-0.06***	-0.04***	-0.06***	-0.06***	-0.08***	-0.09***	-0.04***
variance components:							
constant (σ^2_{a1})	0.23	0.01*	-0.01	0.02	0.03***	0.04***	-0.06
education (σ^2_{b1})	-0.08	-0.05	0.24***	0.26***	0.23***	0.24*	1.60***
age /100($\sigma^2_{c1}/100$)	0.28***	1.22***	0.32***	0.56***	0.49***	0.58***	1.49***

Note: The table reports estimated variance components from an empirical model which fits estimates the conditional mean of log hourly wages and the conditional variance of log hourly wages via nonlinear least squares. Education is measured by a linear spline function in years of education. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 9: Decomposition of the change in the variance of wages between 1984-1989 and 2005-2010

	change in variance		
	between-group	within-group	total
if mismatched:			
price effects			
no vocational degree (1)	-0.004	-0.001	-0.005
high school + vocational degree (3)	0.000	0.000	0.000
master craftsmen (4)	0.001	0.001	0.001
university degree (5)	0.002	0.001	0.003
age	0.012	0.000	0.012
constant	-	-0.026	-0.026
total price effect	0.011	-0.025	-0.014
composition effect	0.013	0.086	0.098
total change	0.024	0.060	0.084
if matched:			
price effects			
no vocational degree (1)	-0.001	0.000	-0.001
high school + vocational degree (3)	0.000	0.000	0.000
master craftsmen (4)	0.000	0.001	0.001
university degree (5)	0.003	0.001	0.004
age	0.019	0.001	0.020
constant	-	-0.026	-0.026
total price effect	0.021	-0.023	-0.002
composition effect	-0.008	0.057	0.049
total change	0.013	0.034	0.047

Table 10: Variance regression considering general skill mismatch across year groups

	1984-1989	1990-1994	1995-1999	2000-2004	2005-2010
no vocational degree (1)	-0.14**	-0.09	-0.09	0.05	0.06
high school + vocational degree (3)	-0.27	0.19	0.24*	0.15*	-0.19**
master craftsmen (4)	0.19*	0.11	0.13	0.00	0.02
university degree (5)	0.56***	0.35***	0.61***	0.34***	0.31***
match	-0.22***	-0.16***	-0.28***	-0.24***	-0.34***
interaction term (1)	-0.19	-0.21	0.16	0.58***	0.55**
interaction term (3)	0.30	-0.38*	-0.26	-0.02	0.50***
interaction term (4)	-0.01	0.09	0.07	0.07	0.07
interaction term (5)	-0.12	-0.31***	-0.24**	-0.19***	-0.20***
age	0.11***	0.04	0.06***	-0.00	-0.02
$age^2/100$	-0.13**	-0.04	-0.06***	0.02	0.02
constant	-4.49***	-3.37***	-3.65***	-2.31***	-1.82***

Note: The table reports estimated coefficients from a variance regression of log hourly wages which is jointly estimated with a mean regression of log monthly wages by iterating a two-stage model via maximum likelihood. The comparison group consists of workers holding a vocational degree together with a lower or middle school degree. The variable *match* takes on the value 0 if the worker is mismatched and 1 if the worker is matched. Interaction terms specify the joint effect of the match indicator and the respective education level. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 11: Variance regression considering general skill mismatch across cohorts

	born 1940-1944	born 1945-1949	born 1950-1954	born 1955-1959	born 1960-1964	born 1965-1969	born 1970-1979
no vocational degree (1)	-0.20**	-0.48***	-0.38***	0.10	0.05	-0.08	0.14
high school + vocational degree (3)	-0.75***	0.79***	-0.20	0.01	-0.03	0.03	0.15
master craftsmen (4)	0.03	-0.35***	-0.10	0.43***	0.24***	-0.17*	0.30
university degree (5)	0.55***	0.24**	0.41***	0.55***	0.31***	0.31***	0.24**
match	-0.03	-0.28***	-0.47***	-0.18***	-0.25***	-0.21***	-0.37***
interaction term (1)	-0.24	0.12	0.02	0.32	-0.06	0.11	0.30
interaction term (3)	-0.27	-1.01***	0.79***	0.11	0.16	0.13	-0.33*
interaction term (4)	0.02	0.22	0.41***	-0.18	-0.09	0.12	-0.23
interaction term (5)	-0.73***	-0.37***	-0.25***	-0.33***	-0.07	-0.18*	0.35***
age	-0.33***	-0.18***	-0.12***	0.03	0.08***	0.17***	0.19*
$age^2/100$	0.35***	0.22***	0.16***	-0.00	-0.06	-0.18***	-0.19
constant	5.28***	1.27	-0.06	-3.65***	-4.37***	-5.79***	-6.16***

Note: The table reports estimated coefficients from a variance regression of log hourly wages which is jointly estimated with a mean regression of log monthly wages by iterating a two-stage model via maximum likelihood. The comparison group consists of workers holding a vocational degree together with a lower or middle school degree. The variable *match* takes on the value 0 if the worker is mismatched and 1 if the worker is matched. Interaction terms specify the joint effect of the match indicator and the respective education level. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 12: Variance regression considering overeducation among university graduates across year groups

	1984-1989	1990-1994	1995-1999	2000-2004	2005-2010
mean coefficients:					
overeducated	0.14***	0.16***	0.10***	0.13***	0.13***
matched	0.36***	0.36***	0.34***	0.40***	0.45***
variance coefficients:					
overeducated	0.57***	0.33***	0.65***	0.53***	0.44***
matched	0.34***	0.01	0.23***	0.04	-0.04

Note: The table reports estimated coefficients from a variance regression of log hourly wages which is jointly estimated with a mean regression of log monthly wages by iterating a two-stage model via maximum likelihood. The comparison group consists of workers holding a vocational degree together with a lower or middle school degree. ***, **, * indicate significance at 1-, 5-, and 10-percent level.

Table 13: Variance regression considering overeducation among university graduates across cohorts

	born 1940-1944	born 1945-1949	born 1950-1954	born 1955-1959	born 1960-1964	born 1965-1969	born 1970-1979
mean coefficients:							
overeducated	0.30***	0.20***	0.05**	0.10***	0.16***	0.12***	0.09***
matched	0.41***	0.43***	0.38***	0.44***	0.41***	0.34***	0.31***
variance coefficients:							
overeducated	0.77***	0.40***	0.60***	0.64***	0.41***	0.31***	0.22*
matched	-0.20***	-0.27***	-0.07	0.11**	0.10**	0.09*	0.47***

Note: The table reports estimated coefficients from a variance regression of log hourly wages which is jointly estimated with a mean regression of log monthly wages by iterating a two-stage model via maximum likelihood. The comparison group consists of workers holding a vocational degree together with a lower or middle school degree. ***, **, * indicate significance at 1-, 5-, and 10-percent level.