Discussion Paper No. 15-022

Mobility Across Firms and Occupations Among Graduates from Apprenticeship

Bernd Fitzenberger, Stefanie Licklederer, and Hanna Zwiener

Zentrum für Europäische Wirtschaftsforschung GmbH

Centre for European Economic Research

Discussion Paper No. 15-022

Mobility Across Firms and Occupations Among Graduates from Apprenticeship

Bernd Fitzenberger, Stefanie Licklederer, and Hanna Zwiener

Mobility across Firms and Occupations among Graduates from Apprenticeship

Bernd Fitzenberger[b,c,d,e,f,g,*], Stefanie Licklederer[a], Hanna Zwiener[b]

a: University of Freiburg; b: Humboldt-University Berlin; c: ZEW; d: IZA, e: IFS, f: ROA, g: CESifo

March 2015

Abstract: Distinguishing carefully between mobility across firms and across occupations, this study provides causal estimates of the wage effects of mobility among graduates from apprenticeship in Germany. Our instrumental variables approach exploits variation in regional labor market characteristics. Pure firm changes and occupation-and-job changes after graduation from apprenticeship result in average wage losses, whereas an occupation change within the training firm results in persistent wage gains. For the majority of cases a change of occupation involves a career progression. In contrast, for job switches the wage loss dominates.

Keywords: Apprenticeship Training, Job Mobility, Occupational Mobility, Wages
JEL Classification: J62, J24, J30, J31.

We thank the editor and two anonymous referees for helpful comments. This paper was presented at the TREE conference 2013 in Basel and at the EALE conference 2014 in Lubljana. We are grateful for useful comments received at these conferences, and we benefited from insightful discussions with Christian Dustmann, Heike Solga, Alexandra Spitz-Oener, and Markus Zimmerman. We acknowledge financial support by the Ministry of Science and Art in Baden-Württemberg through the project "Die Wirkung des Bildungssystems auf den Übergang von Schule in Beruf - Evaluation des Pilotprojekts Erfolgreich in Ausbildung" (No. 31-7532.20/068) and the DFG through the project "Transitions in Education and Vocational Training and Labor Market Outcomes" (No. Fi 692/11-1) as well as through the Research Training Group 1659 "Interdependencies in the regulation of markets". All errors are our sole responsibility.

* Corresponding author: Humboldt-Universität zu Berlin, School of Business and Economics, Institute of Statistics and Econometrics, Chair of Econometrics, Spandauer Straße 1, 10099 Berlin, E-mail: fitzenbb@hu-berlin.de.

1 Introduction

A large literature has documented sizeable mobility across firms and occupations in the US and Western European labor markets.[1] During the time period of 1979–2006 monthly occupational mobility rates in the US were at about 3.5% of overall employment – even higher than the 3.2% average rate of job mobility across firms (Moscarini and Thomsson, 2007). For Denmark, Groes et al. (forthcoming) report that the annual occupational mobility rate lies close to 20%. While a large literature emphasizes the loss of firm-specific or occupation-specific human capital (e.g. Kambourov and Manovskii, 2008; Gathmann and Schönberg, 2010; von Wachter and Bender, 2006; von Wachter et al. 2009), mobility may very well be associated with career progression or job shopping in labor markets with frictions (Topel and Ward, 1992), thus resulting in wage gains after mobility (Groes et al., forthcoming; Fitzenberger and Spitz-Oener, 2004; Fitzenberger and Kunze 2005). Furthermore, mobility across firms and occupations may be an important adjustment mechanism in a dynamic labor market. For instance, the tasked-based approach introduced by Autor et al. (2003) argues that there is a decline in the demand for routine intensive occupations, to which workers may adjust through occupational mobility (Cortes, 2012; Gathmann and Schönberg, 2010). Most of the literature referred to so far is restricted to an analysis of either job mobility or occupational mobility.[2] Based on high-quality administrative data, our analysis allows to distinguish the wage effects of job mobility and occupational mobility. In Germany, vocational training in an apprenticeship involves a job in the training firm and training in a specific occupation. Our analysis estimates the wage effects of mobility right after graduation from an apprenticeship in Germany.

Graduates from apprenticeship constitute a large share of the German workforce, and the apprenticeship combines practical training at the training firm with part-time school-based training, thus involving both general and occupation-specific skills.[3] Graduates may continue to work as a regular employee in their training firm, possibly in their training occupation or in another occupation. At graduation, there is no employment protection in the training firm. Given the combination of firm-based and school-based training the skills acquired during an apprenticeship are often thought to be largely transferable across jobs, thus allowing for worker mobility after graduation from apprenticeship (Euwals and Winkelmann, 2002, 2004; Clark and Fahr, 2002). Indeed, retention rates are only about 60-75% of all graduates (Bougheas and Georgellis, 2004; Euwals and Winkelmann, 2004; von Wachter and Bender, 2006). The high mobility after graduation is a particularly interesting case to analyze. On the one hand, a change across firms involves the loss of the training investment for the training firm (Wolter

[1]Among others, see for the US: Topel and Ward (1992), Neal (1999), Moscarini and Thomsson (2007), Kambourov and Manovskii (2008, 2009); for France: Lalé (2012); for Germany: Fitzenberger and Kunze (2005), von Wachter and Bender (2006), von Wachter et al. (2009), Gathmann and Schönberg (2010); for Denmark: Groes et al. (forthcoming); and for Germany and the UK: Longhi and Brynin (2010).

[2]Studies which investigate mobility across firms and occupations include Neal (1999), Kambourov and Manovskii (2008), Longhi and Brynin (2010), or Müller and Schweri (forthcoming).

[3]For a detailed description of the German dual system of vocational training see e.g. Hoeckel and Schwartz (2010). A graduate from apprenticeship obtains a certified degree in one out of 350 training occupations. In 2009 about 60% of German youths aged between 16 and 24 years entered vocational training (Gericke et al., 2011).

and Ryan, 2011) and a change of occupation (firm) may imply a loss of the occupation- (firm-) specific human capital acquired through apprenticeship training (Kambourov and Manovskii, 2008; Gathmann and Schönberg, 2010). On the other hand, firms may use the apprenticeship as a screening device for young workers, and they may only retain those apprentices after graduation who perform well (Euwals and Winkelmann, 2002; Werwatz 2002; von Wachter and Bender, 2006). Graduates from apprenticeship may search for better job offers as a form of career progression (Topel and Ward, 1992; von Wachter and Bender, 2006; Fitzenberger and Spitz-Oener 2004), and non-training firms may make attractive job offers to well trained graduates from apprenticeship, i.e. there is an incentive for poaching (Wolter and Ryan, 2011). A better match for the employee may also involve working in a different occupation within the training firm, an issue which has received little attention in the literature so far.

Several studies analyze the individual labor market effects of mobility after apprenticeship – mainly for Germany and Switzerland. However, the existing studies typically do not distinguish between a pure firm switch without occupation switch and a simultaneous switch of firm and occupation (a complex switch according to Neal, 1999), and occupational mobility within the training firm is typically ignored. Von Wachter and Bender (2006) estimate a large immediate negative causal wage effect of a switch of firm after graduation. However, the negative effect vanishes five years afterwards. The study emphasizes that OLS estimates of the wage effects after five years are severely downward biased due to the negative selection of the firm switchers. In contrast, a negative wage effect of a firm switch is found by Bougheas and Georgellis (2004) for a six year period after training, and other studies find small positive wage effects of leaving the training firm (Euwals and Winkelmann, 2004; Göggel and Zwick, 2012). For Switzerland, Müller and Schweri (2009, forthcoming) find no wage differential between stayers and pure firm switchers one year after graduation from apprenticeship. Göggel and Zwick (2012) find a small negative immediate wage effect of a switch in occupation. Bougheas and Georgellis (2004) find a positive wage effect of a switch in occupation without switch of firm relative to stayers during the first six years after training. A simultaneous switch of occupation and firm is associated with wage losses both in Germany (Bougheas and Georgellis, 2004) and in Switzerland (Müller and Schweri, forthcoming).

There exist some further studies considering mobility later during the career among prime-aged German workers holding an apprenticeship degree that provide further insights into the topic. Dustmann and Schönberg (2012) estimate the transferability of skills obtained through apprenticeship training for a sample of male workers. The survey data contains information provided by workers on how well they can apply skills obtained through apprenticeship training in their current job. Dustmann and Schönberg (2012) estimate that relative to stayers, pure firm switchers can apply 4.5% less of these skills in their current job. In their current job within-firm occupation switchers can use 8.6% less of their skills obtained through apprenticeship training, while across-firm occupation switchers can use up to 34% less of these skills. These results suggest that occupational mobility is associated with large losses in human capital, especially if a simultaneous firm change occurs. In contrast to this, Clark and Fahr (2002) find that only changes across 1-digit occupations entail wage losses while within 1-digit occupations the skills obtained through apprenticeship training are transferable. Regarding the wage effects of occupational mobility among prime-aged workers, other studies also draw a rather positive picture of occupation changes as they find average wage gains (Werwatz, 2002; Fitzenberger and Spitz-Oener, 2004; Fitzenberger and Kunze, 2005). Werwatz (2002) finds a negative wage effect of occupational mobility only for the small group of occupation switchers who state that in their current job they can only apply very little or none of the skills obtained through training. Similarly, Gathmann and SchÃűnberg (2010) find that the wage loss implied by a switch in occupation increases with the differences in task inputs between the source occupation and the target occupation.

Our study provides causal estimates of the wage effects of mobility across firms and occupations among graduates from apprenticeship in Germany. Our data consist of about 14.200 male graduates who completed apprenticeship training during the period of 1992-1997. We contribute both to the literature on the economic effects of occupational mobility as well as to the literature on labor mobility among young workers. Apprenticeship graduates are very likely selected into the different types of mobility based on unobservables, which may bias OLS estimates. We therefore employ an instrumental variables approach exploiting variation in regional labor market characteristics to estimate the causal short-term and long-term effects of mobility after apprenticeship on wages. We show that local labor market characteristics, such as the unemployment rate, labor market tightness and mobility behavior of the local workforce, are significantly correlated to the incidence of different types of mobility after graduation from apprenticeship.[4] Our paper also contributes to the discussion as to whether an apprenticeship, as part of the school-to-work transition, prepares well for a successful entry into the labor market. This aspect has been the subject of an intensive debate in several EU countries who are discussing reforms of vocational training in order to reduce the high level of youth unemployment (BMBF, 2012; The Economist; 2013).

We contribute to the literature on occupational mobility among young workers by carefully distinguishing between two different dimensions of mobility: mobility across firms and mobility across occupations. The literature on job mobility among young workers as well as the literature on occupational mobility typically does not distinguish these two dimensions and occupational mobility within the training firm is typically ignored.[5] Studies on occupational change often only consider across-firm occupation changes as valid, while within-firm occupation changes are perceived as "spurious" and stemming from coding errors (see e.g. Lalé, 2012, and Longhi and Brynin, 2010). In our analysis, we use high-quality German administra-

[4]Other studies on the individual labor market effects of mobility after apprenticeship in Germany deal with the endogeneity issue using a selection correction approach (Werwatz, 2002; Bougheas and Georgellis, 2004; Fitzenberger and Spitz-Oener, 2004; Müller and Schweri, forthcoming, for Switzerland) or they consider only displaced workers (Clark and Fahr, 2002; Bougheas and Georgellis, 2004; Göggel and Zwick, 2012). Von Wachter and Bender (2006) use differences in firm-specific retention rates as exogenous variation. Neumark (2002) analyzes job mobility among young workers in the U.S. using local unemployment rates as instruments.

[5]An exception are Seibert and Kleinert (2009) who provide a descriptive analysis of mobility at the transition from apprenticeship training into the first job for Germany. Dustmann and Schönberg (2012) use mobility groups similar to our definition to estimate the extent of transferability of human capital across firms and/or occupations. Göggel and Zwick (2012) consider changes across employers and changes across occupations after apprenticeship, but it remains unclear whether these two groups are defined truly exclusively. Müller and Schweri (forthcoming) analyze occupational mobility after apprenticeship in Switzerland considering three well-defined groups similar to our definition of stayers, firm switchers and across-firm occupation switchers.

tive data. We can therefore distinguish four different mobility groups among apprenticeship graduates: stayers, pure firm switchers, within-firm occupation switchers and across-firm occupation switchers. Furthermore, we account for the heterogeneity of the estimated wage effects with regard to the wage position of the training occupation.

Our IV estimates imply that pure firm changes after graduation from apprenticeship lead to average wage losses of about 3.3-4.2% relative to stayers, although the long-term wage losses are reduced once we control for the training occupation. Regarding occupational mobility, the results differ strongly by whether there is a firm change. On average, job-and-occupation changes imply persistent wage losses of about 3.3-4.0% for a period of 7 years after entry into the first job relative to stayers. An occupation change within the firm results in persistent wage gains of about 12%. Within-firm occupation switchers are negatively selected and the switch allows the employee to move to an occupation which matches the employee's skills in a better way. Allowing for heterogeneous wage effects, we find that firm switchers and across-firm occupation switchers tends to lose less/benefit more with a lower relative wage position of the training occupation. In contrast, the wage gain of within-firm occupation switches increases in the relative wage of the training occupation. We further distinguish whether the employee moves to an occupation with a higher relative wage (upgrading) or to an occupation with a lower relative wage (downgrading). The results suggest that in the majority of cases an occupational switch involves a career progression. In contrast, for job switches the wage loss dominates – and the loss does not grow when there is an occupation switch at the same time.

The remainder of the paper is organized as follows: Section 2 discusses our identification strategy and the estimation approach. Section 3 describes the data used. Section 4 contains the empirical results. We present descriptive results and discuss the performance of the instrumental variables as well as the IV estimation results. Section 5 concludes. Appendix A includes Tables and Figures. Appendix B describes the data cleaning procedures and the construction of the sample. The Additional Online Appendix (Tables and Figures starting with "AOA." provides complementary empirical results and further background information.

2 Empirical Approach

2.1 Identification Strategy

We estimate the wage effects of mobility across firms and occupations up to seven years after graduation from apprenticeship. There are four treatments (mobility groups): Stayers, who do not switch neither their job nor their occupation, within-firm occupation switchers, job switchers within occupation, and job-and-occupation switchers (Table AOA.1). A comparison of average wages across the four mobility groups after controlling for observable characteristics would ignore potential selection effects in mobility based on unobservables. On the one hand, Acemoglu and Pischke (1998) and von Wachter and Bender (2006) find that job switchers are a negative selection. During apprenticeship training firms screen the ability of an apprentice and will only retain well-performing apprentices after graduation. By analogy, one would expect a negative selection of occupational switchers. A switch in occupation should be more rewarding for those graduates whose initial match with the training occupation was especially poor (Fitzenberger and Spitz-Oener, 2004; Gathmann and Schönberg 2010).

At the same time, to the extent that graduates choose to change their employer and/or occupation as a form of career advancement, mobility is more likely to occur if it leads to a wage increase relative to staying in the training firm and/or occupation (Topel and Ward, 1992). If this is the case, future wage prospects feed back into the mobility decision. This type of positive selection into mobility serves as another potential source of the endogeneity of mobility decisions. Previous work for Germany finds a positive selection of occupation switchers for older workers (Werwatz, 2002; Fitzenberger and Kunze, 2005).

Von Wachter and Bender (2006) point out that there is sorting into training occupations and training firms. On the one hand, one would expect that training firms with a low retention rate are attracting a worse pool of apprentices. On the other hand, able apprentices may choose a training firm with a low retention rate if the training is particularly useful for their career. Von Wachter and Bender (2006) find that sorting into firms implies a negative selection of job switchers. In contrast, Dustmann and Schönberg (2012) find that including firm fixed effects leaves the regression estimates for the wage effect of mobility among graduates unchanged. Thus, we only account for selection into training occupations by including 2-digit training occupation fixed effects in the wage regressions.

To identify the causal effect of mobility after apprenticeship on wages, we use variation in the local labor market situation in the year of graduation. Our instruments involve both push and pull factors, such as indicators of the tightness of the local labor market and group specific mobility rates.[6] We argue that our instruments provide an exogenous variation in mobility conditional on the sorting of apprentices by 2-digit training occupations, which we account for by including occupation fixed effects.

Our analysis uses data on the graduation cohorts 1992-1997 in West Germany. By the end of 1992 the reunification boom had come to a halt and the West German economy dropped into a deep recession which was accompanied by a worsening of labor market conditions and an increase in the unemployment rate. The recession was followed by a slow recovery until the late 1990's.[7] Thus, in addition to the regional variation, the indicators of the local labor market conditions used as instruments involve sizeable variation over time.

Table 1 summarizes the set of instrumental variables used. We use the aggregate local unemployment rate and the ratio of vacancies per registered unemployed to account for the business cycle in general. In addition, the unemployment rate for those below age 25 accounts specifically for the labor market changes for apprentices who are displaced by their training firm. We also include the shares of high-skilled and low-skilled workers to capture the educational background of the local workforce. The set of instruments also includes dummies for the German federal states, which differ in aggregate labor market conditions. Finally, as proxies for further local labor market characteristics that may affect mobility, we use regional mobility

[6]There are a number of studies which use similar instruments for mobility, see among others Neumark (2002), Müller and Schweri (forthcoming), Werwatz (2002), and von Wachter and Bender (2006).

[7]For a detailed account, see Sachverständigenrat (1993, p. 3), Sachverständigenrat (1996, pp. 1 and 22), and Sachverständigenrat (1998, pp. 84-87).

rates and exit rates into unemployment for male workers aged 25–35, where we exclude our apprenticeship graduates from the calculation.[8] Similar to von Wachter and Bender (2006), we use the mobility rates of other young workers as a proxy for local labor market characteristics that may affect the mobility of graduates from apprenticeship.[9]

The instrumental variables are matched to the sample of graduates from apprenticeship via the administrative district of the training firm and the year of graduation. [10] The way local labor market conditions affect mobility rates may differ across Germany, depending upon the labor market conditions in adjacent administrative districts and mobility patterns between different districts. Therefore, we allow the first stage regressions for the mobility dummies to differ by 26 West German regions.

To justify our identification strategy, our instruments must have a significant impact on mobility, and we need to discuss the necessary conditional exogeneity assumption. Pooled OLS estimations at the national level reveal a statistical significance of the instruments on the mobility dummies, see section 4.3 for details. For the time period under investigation, the exogeneity of the instruments for wages in West Germany (conditional on time effects accounting for the aggregate business cycle) is plausible because wages are basically determined by collective wage bargaining between unions and employer associations at the industry level, and coverage by industry-level wage agreements varies between 70% and 62% of employment (Schnabel, 2005). Consistent with our line of argument, Mertens (2002) finds that in West Germany wages are rigid at the level of federal states, and that regional labor demand shocks have no significant effect on wages.

2.2 Estimation

We estimate the following pooled wage regressions separately for the time period 0-2 years (short term) and the time period 3-7 years (long term) in employment after graduation from apprenticeship:

$$\begin{aligned}\log\left(wage_{it}\right) =&\alpha + \beta_1 \cdot job\_sw_i + \beta_2 \cdot occ\_sw_i + \beta_3 \cdot occ\_job\_sw_i + \gamma \cdot X_i \\ &+ \sum_j \delta_j \cdot occup_{j,i} + \epsilon \cdot yograd_i + \zeta \cdot yoempl_{it} + \eta \cdot year_{it} + u_{it}\end{aligned}$$

[8]The exit rates into unemployment, where the unemployment spell lasts at least 92 days, are calculated only for workers who were full-time employees at the end of the previous year. Observations in years with at least one apprenticeship training episode are excluded.

[9]The set of instrumental variables further contains dummy variables for a small cell size. Year-administrative district-economic sector cells are small ($n < 10$ persons) for about 7.4% of all graduates. Furthermore, the distributions of mobility rates show spikes at zero (these results are available upon request), for which we also include dummy variables.

[10]For variables measured at the level of employment agencies, we constructed a key that allows us to match employment agency districts to administrative districts (details are available upon request).

with the dummy variables job_sw_i, occ_sw_i, $occ\_job\_sw_i$ representing the three mobility dummies. In addition, we control for the following set of covariates (X_i): age at the beginning of the first job, diploma from upper track secondary schools (Abitur), non-German citizenship, and citizenship missing. All specifications include a set of dummies for year of graduation ($yograd_i$). We also add a dummy for each 2-digit training occupation j ($occup_{j,i}$) to control for selection into training occupations. Furthermore, all regressions control for the year since start of employment after graduation ($yoempl_{it}, t = 0, ..., 7$) and the calendar year ($year_{it}$). Standard errors are clustered at the person level.

To increase efficiency of the estimator, our instrumental variables (IV) approach takes account of the binary nature of the endogenous variables by estimating a Probit model in the first stage and by adopting GMM estimation in the second stage (Angrist, 2001; Wooldridge, 2010, chapter 21). Specifically, we adapt Wooldridge's Procedure 21.1 as follows:

1. Estimate a Probit model separately for 26 regions for each mobility dummy controlling for the exogenous covariates X_i and the local labor market characteristics IV_i and calculate the predicted probabilities $\hat{P}_{i,mobtype}$:

$$\hat{P}_{i,mobtype} = \alpha + \gamma \cdot X_i + \sum_j \delta_j \cdot occup_{j,i} + \lambda \cdot IV_i + \epsilon \cdot yograd_i + u_{it}$$

2. Estimate optimal cluster-robust GMM[11] using the three predicted probabilities $\hat{P}_{i,mobtype}$ from step 1 as excluded instruments for the endogenous mobility dummies.

This two-step procedure allows to use the usual GMM standard errors and test statistics and it is robust against a misspecification in the Probit models (Wooldridge, 2010, chapter 21).

In a second set of results, we allow the mobility effects to differ by the relative wage position of the training occupation. To obtain the relative wage position, we regress log-wages on age, age^2, a full set of year dummies, and a full set of occupation dummies (without intercept) for full-time working males below age 30:

$$\log(wage_i) = \sum_j \beta_j \cdot occup_{j,i} + \alpha_1 \cdot age + \alpha_2 \cdot age^2 + \eta \cdot year_{it} + u_i$$

where β_j is the estimated relative wage position for occupation j. We define $tw(occup)_i = \sum_j \beta_j \cdot occup_{j,i}$ as the relative wage position of the training occupation of individual i, and we calculate the average relative wage position within each mobility group, denoted by $\overline{tw}_{mobtype}$. The wage regression now includes both the three mobility dummies and three interaction terms with the mobility dummy for mobtype times $\left(tw(occup)_i - \overline{tw}_{mobtype}\right)$. Adapting Wooldridge (2010, Procedure 21.2), the second-step GMM estimation now uses both the three predicted probabilities $\hat{P}_{i,mobtype}$ and the three interaction terms $\hat{P}_{i,mobtype} \cdot \left(tw(occup)_i - \overline{tw}_{mobtype}\right)$ as instruments. In addition, the set of instruments includes a third order polynomial of the relative wage position. The normalization of the relative wage position allows us to use the

[11]We estimate optimal cluster-robust GMM using Stata command *ivregress* with clustered standard errors.

coefficient of the mobility dummy as the estimate of the average wage effect of mobility among the corresponding mobility group (ATT: average effect of treatment for the treated).

Based on the GMM estimates of the model with interaction effects, we calculate the estimated heterogeneous mobility effects at different deciles (q_j, with $j = 1, ..., 9$) of the relative wage of the training occupation as:

$$ATT_{q_j,mobtype} = coef_{mobtype} + (tw_{q_j,mobtype} - \overline{tw}_{mobtype}) \cdot coef_{(tw(occup)_i - \overline{tw}_{mobtype}) \cdot mobtype}$$

where $coef_{mobtype}$ is the coefficient of the mobility dummy and $coef_{(tw(occup)_i - \overline{tw}_{mobtype}) \cdot mobtype}$ is the coefficient of the interaction effect. We also calculate the treatment effects at different deciles of the entire sample.

3 Data

Our analysis is based on the IAB Employment Sample (IABS) regional file 1975-2004, a 2% random sample of all employees paying social security taxes (see Drews, 2008). The basic data involves employment spells and spells of unemployment benefit receipt. We restrict our sample to full-time working men in West Germany who completed their vocational training sometime during the period of 1992-1997 (Berlin is excluded). For employment spells, we observe daily wages, indicators of full-time and part-time work, the three-digit occupation code (about 130 occupations), and the industry. The dataset records a switch of establishment, but we do not know if two employees work in the same establishment. This prevents us from estimating establishment fixed effects.

An ongoing apprenticeship is recorded as a regular employment spell with the status information apprentice. To identify the completion of the first apprenticeship training, we use the information about when there is change in the reported education to vocational training degree. Because of potential misclassification problems, we implement a series of data cleaning procedures and sampling conditions. A further complication stems from the fact that there can be a time lag between completion of the vocational training degree and the fact being recorded in the education variable in the IABS. Appendix B provides an overview of the data cleaning procedures and detailed further data preparation steps.

We determine mobility after apprenticeship based on changes in the occupational code (occupation switch) and changes in the establishment ids (job switch) between the employment spell recording the apprenticeship and the first job spell after graduation. Figure AOA.1 illustrates the timing of spells in a case with an employment interruption between apprenticeship and first job after graduation.

There is a lot of concern in the literature about measurement error in occupational codes when using survey data which is self-reported by the employee, see e.g. Neal (1999) for the US. In fact, Longhi and Brynin (2010) argue that occupational switches within firms are not well measured in household panel data of the SOEP for Germany and the BHPS for the UK. Our administrative data involve occupational codes reported by the employer, for which measurement error is likely to be very small (similar data are used by Fitzenberger and Kunze, 2005, and Gathmann and Schönberg, 2010). It is likely that employers report precisely the occupation of the first regular job of an employee after graduation from apprenticeship. In fact, our data show a sizeable number of occupational switches within firms, which we can analyze in contrast to Longhi and Brynin (2010).

We construct an unbalanced wage panel for full-time working males with a yearly frequency (Table AOA.2). Starting with the wage in the year of the first employment spell after graduation, we record the wage up to seven years after the year of the first employment spell. Wages are averaged across all employment spells observed in one year. Since the IABS data only contains information on daily wages, we only take full-time employment spells into account. In case of parallel employment spells, we only use the spell with the highest recorded wage. We drop records with zero wages and jobs where employees work at home (*Heimarbeit*, typically part-time). Wages are deflated by the consumer price index (2005=100) and measured in Euros.[12] We impute top-coded wages based on a Tobit model, for which we only know that the wage exceeds the social security contribution.

4 Empirical Results

4.1 Descriptive Results

Table 2 shows descriptive statistics for the four mobility groups. Our sample consists of about 14.200 male apprenticeship graduates. While the four mobility groups differ in size, the sample shares do not vary a lot over the graduation years 1992 to 1997.[13] The stayers, i.e. those who stay with their training firm and their training occupation, form the largest mobility group. They will also serve as the comparison group in all further econometric analysis. Table 2 shows that, in comparison to stayers and job switchers, occupation switchers less often hold an upper secondary school degree and more often are of foreign citizenship. The average apprenticeship duration as well as the average age at the beginning of the first job after graduation are fairly similar across the four mobility groups. However, regarding the time it takes to start the first job, we observe strong differences between the four mobility groups. Stayers and within-firm occupation switchers quickly start their first job after graduation. In contrast, to start the first job after apprenticeship, it takes about 15 weeks for job switchers and 23 weeks for job-and-occupation switchers.

Figure 1 displays the descriptive wage profiles for the four mobility groups weighted by the individual length of employment spells. All mobility groups show average wages that increase almost linearly with years of employment. However, wage levels differ across mobility groups.

[12]The consumer price index is obtained from Statistisches Bundesamt (2010, p. 214).

[13]The overall share of graduates leaving the training firm in our sample is similar to that reported byn von Wachter and Bender (2006) for the German apprenticeship graduation cohorts 1992-1994. The shares of mobility groups in our sample are also roughly consistent with the ones reported for mobility among German apprenticeship graduates in Seibert and Kleinert (2009).

Within-firm occupation switchers earn higher wages than stayers. The two groups of apprenticeship graduates who leave their training firm, job switchers and across-firm occupation switchers do worse than the stayers.

4.2 OLS Results

Table 3 shows the estimated wage effects of mobility obtained by a Pooled OLS wage regressions controlling for a set of socio-economic covariates. The results reported in columns (1) and (3) imply that on average within-firm occupation switchers earn about 7.5% higher wages than stayers in the short run (up to two years after entry into first job), and about 6.9% higher wages in the long run (years three to seven after entry into first job). In contrast, firm switchers do worse than stayers in terms of wages. Relative to stayers, wage losses for job switchers amount to about 3.5% in the short run and about 3.8% in the long run. Relative wage losses for job-and-occupation switchers are slightly more pronounced with losses of about 4% in the short run and about 4.9% in the long run. A comparison of short-run and long-run results suggests that wage differences are persistent and for both job switchers and job-and-occupation switchers no catching up takes place over a seven-year horizon after entry into the first job. However, as the results in Table 3 show, within each time window on average wages tend to increase over years of employment.

In addition, the specifications in columns (2) and (4) of Table 3 control for the 2-digit training occupation to account for possible sorting of apprentices into training occupation. The wage gains of within-firm occupation switches are stronger, both in the short and long run, compared to the results without controlling for the 2-digit training occupations. This suggests a negative selection regarding the training occupations of within-firm occupation switchers. The relative wage losses of job switchers and job-and-occupation switchers are less pronounced after controlling for the 2-digit training occupation. This suggests that also these two mobility groups are negatively selected with respect to their training occupations. These results are similar to the findings of von Wachter and Bender (2006) regarding the negative selection of firm switchers regarding the training firms.

4.3 First Stage of IV Estimation

We exploit exogenous variation in local labor market conditions to instrument the different potentially endogenous mobility dummies. Our identification strategy is based on the assumption that the local labor market situation in the year of graduation is significantly correlated with graduates' propensity to leave the training firm and/or to switch occupation. From the first-step (stage zero) Probit regressions of the mobility decisions on the exogenous covariates and the local labor market conditions described in subsection 2.2, we obtain predicted probabilities $\hat{P}$ that then serve as the excluded instruments in the GMM estimation approach. When checking the validity of the above-mentioned assumption, we thus have to consider both the statistical relationship between the local labor market conditions (our original instruments) and the mobility decisions as well as the relationship between the predicted probabilities (our "constructed" instruments) and the mobility decisions.

As explained in subsection 2.2, in the first step of the IV procedures we also allow for heterogeneity regarding the influence of local labor market conditions on mobility decisions by estimating separate Probit regressions for 26 West German regions. We thus exploit the fact that the broader economic environment of the larger regions may mediate the way in which local labor market conditions (at the administrative district level) influence graduates' mobility decisions.

To summarize the relationship between the local labor market conditions and the mobility decisions, we run an OLS estimation at the national level for each of the three mobility groups. More specifically, we regress the predicted probabilities $\hat{P}$ obtained from the respective 26 initial Probit regressions on the set of exogenous covariates X_i and the local labor market conditions while pooling observations from all 26 regions:

$$\hat{P}_{i,mobtype} = \alpha + \gamma \cdot X_i + \lambda \cdot IV_i + \delta \cdot occup_i + \epsilon \cdot yograd_i + u_{it}$$

The estimation results displayed in Table 4 show a statistically significant correlation between the local labor market conditions and the three different mobility dummies. When testing for joint statistical significance of the local labor market conditions, we obtain large F-statistics with values above 25. Patterns of individual significance and the signs of coefficients of local labor market conditions vary across the three regressions, thus showing that the different kinds of mobility decisions are affected in a different way by the local labor market conditions. The predicted probability of job switches within occupation appears to be driven by push factors. Whenever and wherever the local labor market conditions are worsening (increasing unemployment rates, lower labor market tightness), the predicted probability of firm change increases.[14] The opposite seems to hold for within-firm occupation switches. Here, an improving local labor market situation is correlated with a higher propensity to change occupation within the training firm. For job-and-occupation switches the picture is mixed. The predicted probability of job-and-occupation switches increases with higher overall unemployment, but decreases with higher youth unemployment ($<$ 25 years), and it increases as the ratio of vacancies to unemployed improves. Thus, in the case of job-and-occupation switches both push and pull factors are significant.

As a proxy for further unobserved local labor market conditions that affect mobility, we have also included transition rates that vary at the local as well as the industry level for male workers aged 25–35. Thus, similar to von Wachter and Bender (2006) we use the mobility behavior of other young workers in the local labor market as a proxy for the individual graduate's propensity to change the firm and/or occupation. As Table 4 shows a certain higher overall mobility rate of young workers is always significantly positively correlated with the predicted probability of the respective mobility decision for apprenticeship graduates. Very clearly, within-firm occupation switches are less likely to occur in an environment with a higher exit rate into unemployment, with more job switches, or with more job-and-occupation switches. Regarding the determinants of job switches and job-and-occupation switches, the picture is somewhat

[14]Mertens and Haas (2006) find a similar average relationship between regional unemployment rates and job mobility of workers for the period 1984-1999 in Germany. Furthermore, the workers were explicitly asked whether the job change was voluntary or involuntary. The authors find that rising local unemployment rates are related to higher involuntary job mobility and lower voluntary job mobility.

mixed. Also, Table 4 implies that each type of mobility is more likely to occur if the local workforce involves a higher share of highly qualified employees and a lower share of employees with low qualifications.

Considering the statistical relationship between the predicted probabilities (our "constructed" instruments) and the mobility decisions, we find strong regional differences (Figures AOA.3-AOA.5). We exploit this variation in the instrumental variables approach and find highly statistically significant F-statistics for the excluded instruments (the "constructed" instruments) in the first stage of the GMM estimator (Table 5).

4.4 IV Estimates without Heterogeneous Treatment Effects

We cannot assume random assignment into the four mobility groups for our sample of apprenticeship graduates conditional on the control variables considered in the OLS regressions. There is very likely selection into mobility, and from a theoretical perspective, both negative as well as positive selection effects could arise. Since an across-group comparison of average wage levels is likely to results in a biased estimate of the wage effects of mobility, we continue our analysis with estimating the causal effects of mobility after apprenticeship using an instrumental variables approach.

Table 6 displays the estimation results of the IV procedure (GMM, Wooldridge Procedure 21.1) discussed in subsection 2.2. On average, wage losses due to job switches amount to about 4.3% (column (2)) in the short run and are largely persistent over time. This suggests that no catching up takes place relative to stayers, a result which differs from the results obtained by von Wachter and Bender (2006) for all job switches. The negative wage effect of a job switch is more pronounced than in the OLS regression (compare Table 3). This suggests a positive selection of job switchers into mobility.

The IV estimates also imply a causal wage effect of within-firm occupation switches that is much stronger than in the OLS regression. An occupation switch within the training firm results in an average wage gain of about 14.3% in short run relative to stayers (column (2)). These gains are largely persistent for a period of up to seven years after entry into the first job. A comparison of IV and OLS estimation results suggests a negative selection of within-firm occupation switchers.

Regarding the job-and-occupation switchers, the IV estimation results reveal a negative causal wage effect of leaving both the training firm and the training occupation. However, the effect is only statistically significant in the short run, amounting to an average wage loss of about 3.3% relative to stayers (column (2)). Since the long-run estimate is insignificant, some catching up relative to stayers may be possible in the long run (column (4)). The comparison to OLS results tends to imply a negative selection of job-and-occupation switchers.[15]

[15] As a robustness check, we re-estimate the model shown in Table 6 with different clustered standard errors. We are grateful to a referee for this suggestion. Once we cluster the standard errors in the last step of the GMM estimation at the region - interacted with year-of-graduation level, which is the level at which a number of the instruments vary, instead of at the person level, standard errors do increase slightly. However, the significance level of the estimated

While the OLS estimation results suggest that all three mobility groups are negatively selected with respect to the 2-digit training occupation, comparison of IV specifications in Table 6 with and without 2-digit training occupation fixed effects shows a different pattern for job switchers. Here, job switchers are revealed to be positively selected into the training occupation. The IV results still indicate that within-firm occupation switchers are negatively selected with respect to the training occupation, while the results are somewhat inconclusive for job-and-occupation switchers.

4.5 Overidentification Test and Reducing the Number of Instruments

We use a large number of instruments when constructing the predicted probabilities $\hat{P}_{i,mobtype}$. This provides the opportunity to investigate the validity of the instruments by means of an overidentification test.[16] However, a standard overidentification test is not applicable for two reasons. First, we implement a GMM estimation approach which is formally based on the predicted probabilities as instruments. Thus, the GMM objective in the second stage can not be used for an overidentification test, simply because formally we have an exactly identified case. Second, even though our estimates are second stage GMM estimates building on the weighting matrix estimated in the first stage, we argue that inference has to take account of clustering at the individual level. This is because the weighting matrix estimated in the first stage and used in the second stage does not account of clustering. Note further that our instruments are assumed to affect the endogenous treatment dummies through the nonlinear function yielding $\hat{P}_{i,mobtype}$.

As a simple approach to implement an overidentification test, we extend the heteroscedasticity-robust test of overidentifying restrictions for the two-stage least squares estimator suggested by Wooldridge (2010, p. 136) to our setting as follows. (i) We first run a panel regression of all instruments on the three $\hat{P}_{i,mobtype}$'s and on the exogenous regressors in the wage equation. Denote the residuals from this auxiliary regression as $\hat{r}_2$. (ii) Next, we regress the estimated residuals of the wage regression (these residuals are based on the GMM coefficient estimates and the actual treatment dummies plugged into the wage regression) on the residual variation of the instruments $\hat{r}_2$. (iii) We use the cluster robust Wald test statistic for the joint significance of all instruments in the regression under (ii). Because we estimate separate probit regressions by 26 regions to estimate $\hat{P}_{i,mobtype}$, our instruments are fully interacted with the regions. For this reasons, we implement the overidentification test (i)-(iii) separately by the 26 regions. Furthermore, we weight all regressions by the employment weight for each single wage observations. As a caveat, it should be noted that the auxiliary regressions involve a linear approximation of the possibly nonlinear relationship between the instruments and the error term in the wage regression.[17] Furthermore, we differ from Wooldridge (2010, chapter 6)

---

coefficients only changes in two cases (Table AOA.3). Thus, we conclude that our results are basically robust to this change.

[16]We are grateful to one referee for suggesting to implement an overidentification test for our case.

[17]Thus, a misspecification of the probit models for the treatment dummies may also cause a rejection of the overidentification test even though the instruments may still be strictly exogenous.

by using all instruments in our auxiliary regressions in (i) and (ii), simply because $\hat{P}_{i,mobtype}$ is not a linear function of the instruments and therefore the matrix spanned by $\hat{r}_2$ has full rank.

The benchmark specification discussed in table 6 is based on 22 instruments. The overidentification test (see Table AOA.4 in the additional appendix, Panel '0. Original set of 22 IV's') typically does not lead to a rejection at the 1% significance level for a majority of regions, but depending on the case considered there are between 5 and 13 rejections among 26 regions. The rejection rate is considerably higher than the significance level of 1%, and also the joint test for the national level involves a rejection. Thus, strictly speaking, our IV approach does not pass the overidentification test.

To address the problem in more detail, we now reduce sequentially the set of instruments from 22 to 7 instruments, as described in the notes of Table AOA.4. The 7 core instruments involve the general indicators of regional labor market conditions such as the unemployment rate, the unemployment rate below age 25, and labor market tightness at the district level (as well as powers of these variables). The 9 IV's also involve the skill shares among employees. And the 12 IV's involve information on the mobility share regarding unemployment longer than 3 months. For the 12 IV's, we have excluded the information on mobility shares regarding job switches and occupational switches whose effects we are estimating. Thus, one might be most concerned about the validity of these group instruments. When we reduce the number of instruments, the number of rejections of the overidentification test falls dramatically. With 7 instruments there is no rejection any more for the long run with occupation dummies and there are only between one to three rejections for the other case (for all rejections the p-value lies around 0.005, except for one case with 0.0005). We take the model with 7 instruments as basically passing the overidentification test. One could make a similar argument for the model with 9 instruments. Regarding the number of rejections, the model with 12 IV's lies somewhere in between the model with 22 IV's and the model with 9 IV's.

Should we now use the model estimates based on 7 or 9 instruments as our benchmark model? A comparison of the estimated treatment effects in table 7 suggests that the variation of the set of instruments does not change the estimated treatment effects in a considerable way.[18] In particular, there is a striking qualitative similarity of the results (in light of the estimated standard errors) for the estimates with fixed effects for the 2-Digit training occupations). If the validity of the instruments were to be questioned, we would expect that the estimated treatment effects would change strongly. However, that is not the case. We rather think that the rejections suggest a slight misspecification of the probit model used to construct the predicted treatment probabilities $\hat{P}_{i,mobtype}$. Put differently, we do not have a problem of endogenous instruments but rather the nonlinearity of the relationship between the instruments and the treatment dummy variables may not be fully captured by our probit model. This problem may be aggravated by the fact that our mobility shares are noisy estimates and that we account for the fact that cell sizes are too small.

Because the IV approach suggested by Wooldridge (2010, p. 939), which we use in this paper, explicitly allows for a misspecification of the Probit model, we stick to our benchmark estimates with 22 instruments. Furthermore, the subsequent analysis will also be based on the set of 22

[18]The strongest difference involves the absolute reduction in the negative wage effect of a job switch in the models without fixed effects for the 2-Digit training occupations.

instruments.

4.6 IV Results with Heterogeneous Treatment Effects

The IV estimations presented so far estimate treatment effects that are homogeneous with respect to the relative wage position of the training occupation. However, differences in the relative wages of training occupations may reflect differences in the amount of occupation-specific capital typically obtained trough training as well differences in the occupation-specific ratio of labor supply and demand. In the following we will thus drop the assumption of homogeneous starting conditions within mobility groups by taking account of the relationship between the relative wage of the training occupation and the wage effects of mobility. The IV procedure (Wooldridge Procedure 21.2) discussed in Subsection 2.2 estimates the ATT, taking account of the effect heterogeneity by the relative wage level of the training occupation.

The main mobility effects shown in Table 8 are calculated as average effects among the corresponding mobility group. Regarding these average causal mobility effects the results do not change much relative to the IV results without heterogeneous treatment effects. For job switchers and within-firm occupation switchers the effects are a bit less pronounced than before. For job-and-occupation switchers the negative long-term wage effect now becomes statistically significant.

Regarding the relevance of the training occupation, Table 8 shows that on average the relative wage of the training occupation $tw(occup)_i$ is positively related to current wages both in the short and long run.[19] This means that apprenticeship graduates from training occupations with a higher relative wage also earn higher wages during the first seven years of their labor market careers.

Most importantly, the interaction effects between the relative wage distance and the mobility type reveal interesting results. Job-and-occupation switchers display negative interaction effects. For job switchers, the interaction effect is close to zero in the short run, but becomes negative in the long run. Interestingly, the relationship is reversed for within-firm occupation switchers. Here, we find a positive interaction effect.

To illustrate the meaning of these findings, Figure 2 shows the ATT at deciles of the overall distribution of wages in the training occupation for each of the three treatment groups.[20] For job switchers and job-and-occupation switchers we find that those members of the mobility group who have been trained in a low-wage training occupation suffer relatively less from being mobile (relative to those having been trained in better-paid training occupations). We cannot rule out, that for the most ill-positioned graduates the respective mobility decision may even be neutral relative to stayers in terms of wages. Interestingly, the ATTs for job switchers and job-and-occupation switchers are not statistically different from each other. This

[19] The respective coefficients on $tw(occup)_i$ have to be interpreted as elasticities: On average a 1% higher wage in the training occupation is associated with a β% higher wage after graduation.

[20] A conditional version of Figure 2, where we compute the ATT at deciles of the group-specific distribution of wages in the training occupation, shows similar results (Figure AOA.6).

suggests that additional to leaving the training firm a change of occupation does not have any further negative wage effects for the apprenticeship graduate. A change of occupation within the training firm is clearly beneficent for the apprenticeship graduates. Those apprenticeship graduates who, regarding their choice of training occupation, are already in a favorable initial position profit most from an occupational switch within the training firm. Even the initially most ill-positioned graduates profit from a within-firm occupation switch relative to stayers. [21]

Note that Figure 2 shows very similar patterns both for the short-term and the long-term results. Within-firm occupation switchers persistently perform better than stayers in terms of wages. For job switchers and job-and-occupation switchers the treatment effects appear to be largely persistent over time. However, one has to keep in mind that the average long-term effect of pure firm changes reported above turns insignificant once we control for selection into training occupations, so that for job switchers catching up relative to stayers could be possible.

A comparison of the IV estimation results relative to the corresponding OLS results in Table 9 shows similar selection patterns as discussed above in Section 4.4 for IV estimation without heterogeneous treatment effects. The estimation results indicate a positive selection of job switchers. Within-firm occupation switchers are negatively selected with respect to unobservables. This result holds in particular in light of the differences in the corresponding interaction effects between OLS and IV estimation with heterogeneous treatment effects. However, the IV results imply that there is no significant selection on unobservables for job-and-occupation switchers. A comparison of IV specifications in Table 8 with and without 2-digit training occupation fixed effects again suggests sorting into training occupations for all mobility groups.

4.7 Occupational Upgrading and Downgrading

Not only may the effects of occupational mobility depend on the initial occupational position of the apprenticeship graduates, but they may also be related to the direction of the occupational move. We explore this aspect of occupational mobility by distinguishing between upward and downward switches. Based on relative wages, we ordinarily rank all 130 occupations observed in the IABS from lowest paid (1) to highest paid (130). For each apprenticeship graduate we then compare the rank of his training occupation to the rank of his occupation in the first job after graduation and thereby determine whether they performed an upward or downward occupational switch. We find that in both mobility groups a significant proportion of occupational switches is directed towards higher ranked occupations. About 60% of within-firm occupation changes are upward. Surprisingly, even in the group of job-and-occupation switchers about 48% of all cases are associated with an upward move.

In light of these results, we estimate a modified version of the IV procedure without heterogeneous treatment effects that distinguishes between upward and downward occupation

[21]Relative to the group of stayers, in the group of job-and-occupation switchers and even more so in the group of job switchers weakly ranked training occupations are more frequent. The group of within-firm occupation switchers is more dominant in the upper part of the ranking of training occupations (Figure AOA.7).

switches. As Table AOA.6 shows, the wage effects of occupational mobility are indeed heterogeneous with respect to the direction of the occupational move.[22] For within-firm occupation switchers we find that even those apprenticeship graduates who move towards a lower ranked occupation on average still realize significant relative wage gains of about 6.6% that largely persist over a seven year period after graduation.

Most importantly, we find that an occupation switch across firms does not necessarily cause a negative wage effect. Those job-and-occupation switchers who move towards a higher ranked occupation do not suffer wage losses on average. In the short run, they even realize significant average wage gains of about 6.7% relative to the stayers. In the long run, upward job-and-occupation switches appear to be at least wage neutral. These effects are strongest when we include fixed effects for the 2-digit training occupation and, thus, only compare job-and-occupation switchers moving away from the same initial 2-digit training occupation.

5 Conclusions

Distinguishing carefully between mobility across firms and across occupations, this study provides causal estimates of the wage effects of mobility among graduates from apprenticeship in Germany during the first seven years after starting the first regular job after graduation. Our analysis distinguishes between pure firm switchers, within-firm occupation switchers, and across-firm occupation switchers. Mobility across firms and occupations may be associated with a loss of human capital implying a wage loss or with finding a better job match implying a wage gain. Due to the likely presence of selection based on unobservables, OLS estimates are likely to be biased and we employ an instrumental variables approach exploiting variation in regional labor market characteristics. We show that local labor market conditions, such as the unemployment rate, labor market tightness and mobility behavior of the local workforce, are significantly correlated with mobility after graduation from apprenticeship. Our analysis accounts for the heterogeneity of the estimated wage effects with regard to the wage position of the training occupation.

Our IV estimates imply that pure firm changes after graduation from apprenticeship lead to average wage losses of about 3.3-4.2% relative to stayers, although the long-term wage losses are reduced once we control for the training occupation. Job switchers are positively selected into mobility with respect to unobservable characteristics relative to stayers.

Regarding occupational mobility, the results differ strongly by whether there is a firm change. On average, job-and-occupation switches imply persistent wage losses of about 3.3-4.0% for a period of 7 years after entry into the first job relative to stayers. An occupation switch within the training firm results in persistent wage gains of about 12%. Our results indicate that Across-firm occupation switchers basically show no selection on unobservables, while within-firm occupation switchers are negatively selected. During the training period the employer

[22]The corresponding OLS estimation results can be found in Table AOA.5. Table AOA.7 provides results on the correlation between the local labor market conditions and upward/downward occupational mobility. Table AOA.8 shows the F-statistics for the excluded instruments in the first stage of the GMM estimation.

can observe the apprentice's ability and then decide, whether the employee should switch to an occupation which matches the employee's skills in a better way. This occurs in particular when the initial match with the training occupation was poor.

Allowing for heterogeneous wage effects, we find that job switchers and across-firm occupation switchers tend to lose less/benefit more with a lower relative wage position of the training occupation. In contrast, the wage gain of within-firm occupation switchers increases in the relative wage of the training occupation. Furthermore, we find that the wage effects of occupational mobility differ by the direction of the move. Occupational upgrading across firms, which comprises 48% of all job-and-occupation switches, actually causes an average wage gain of 6.7%.

While our results indicate that pure firm changes after apprenticeship lead to wage losses, our conclusions regarding the wage effects of occupational mobility after apprenticeship are somewhat more positive. Occupational mobility within the training firm can be interpreted as a career progression involving persistent wage gains. The positive wage effects of occupation switches within the firm and occupational upgrading across firms suggest that for the majority of cases a change of occupation involves a career progression. In contrast, for job switches the loss of firm-specific human capital seems to dominate – and the loss does not grow when there is an occupation switch at the same time. At a more general level, our results suggest that the skills acquired through apprenticeship training in a specific occupation are sufficiently general to be useful when working in another occupation.

References

Acemoglu, D. and J.-S. Pischke (1998). "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113 (1), 79-119.

Angrist, J. D. (2001). "Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business & Economic Statistics*, 19(1), 2-28.

Autor, D., F. Levy, and R. Murnane (2003). "The Skill Content of Recent Technological Change: An empirical exploration." *Quarterly Journal of Economics*, 118, 1279-1333.

BMBF (2012). "Vocational Education and Training in Europe – Perspectives for the Young Generation." *Memorandum on Cooperation in Vocational Education and Training in Europe*, Berlin, 10-11/12/2012. Available online:
http://www.bmbf.de/pubRD/memorandumvocational_education_and_training_2012.pdf, Last accessed 24.01.2014.

Bougheas, S. and Y. Georgellis (2004). "Early Career Mobility and Earnings Profiles of German Apprentices: Theory and Empirical Evidence." *LABOUR*, 18 (2), 233-263.

Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, Ed. (1997). *Berufsbildungsbericht 1997.* Bonn.

Clark, D. and R. Fahr (2002). "The Promise of Workplace Training for Non-College Bound Youth: Theory and Evidence from German Apprenticeship." *CEP Discussion Papers dp0518*, Centre for Economic Performance, LSE.

Cortes, G. M. (2012). "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data." Manchester: The University of Manchester, *Working Paper No. EDP-1224.*

Drews, N. (2008). "Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004 – Handbuch-Version 1.0.2." *FDZ Datenreport*, (02/2008).

Dustmann, C. and U. Schönberg (2012). "What makes firm-based vocational training schemes successful? The role of commitment." *American Economic Journal. Applied Economics*, 4(2):36-61.

The Economist (2013). "German education and training – Ein neuer Deal?" *The Economist*, Print edition 01/06/2013. Available online:
http://www.economist.com/printedition/2013-06-01, Last accessed 24.01.2014.

Euwals, R. and R. Winkelmann (2002). "Mobility after Apprenticeship – Evidence from Register Data." *Applied Economics Quarterly (Konjunkturpolitik)*, 48 (3-4), 256-278.

Euwals, R. and R. Winkelmann (2004). "Training intensity and first labor market outcomes of apprenticeship graduates." *International Journal of Manpower*, 25 (5), 447-462.

Fitzenberger, B. and A. Kunze (2005). "Vocational Training and Gender: Wages and Occupational Mobility among young Workers." *Oxford Review of Economic Policy*, 21 (3), 392-415.

Fitzenberger, B., A. Osikominu, and R. Völter (2006). "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample." *Schmollers Jahrbuch*, 126 (3), 405-436.

Fitzenberger, B. and A. Spitz-Oener (2004). "Die Anatomie des Berufswechsels: Eine empirische Bestandsaufnahme auf Basis der BIBB/IAB-Daten 1998/1999." In: Franz, W. and H. Ramser and M. Stadler (Eds.): "Bildung." Tübingen: J.C.B. Mohr (Paul Siebeck), *Schriftenreihe des wirtschaftswissenschaftlichen Seminars Ottobeuren*, 29-55.

Gathmann, C. and U. Schönberg (2010). "How General Is Human Capital? A Task-Based Approach." *Journal of Labor Economics*, 28 (1), 1-49.

Gericke, N. and A. Uhly and J. G. Ulrich (2011). "Wie hoch ist die Quote der Jugendlichen, die eine duale Berufsausbildung aufnehmen? Indikatoren zur Bildungsbeteiligung." *Berufsbildung in Wissenschaft und Praxis*, 1/2011, 41-43.

Göggel, K. and T. Zwick (2012). "Heterogeneous Wage Effects of Apprenticeship Training." Scandinavian Journal of Economics, 114 (3), 756-779.

Groes, F., P. Kircher, and I. Manovskii (forthcoming). "The U-shapes of Occupational Mobility." *The Review of Economic Studies*, 2015, forthcoming.

Hoeckel, K. and R. Schwartz (2010). "Learning for Jobs: OECD Reviews of Vocational Education and Training – Germany." OECD, Paris.

Kambourov, G. and I. Manovskii (2008). "Rising Occupational and Industry Mobility in the United States: 1968–97." *International Economic Review*, 49, 41-79.

Kambourov, G. and I. Manovskii (2009). "Occupational Specificity of Human Capital." *International Economic Review*, 50 (1), 63-115.

Lalé, E. (2012). "Trends in Occupational Mobility in France: 1982–2009." *Labour Economics*, 19, 373-387.

Longhi, S. and M. Brynin (2010). "Occupational Change in Britain and Germany." *Labour Economics*, 17, 655-666.

Mertens, A. (2002). "Regional and Industrial Wage Dynamics in West Germany and the United States." *Jahrbücher für Nationalökonomie und Statistik*, 222 (5), 584-608.

Mertens, A. and A. Haas (2006). "Regionale Arbeitslosigkeit und Arbeitsplatzwechsel in Deutschland – Eine Analyse auf Kreisebene." *Jahrbuch für Regionalwissenschaft*, 26, 147-169.

Moscarini, G. and K. Thomsson (2007). "Occupational and Job Mobility in the US." *Scandinavian Journal of Economics*, 109, 807-836.

Müller, B. and J. Schweri (2009). "Berufswechsel beim Übergang von der Lehre in den Arbeitsmarkt." University of Zurich, Institute for Strategy and Business Economics (ISU), *Economics of Education Working Paper Series No. 44.*

Müller, B. and J. Schweri (forthcoming). "How Specific is Apprenticeship Training? Evidence from Inter-Firm and Occupational Mobility after Graduation" *Oxford Economic Papers*, 2015, forthcoming.

Neal, D. (1999). "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics*, 17(29), 237-261.

Neumark, D. (2002). "Youth Labor Markets in the United States: Shopping Around vs. Staying Put." *The Review of Economics and Statistics*, 84, 462-482.

Sachverständigenrat, Ed. (1993). "Zeit zum Handeln – Antriebskräfte stärken. Jahresgutachten des Sachverständigenrats zur Begutachtung der gesamtwirtschaftlichen Entwicklung 1993/94".

Sachverständigenrat, Ed. (1996). "Reformen voranbringen. Jahresgutachten des Sachverständigenrats zur Begutachtung der gesamtwirtschaftlichen Entwicklung 1996/97".

Sachverständigenrat, Ed. (1998). "Vor weitreichenden Entscheidungen. Jahresgutachten des Sachverständigenrats zur Begutachtung der gesamtwirtschaftlichen Entwicklung 1998/99".

Schnabel, C. (2005). "Gewerkschaften und Arbeitgeberverbände: Organisationsgrade, Tarifbindung und Einflüsse auf Löhne und Beschäftigung." *Zeitschrift für Arbeitsmarkt-Forschung*, 38 (2/3), 181-196.

Seibert, H. and C. Kleinert (2009). "Duale Berufsausbildung – Ungelöste Probleme trotz Entspannung." *IAB-Kurzbericht*, 10/2009.

Statistisches Bundesamt, Ed. (2010). "Preise – Verbraucherpreisindizes für Deutschland – Lange Reihen ab 1948". Wiesbaden.

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

Uhly, A., L. Lohmüller, and U. M. Arenz (2006). "Schaubilder zur Berufsbildung – Strukturen und Entwicklungen – Ausgabe 2006". Available online: www.bibb.de/schaubilder, Last accessed 20.01.2014.

von Wachter, T., J. Schmieder, and S. Bender (2009). "Job Displacement "The Long-Term Impact of Job Displacement in Germany During the 1982 Recession on Earnings, Income, and Employment" *Columbia University, Department of Economics Discussion Paper Series DP0910-07*.

von Wachter, T. and S. Bender (2006). "In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers." *The American Economic Review*, 96 (5), 1679-1705.

Werwatz, A. (2002). "Occupational Mobility after Apprenticeship – How Effective ist the German Apprenticeship System?" *Applied Economics Quarterly (Konjunkturpolitik)*, 48 (3-4), 279-303.

Wolter, S. C. and P. Ryan (2011). "Apprenticeship." In: Hanushek, E. A. and S. Machin and L. Woessmann (Eds.): *Handbook of the Economics of Education, Volume 3*, Elsevier, 1st edition, 521-576.

Wooldridge, J. M. (2010). "Econometric Analysis of Cross Section and Panel Data". Cambridge, Massachusetts: The MIT Press, 2nd edition.

Appendix A: Tables and Figures

Table 1: Main Instrumental Variables

Instrumental variable	Level of variation	Data source
unemployment rate	iabs-districts	FEA
unemployment rate $<$ 25 years	iabs-districts	IABS, FEA
vacancies/unemployed	empl. agency	FEA
mobility rates: job switch within-firm occ. switch job-and-occ. switch exit into unemployment $>$ 3 months	iabs-districts, economic sector	IABS
share of low-skilled workers share of high-skilled workers	empl. agency	FEA

Notes: FEA: Federal Employment Agency, IABS: IAB Employment Sample regional file 1975–2004; Dummies for German federal states also included; Regarding the mobility rates, the set of instrumental variables further contains dummy variables (and interactions thereof with the mobility groups) controlling for small cell size and mobility rates of zero.

Table 2: Summary Statistics for Four Groups of Apprenticeship Graduates

Variable	All graduates	Mobility type stayers	job switchers	within-firm occ. switch	job-and-occ. switch
Total	14234	8316	2225	1198	2495
Share	1	0.58	0.16	0.08	0.18
Year of graduation					
1992	2362	0.61	0.15	0.08	0.16
1993	2483	0.60	0.16	0.08	0.16
1994	2495	0.56	0.16	0.08	0.19
1995	2342	0.60	0.14	0.10	0.16
1996	2237	0.56	0.17	0.09	0.19
1997	2315	0.57	0.16	0.08	0.19
High school diploma	0.10	0.10	0.12	0.08	0.07
Foreign citizenship	0.10	0.08	0.10	0.12	0.16
Citizenship missing	0.02	0.02	0.02	0.02	0.03
Apprenticeship duration	1076	1071	1072	1096	1083
Distance between graduation and first job (days)	49	6	107	6	160
Age at beginning of first job	20.83	20.72	21.08	20.77	21.01

Table 3: Pooled OLS Estimates without Heterogeneous Treatment Effects

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.0346***	-0.0251***	-0.0378***	-0.0222***
	[0.0056]	[0.0051]	[0.0066]	[0.0063]
Within-firm occ. switch	0.0753***	0.0841***	0.0690***	0.0734***
	[0.0077]	[0.0070]	[0.0087]	[0.0083]
Job-and-occ. switch	-0.0404***	-0.0353***	-0.0492***	-0.0395***
	[0.0061]	[0.0059]	[0.0069]	[0.0068]
Age at job entrance	0.0105***	0.0109***	0.0124***	0.00799***
	[0.0014]	[0.0013]	[0.0016]	[0.0016]
High school diploma	0.0388***	0.0489***	0.127***	0.0973***
	[0.0080]	[0.0083]	[0.0099]	[0.0105]
Foreigner	0.0244***	0.0106*	0.0270***	0.0175**
	[0.0065]	[0.0059]	[0.0075]	[0.0072]
Foreigner missing	-0.111***	-0.0942***	-0.124***	-0.0942***
	[0.0137]	[0.0128]	[0.0166]	[0.0165]
Year of employment 1	0.103***	0.0983***		
	[0.0067]	[0.0064]		
Year of employment 2	0.192***	0.180***		
	[0.0126]	[0.0121]		
Year of employment 4			0.0588***	0.0517***
			[0.0068]	[0.0066]
Year of employment 5			0.114***	0.100***
			[0.0134]	[0.0129]
Year of employment 6			0.165***	0.144***
			[0.0199]	[0.0191]
Year of employment 7			0.214***	0.186***
			[0.0265]	[0.0254]
Constant	4.182***	3.932***	4.274***	3.991***
	[0.0070]	[0.0289]	[0.0078]	[0.0322]
Fixed effects				
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
N	14225	14225	13378	13378
R-sq	0.060	0.192	0.067	0.134

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell.

Table 4: OLS Regression of Predicted Probabilities of Mobility on the Local Labor Market Conditions at the National Level (Pooling 26 Regions)

Dependent variable: Predicted probability of	Job switch (1)	Within-firm occ. switch (2)	Job-and-occ. switch (3)
Unemployment rate	0.0346***	-0.00345	0.0278***
	[0.0070]	[0.0065]	[0.0070]
Unemployment rate^2	-0.00371***	0.000582	-0.00198***
	[0.0007]	[0.0006]	[0.0007]
Unemployment rate^3	0.000114***	-0.0000101	0.0000663***
	[0.0000]	[0.0000]	[0.0000]
Unemployment rate < 25 years	0.00565***	-0.00387***	-0.00379***
	[0.0008]	[0.0008]	[0.0008]
Labor market tightness	-0.00180***	0.00118***	0.00130***
	[0.0004]	[0.0004]	[0.0005]
Labor market tightness^2	0.0000492***	-0.0000300***	0.00000746
	[0.0000]	[0.0000]	[0.0000]
Labor market tightness^3	-0.000000274***	0.000000170***	-0.000000136**
	[0.0000]	[0.0000]	[0.0000]
Share low qualified	-0.000840**	-0.000159	-0.0000146
	[0.0004]	[0.0004]	[0.0004]
Share highly qualified	0.00295***	0.000832**	0.00431***
	[0.0004]	[0.0004]	[0.0004]
Mobility shares			
Unemployment	-0.000214	-0.00374***	0.000407
	[0.0005]	[0.0005]	[0.0005]
Job switch	0.00236***	-0.00163***	0.00268***
	[0.0002]	[0.0002]	[0.0002]
Within-firm occ. switch	-0.00277***	0.00324***	0.00188***
	[0.0005]	[0.0005]	[0.0005]
Job-and-occ. switch	0.00190***	-0.00347***	0.0000289
	[0.0003]	[0.0003]	[0.0003]
Further instrumental variables			
Interaction effects indicating small cells for mobility shares	Yes	Yes	Yes
Interaction effects indicating mobility share zero	Yes	Yes	Yes
Fixed effects			
Year of graduation	Yes	Yes	Yes
2-Digit training occupation	Yes	Yes	Yes
Other controls	Yes	Yes	Yes
N	14225	14225	14225
Adj. R-sq	0.280	0.212	0.295
F-test excl. IVs	25.78	42.01	32.36

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant; Year and year of employment dummies are not required since only one observation per apprenticeship graduate is included.

Table 5: Key Performance Measures for First Stages of IV Estimates

F-Test excl. IVs	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
A. Without Heterogeneous Treatment Effects:				
Job switch	265.6	492.1	250.9	467.3
Within-firm occ.-switch	221.6	402.9	226.5	411.3
Job-and-occ. switch	227.3	408.0	222.8	388.8
B. With Heterogeneous Treatment Effects:				
Job switch	158.3	271.5	140.7	250.7
Within-firm occ. switch	122.6	225.2	122.7	227.2
Job-and-occ. switch	125.6	215.2	120.4	205.3
$\left(tw(occup)_i - \overline{tw}_{job\_sw}\right) \cdot job\_sw_i$	239.9	222.6	226.1	236.2
$\left(tw(occup)_i - \overline{tw}_{occ\_sw}\right) \cdot occ\_sw_i$	60.3	89.9	65.2	89.9
$\left(tw(occup)_i - \overline{tw}_{occ\_job\_sw}\right) \cdot occ\_job\_sw_i$	180.0	142.9	165.9	134.7
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes

Table 6: Coefficient Estimates for IV Procedure without Heterogeneous Treatment Effects

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.109***	-0.0429***	-0.123***	-0.0373**
	[0.0231]	[0.0155]	[0.0271]	[0.0184]
Within-firm occ. switch	0.232***	0.143***	0.238***	0.124***
	[0.0233]	[0.0179]	[0.0285]	[0.0219]
Job-and-occ. switch	-0.0241	-0.0333*	-0.0327	-0.0305
	[0.0257]	[0.0184]	[0.0281]	[0.0215]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14225	14225	13378	13378
Adj. R-sq	0.011	0.186	0.026	0.131

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Table 7: Comparison of Coefficient Estimates for IV Procedure without Heterogeneous Treatment Effects for Various Sets of IVs

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
A. 12 IVs (instead of 22 as in Table 6)				
Job switch	-0.0702**	-0.0303*	-0.0745**	-0.0212
	[0.0276]	[0.0176]	[0.0325]	[0.0206]
Within-firm occ. Switch	0.266***	0.150***	0.283***	0.114***
	[0.0297]	[0.0205]	[0.0351]	[0.0263]
Job-and-occ. Switch	0.0170	-0.0198	-0.0171	-0.0261
	[0.0302]	[0.0209]	[0.0338]	[0.0245]
B. 9 IVs (instead of 22 as in Table 6)				
Job switch	-0.0421	-0.0322*	-0.0579*	-0.0248
	[0.0296]	[0.0182]	[0.0347]	[0.0215]
Within-firm occ. Switch	0.265***	0.146***	0.261***	0.106***
	[0.0323]	[0.0227]	[0.0373]	[0.0281]
Job-and-occ. Switch	0.0393	-0.0215	0.00561	-0.0198
	[0.0320]	[0.0214]	[0.0362]	[0.0252]
C. 7 IVs (instead of 22 as in Table 6)				
Job switch	-0.0430	-0.0330*	-0.0727**	-0.0277
	[0.0311]	[0.0187]	[0.0369]	[0.0221]
Within-firm occ. Switch	0.229***	0.124***	0.233***	0.0892***
	[0.0334]	[0.0234]	[0.0390]	[0.0292]
Job-and-occ. Switch	0.0199	-0.0314	-0.0181	-0.0284
	[0.0335]	[0.0221]	[0.0382]	[0.0261]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Table 8: Coefficient Estimates for IV Procedure with Heterogeneous Treatment Effects

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.0317	-0.0354**	-0.0522**	-0.0270
	[0.0204]	[0.0150]	[0.0253]	[0.0179]
Within-firm occ. switch	0.203***	0.122***	0.211***	0.116***
	[0.0207]	[0.0161]	[0.0248]	[0.0185]
Job-and-occ. switch	-0.0106	-0.0399**	-0.0155	-0.0362*
	[0.0225]	[0.0173]	[0.0260]	[0.0206]
$\left(tw(occup)_i - \overline{tw}_{job\_sw}\right) \cdot job\_sw_i$	-0.100	0.00127	-0.160*	-0.126
	[0.0626]	[0.0642]	[0.0848]	[0.0856]
$\left(tw(occup)_i - \overline{tw}_{occ\_sw}\right) \cdot occ\_sw_i$	0.300	0.484***	0.513**	0.619***
	[0.1953]	[0.1827]	[0.2432]	[0.1965]
$\left(tw(occup)_i - \overline{tw}_{occ\_job\_sw}\right) \cdot occ\_job\_sw_i$	-0.342***	-0.199*	-0.291**	-0.190
	[0.1092]	[0.1063]	[0.1206]	[0.1193]
$tw(occup)_i$	0.989***	0.839***	0.893***	0.742***
	[0.0382]	[0.0514]	[0.0447]	[0.0581]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14221	14221	13374	13374
Adj. R-sq	0.172	0.234	0.123	0.156

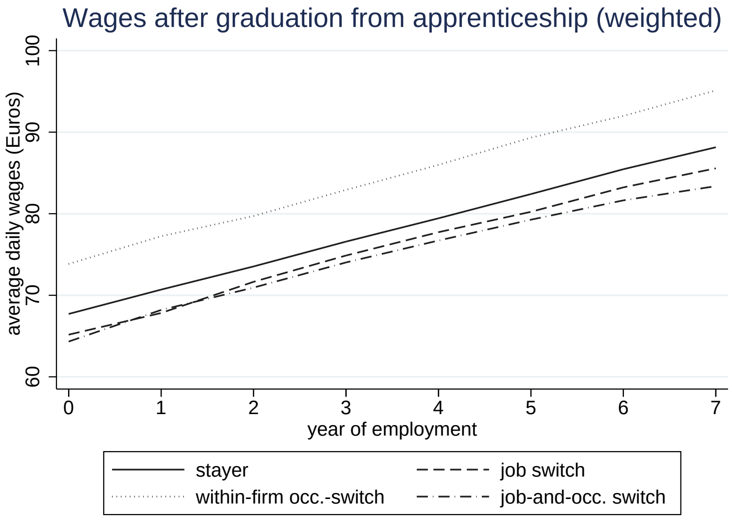
Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Table 9: Pooled OLS Estimates with Heterogeneous Treatment Effects

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.00510	-0.0147***	-0.0129**	-0.0131**
	[0.0051]	[0.0050]	[0.0063]	[0.0064]
Within-firm occ. switch	0.0658***	0.0668***	0.0614***	0.0585***
	[0.0074]	[0.0072]	[0.0083]	[0.0083]
Job-and-occ. switch	-0.0279***	-0.0348***	-0.0370***	-0.0391***
	[0.0058]	[0.0058]	[0.0068]	[0.0068]
$(tw(occup)_i - \overline{tw}_{job\_sw}) \cdot job\_sw_i$	-0.103**	-0.0622	-0.121*	-0.124*
	[0.0470]	[0.0476]	[0.0692]	[0.0703]
$(tw(occup)_i - \overline{tw}_{occ\_sw}) \cdot occ\_sw_i$	0.211*	0.170	0.272**	0.278**
	[0.1168]	[0.1193]	[0.1370]	[0.1320]
$(tw(occup)_i - \overline{tw}_{occ\_job\_sw}) \cdot occ\_job\_sw_i$	-0.393***	-0.378***	-0.346***	-0.314***
	[0.0725]	[0.0730]	[0.0835]	[0.0846]
$tw(occup)_i$	1.030***	0.948***	0.941***	0.830***
	[0.0298]	[0.0427]	[0.0343]	[0.0480]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14221	14221	13374	13374
R-sq	0.200	0.240	0.148	0.161

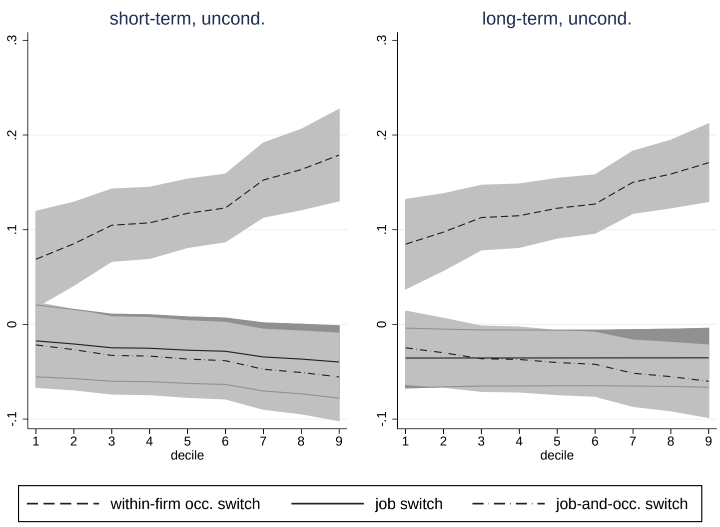
Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Figure 1: Wages after Graduation from Apprenticeship



Notes: Observations weighted by length of employment spell.

Figure 2: Average Treatment Effect on the Treated at Deciles of the Overall Distribution of Wages in the Training Occupation (Showing 95% Confidence Bands)



Notes: Calculations based on results from 3-step IV estimation controlling for 2-digit training occupations.

Appendix B: Data Cleaning Procedures and Identification of Completed Apprenticeships

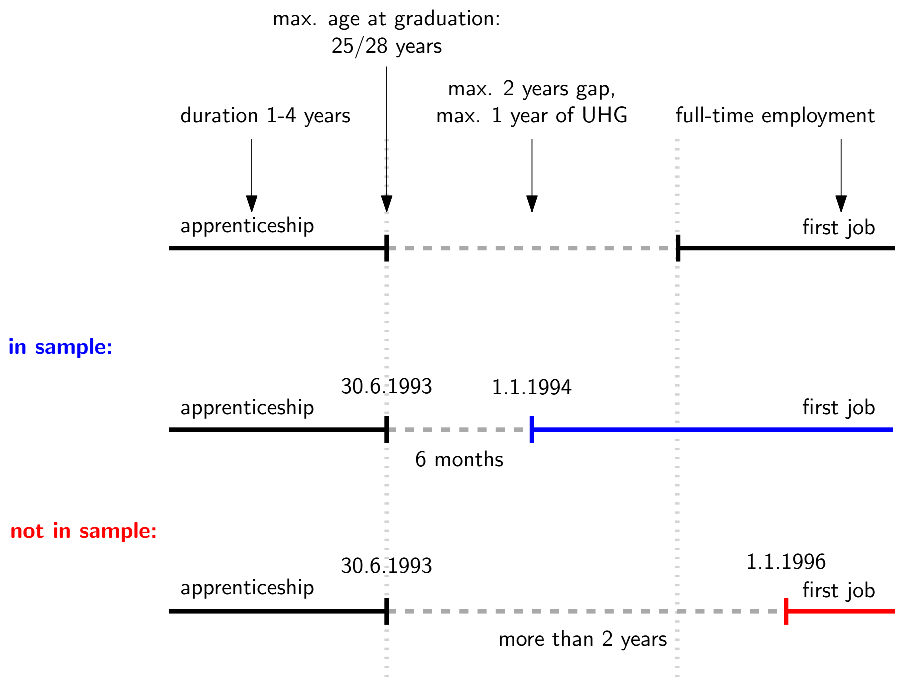
To identify an individual's first completed apprenticeship training, we apply a set of data cleaning procedures and restrictions to the IABS data. In order to identify whether an apprenticeship was successfully completed we need to observe a change in the education variable. Due to certain deficiencies of the education information provided by the IABS we use an imputed education variable based on imputation strategy ip1 proposed by Fitzenberger et al. (2006).

An apprenticeship episode observed in the data is identified as a person's first completed apprenticeship training if the following conditions are met. Figure 3 provides a summary of these conditions.

1. During the apprenticeship period, the individual is still observed as holding no vocational degree.
2. The information on the training occupation is non-missing in the last training spell.
3. The duration of training is at least one year. Also, we allow for a maximum duration of four years. For the observation period, the scheduled training duration lies between two and three and a half years (depending on the occupation) with an average of about three years. However, the training duration could be further shortened due to previous educational attainments such as holding a high-school diploma (*Abitur*). During the observation period about 19% of apprenticeship durations were shortened per year (see Uhly et al., 2006, figures 7.1 and 7.2).
4. Age at completion of training may not be more than 25 years for persons with no more than secondary education (*Hauptschulabschluss* or *Realschulabschluss*) and 28 years for persons with high school diploma (*Abitur*).
5. The education information changes to the status "holding vocational degree" within a period of two years after graduation from apprenticeship. This two-year window is long enough to allow us to observe changes in the education variable also for individuals doing military or civilian service right after their vocational training. At the same time, limiting the analysis to a two-year window makes it very unlikely that after graduation from apprenticeship the individual obtained a second vocational degree in a different occupation through types of training unobservable to us. Most importantly, fully school based vocational training would be unobservable to us. However, during the observation period most trainees in fully school based vocational training were female.[23] Another form of training unobservable to us would be further training programs, in which case participants could apply to the employment agency for a training allowance (*Unterhaltsgeld*). Thus, as a further restriction, during the two-year period individuals should not have received more than one year of training allowance.

[23]According to Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (1997, p. 67) during the years 1992–1995 about 80% of persons in fully school based vocational training (learning an occupation outside the dual system) were female.

Figure 3: Sampling Conditions



Furthermore, the following individuals are excluded from the sample:

1. Individuals whose training occupation of identified as "occupational code 130", since according to Drews (2008, p. 85) this category is also used for individuals whose training occupation is not defined yet.
2. Individuals who show earlier apprenticeship episodes lasting for longer than one year in a different occupation before the start of the main completed apprenticeship. (Shorter previous apprenticeship spells are allowed for, since they may well be internship spells that have been misclassified as apprenticeship training.)
3. Individuals for whom we observe further apprenticeship spells after graduation from apprenticeship.
4. Individuals who complete tertiary education (university degree, technical college degree) sometime during their further career.

Additional Online Appendix for:

Mobility across Firms and Occupations among Graduates from Apprenticeship

Forthcoming in: *Labour Economics*

By Bernd Fitzenberger, Stefanie Licklederer, and Hanna Zwiener

March 2015

Appendix C: Additional Tables and Figures

Table AOA.1: Definition of Four Mobility Groups (Number of Apprenticeship Graduates Sampled per Group in Parentheses)

		Change of firm	
		no	yes
Change of 3-digit occupation	no	stayer (n=6865)	job switcher (n=1961)
	yes	within-firm occupation switcher (n=1001)	job-and-occupation switcher (n=2187)

Table AOA.2: Distribution of Person-year Observations in the Wage Panel Across Four Mobility Groups by Year of Employment

	Year of employment							
Mobility type	0	1	2	3	4	5	6	7
stayer	58.41	57.72	58.46	58.98	59.13	59.16	59.29	60.11
job switcher	15.64	16.38	15.93	15.74	15.54	15.53	15.24	14.94
within-firm occ. switcher	8.41	8.33	8.46	8.44	8.48	8.48	8.65	8.80
job-and-occ. switcher	17.53	17.57	17.15	16.83	16.84	16.83	16.82	16.16
Total (N)	14225	12103	12251	12202	12141	12134	11971	11561

Notes: Sample share in the respective year of employment. Year of employment 0 refers to the year during which graduation occurred.

Table AOA.3: Coefficient Estimates for IV Procedure without Heterogeneous Treatment Effects (Standard Errors Clustered at Region-Year-of-graduation Level)

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.109***	-0.0429***	-0.123***	-0.0373*
	[0.0269]	[0.0165]	[0.0304]	[0.0193]
Within-firm occ. switch	0.232***	0.143***	0.238***	0.124***
	[0.0297]	[0.0213]	[0.0358]	[0.0256]
Job-and-occ. switch	-0.0241	-0.0333	-0.0327	-0.0305
	[0.0284]	[0.0209]	[0.0304]	[0.0234]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14225	14225	13378	13378
Adj. R-sq	0.011	0.186	0.026	0.131

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at region- interacted with year-of-graduation-level; Observations weighted by length of employment spell; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Table AOA.4: Overidentification Tests: Number of Rejections at 1% Significance Level among 26 Regions (Standard Errors Clustered at Individual Level)

No. IVs	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
0. Original set of 22 IVs				
	13	8	5	7
A. 12 IVs				
	5	4	5	2
B. 9 IVs				
	2	3	1	0
C. 7 IVs				
	1	2	3	0
Fixed effects for 2-Digit training occupation	No	Yes	No	Yes

Notes: Results of the overidentification test described in section 4.5 for the specifications discussed in Tables 6 and 7. 7 IV's: polynomial in unemployment rate, unemployment below 25, polynomial in labor market tightness. 9 IV's: 7 IV's plus share of low-skilled and high-skilled employees. 12 IV's: 9 IV's plus exit rate from employment into unemployment, indicator small cells, interaction mobility zero. 22 IV's: 12 IV's plus further mobility shares, interaction mobility zero, indicator small cells

Table AOA.5: Pooled OLS Estimates Accounting for Upward and Downward Mobility

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.0347***	-0.0249***	-0.0379***	-0.0220***
	[0.0056]	[0.0051]	[0.0066]	[0.0063]
Within-firm occ. switch up	0.107***	0.114***	0.0970***	0.0990***
	[0.0100]	[0.0093]	[0.0112]	[0.0110]
Within-firm occ. switch down	0.0430***	0.0534***	0.0398***	0.0470***
	[0.0110]	[0.0096]	[0.0128]	[0.0115]
Job-and-occ. switch up	-0.0123	0.00563	-0.0243**	-0.00455
	[0.0087]	[0.0084]	[0.0101]	[0.0100]
Job-and-occ. switch down	-0.0615***	-0.0653***	-0.0680***	-0.0654***
	[0.0077]	[0.0074]	[0.0086]	[0.0083]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14225	14225	13378	13378
R-sq	0.063	0.196	0.068	0.136

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell.

Table AOA.6: Coefficient Estimates for Two-step IV Procedure (no Heterogeneous Treatment Effects) Distinguishing Upward and Downward Occupational Mobility

Dependant variable: log(wage)	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	-0.108***	-0.0407***	-0.124***	-0.0374**
	[0.0229]	[0.0155]	[0.0268]	[0.0185]
Within-firm occ. switch UP	0.267***	0.175***	0.264***	0.159***
	[0.0268]	[0.0189]	[0.0324]	[0.0223]
Within-firm occ. switch DOWN	0.113***	0.0655***	0.112**	0.0509*
	[0.0367]	[0.0230]	[0.0438]	[0.0282]
Job-and-occ. switch UP	0.0383	0.0674***	0.00559	0.0227
	[0.0350]	[0.0244]	[0.0389]	[0.0293]
Job-and-occ. switch DOWN	-0.0472	-0.0808***	-0.0341	-0.0635**
	[0.0323]	[0.0220]	[0.0342]	[0.0250]
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes
N	14225	14225	13378	13378
Adj. R-sq	0.024	0.188	0.037	0.134

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Standard errors clustered at person-level; Observations weighted by length of employment spell; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant.

Table AOA.7: OLS Regression of Predicted Probabilities of Mobility on the Local Labor Market Conditions at the National Level (Pooling 26 Regions) Accounting for Upward and Downward Mobility

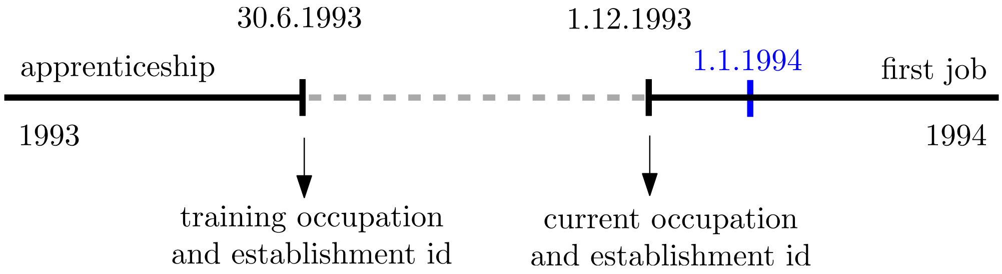
Dependent variable: Predicted probability of	Job switch (1)	Within-firm occ. switch up (2)	Within-firm occ. switch down (3)	Job-and-occ. switch up (4)	Job-and-occ. switch down (5)
Unemployment rate	0.0346***	-0.0157***	0.0120**	0.00549	0.0217***
	[0.0070]	[0.0054]	[0.0053]	[0.0055]	[0.0058]
Unemployment rate^2	-0.00371***	0.00142***	-0.000787	-0.000553	-0.00139**
	[0.0007]	[0.0005]	[0.0005]	[0.0005]	[0.0006]
Unemployment rate^3	0.000114***	-0.0000407**	0.0000281*	0.0000234	0.0000413**
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Unemployment rate < 25 years	0.00565***	-0.0000687	-0.00375***	-0.00142**	-0.00210***
	[0.0008]	[0.0006]	[0.0006]	[0.0007]	[0.0007]
Labor market tightness	-0.00180***	-0.000238	0.00138***	0.000337	0.000914**
	[0.0004]	[0.0003]	[0.0003]	[0.0004]	[0.0004]
Labor market tightness^2	0.0000492***	0.00000598	-0.0000353***	0.0000146*	-0.00000723
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Labor market tightness^3	-0.000000274***	-4.23e^{-08}	0.000000208***	-0.000000120***	-1.35e^{-08}
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Share low qualified	-0.000840**	-0.000365	0.000276	-0.000348	0.000372
	[0.0004]	[0.0003]	[0.0003]	[0.0003]	[0.0003]
Share highly qualified	0.00295***	-0.000435	0.00128***	0.00110***	0.00321***
	[0.0004]	[0.0003]	[0.0003]	[0.0003]	[0.0004]
Mobility shares					
Unemployment	-0.000214	-0.00155***	-0.00213***	0.000727*	-0.000347
	[0.0005]	[0.0004]	[0.0004]	[0.0004]	[0.0004]
Job switch	-0.00277***	0.00220***	0.00101***	0.000754*	0.00125***
	[0.0005]	[0.0004]	[0.0004]	[0.0004]	[0.0004]
Within-firm occ. switch	0.00236***	-0.000794***	-0.000794***	0.000701***	0.00209***
	[0.0002]	[0.0002]	[0.0002]	[0.0002]	[0.0002]
Job-and-occ. switch	0.00190***	-0.00287***	-0.000718***	0.000362	-0.000381
	[0.0003]	[0.0002]	[0.0002]	[0.0002]	[0.0002]
Further instrumental variables					
Interaction effects indicating small cells for mobility shares	Yes	Yes	Yes	Yes	Yes
Interaction effects indicating mobility share zero	Yes	Yes	Yes	Yes	Yes
Fixed effects					
Year of graduation	Yes	Yes	Yes	Yes	Yes
2-Digit training occupation	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes
N	14225	14225	14225	14225	14225
Adj. R-sq	0.283	0.123	0.110	0.165	0.199
F-test excl. IVs	25.78	25.87	14.71	8.50	27.87

Notes: Standard errors in brackets; * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$; Other controls include age at job entrance and dummies for high-school diploma, foreign citizenship, foreign citizenship missing and a constant; Year and year of employment dummies are not required since only one observation per apprenticeship graduate is included.

Table AOA.8: Key Performance Measures for First Stages of IV Estimates without Heterogeneous Treatment Effects Accounting for Upward and Downward Occupational Mobility

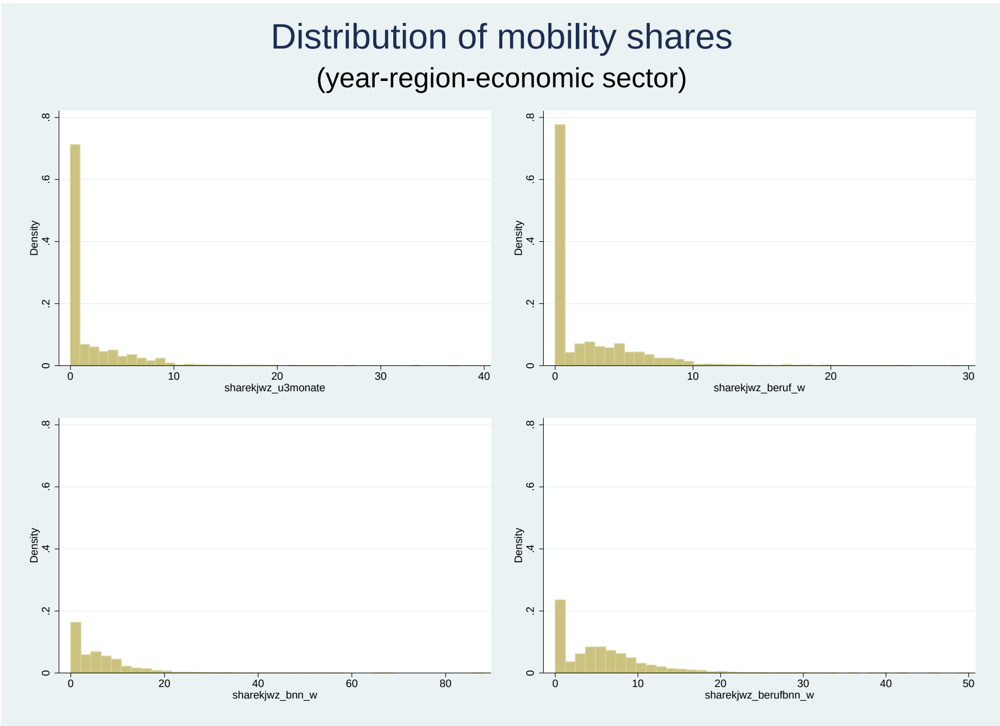
F-Test excl. IVs	Short term (0-2)		Long term (3-7)	
	(1)	(2)	(3)	(4)
Job switch	160.54	301.56	151.51	287.11
Within-firm occ. switch up	84.76	159.83	89.69	177.86
Within-firm occ. switch down	58.51	166.39	55.81	156.63
Job-and-occ. switch up	82.67	150.25	83.42	148.64
Job-and-occ. switch down	89.61	183.95	91.01	183.21
Fixed effects				
Year	Yes	Yes	Yes	Yes
Year of graduation	Yes	Yes	Yes	Yes
Year of employment	Yes	Yes	Yes	Yes
2-Digit training occupation	No	Yes	No	Yes
Other controls	Yes	Yes	Yes	Yes

Figure AOA.1: Apprenticeship and First Employment Spell with Interruption



Notes: This example shows the measurement of occupation and establishment id for an apprentice who graduated in June 1993. His first job held after apprenticeship starts in December 1993 and, thus, lies within the required two-year window after graduation.

Figure AOA.2: Distribution of Mobility Shares Showing Spikes at Zero for Each of the Four Mobility Groups



Notes: Upper panel, left: mobility share unemployment duration of at least 3 months. Upper panel, right: mobility share within firm occupational switch. Lower panel, left: mobility share job switch. Lower panel, right: mobility share job-and-occupation switch.

Figure AOA.3: Regional Distribution of Probability Scores for Job Switches (Resulting from Step 1 of IV Procedures, Short Run, Weighted)

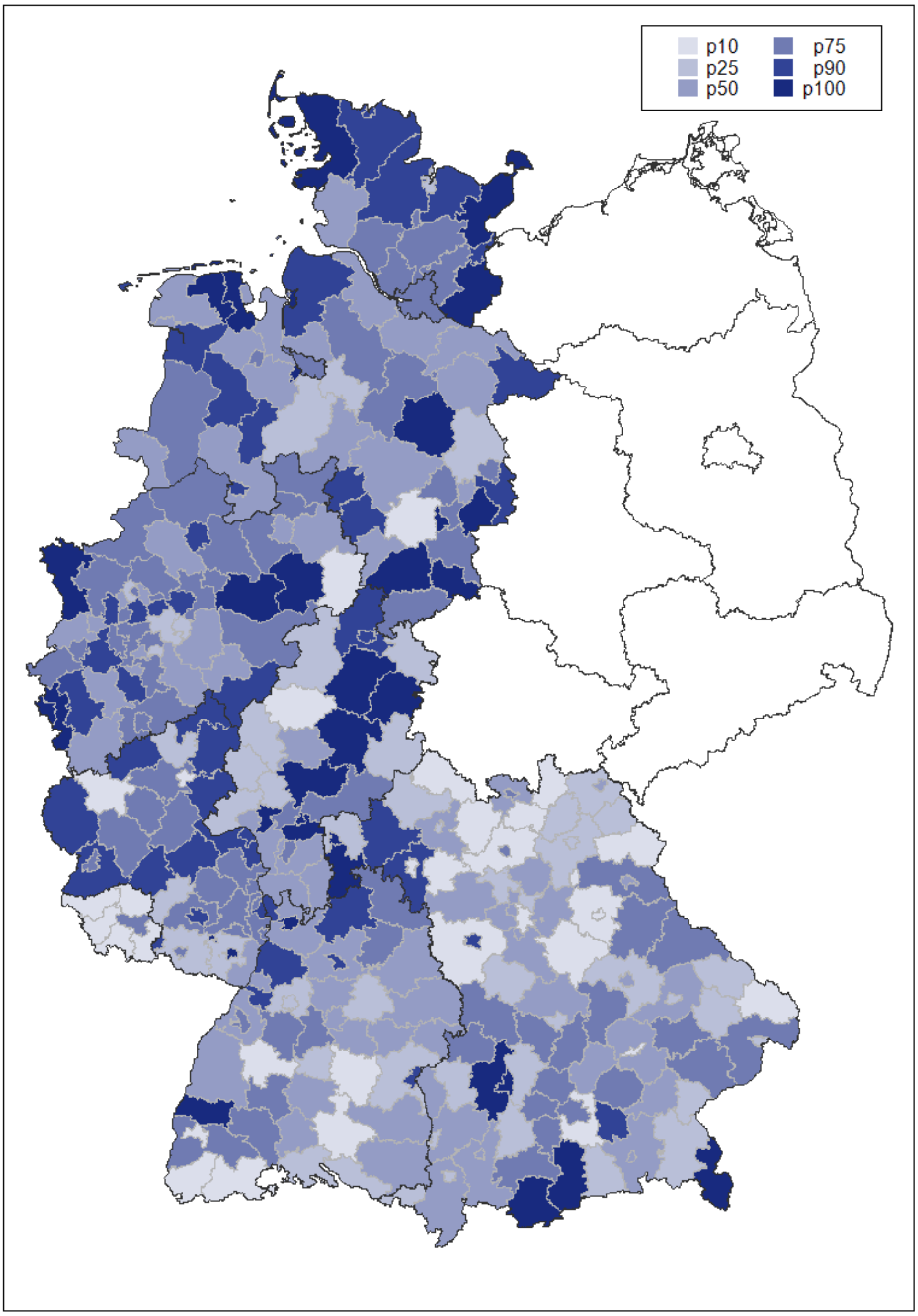


Figure AOA.4: Regional Distribution of Probability Scores for Within-firm Occupation Switches (Resulting from Step 1 of IV Procedures, Short Run, Weighted)

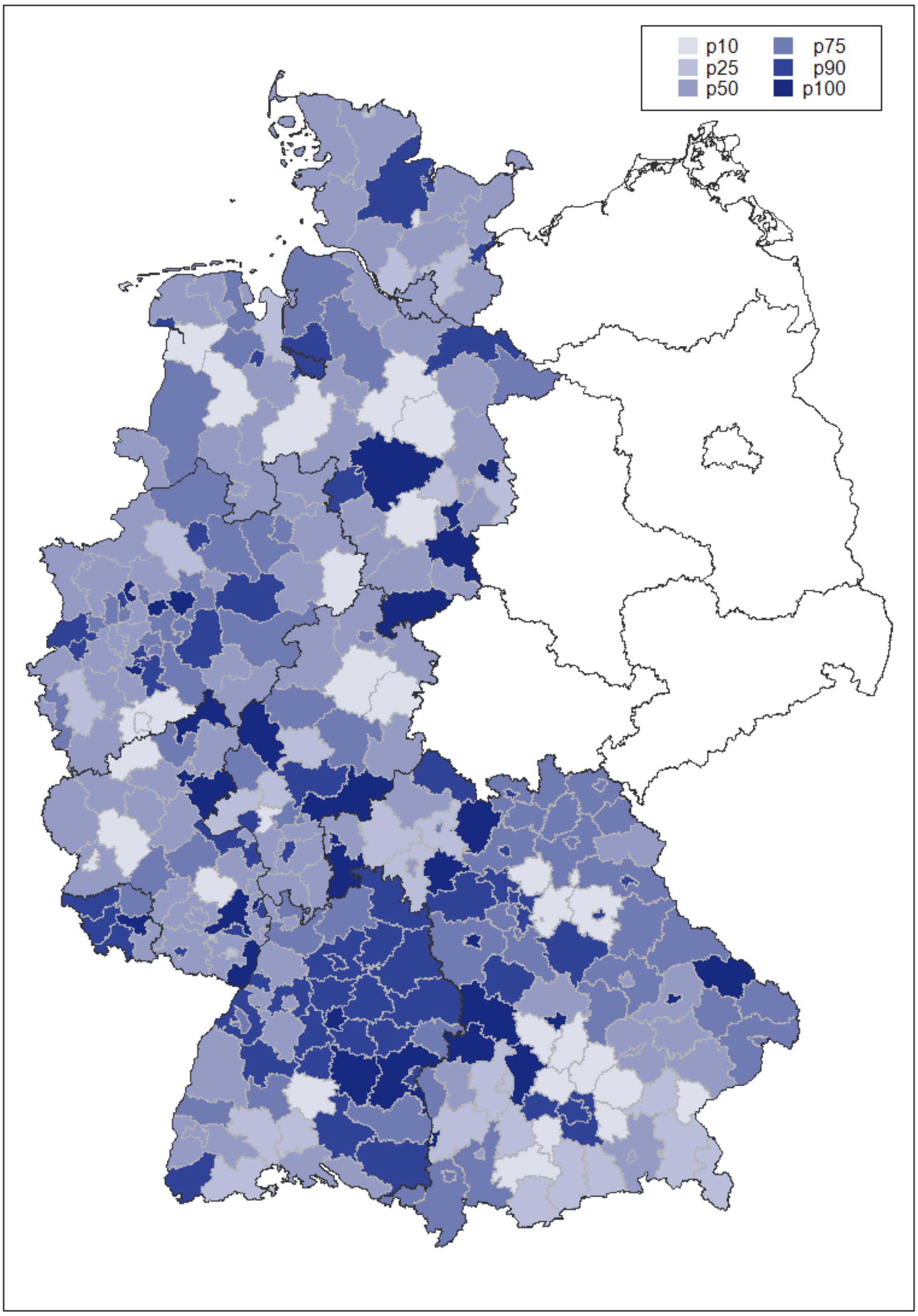


Figure AOA.5: Regional Distribution of Probability Scores for Job-and-occupation Switches (Resulting from Step 1 of IV Procedures, Short Run, Weighted)

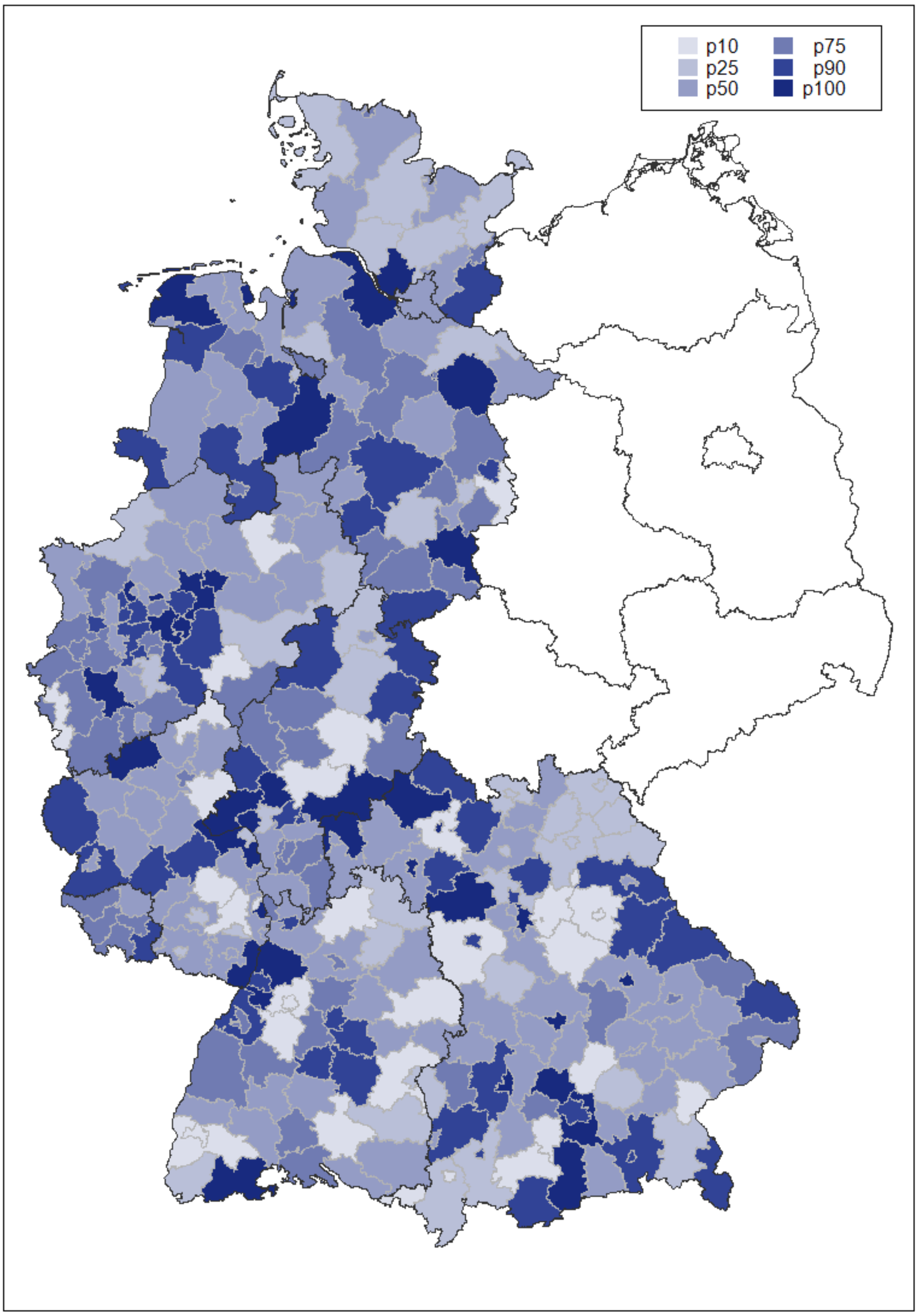
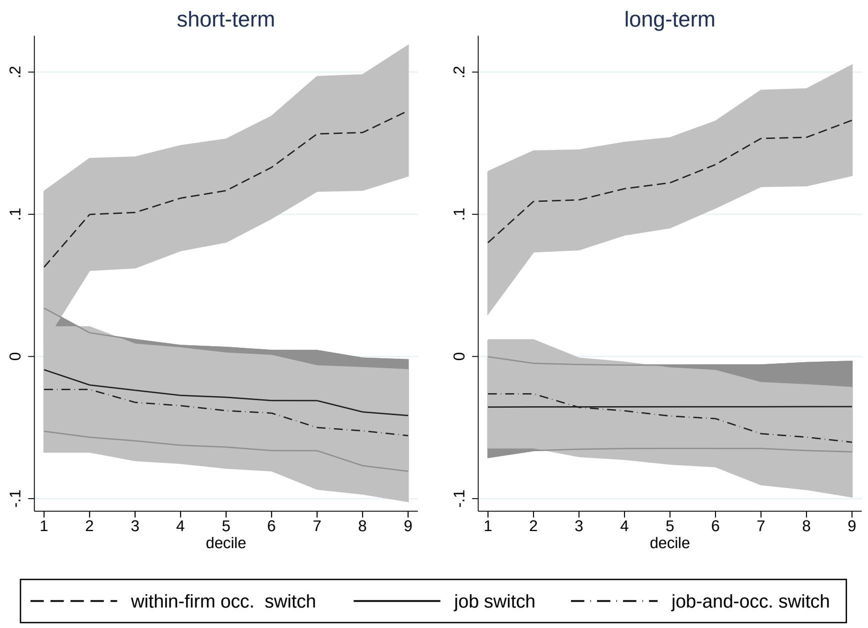
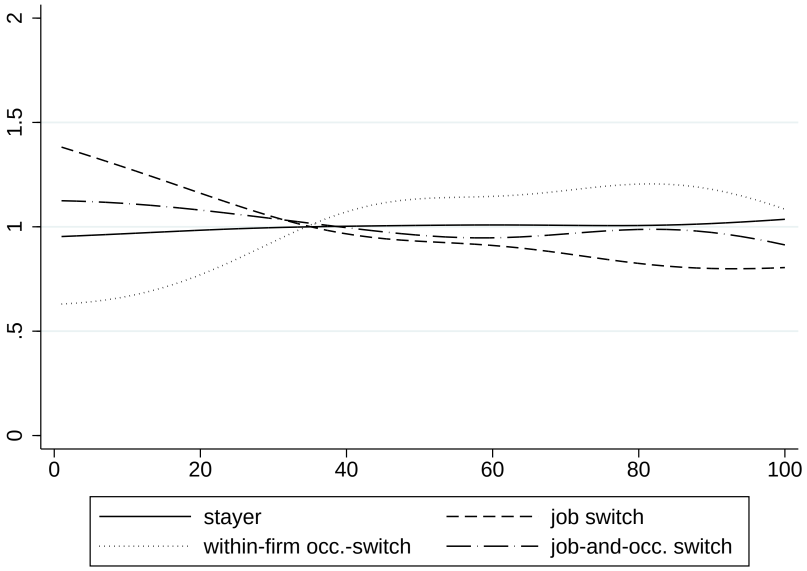


Figure AOA.6: Average Treatment Effect on the Treated at Deciles of the Group-specific Distribution of Wages in the Training Occupation (Showing 95% Confidence Bands)



Notes: Calculations based on results from 3-step IV estimation controlling for 2-digit training occupations.

Figure AOA.7: Relative Frequency of Wage Position of Training Occupation by Mobility Group



Notes: Occupations ranked from lowest paid (0) to highest paid (100).

Appendix D: Matching of Instrumental Variables Across Different Spatial Classifications

For reasons of data anonymization, regional information in the IABS regional file is not coded at the original level of administrative districts (Kreise), but at a slightly aggregated level (grouped districts) which ensures that the dataset only contains regional units of at least 100.000 inhabitants. We aggregate all instrumental variables which are provided at the original administrative district (*Kreise*) level to the grouped-district level. In this, we weight districts by their relative size in terms of the number of inhabitants. The required key matching administrative districts to grouped districts is provided in Drews (2008, pp. 69-78).

Additionally, some of the instrumental variables, such as the labor market tightness measure, are only available at the level of employment agency districts (*Agenturbezirke*). This creates a problem, since administrative districts and agency district may overlap. Some administrative districts actually belong to four different agency districts. This is farther complicated by the grouping of administrative districts in the IABS regional file. Taking all these complications and spatial overlaps into account, based on the comparison of maps of administrative districts and agency districts we create a key matching agency districts and grouped districts in the IABS regional file. For simplification, in the case that an administrative district strongly overlaps with several agency districts, we assume that the administrative district is equally distributed across all relevant agency districts. The key takes into account changes at the administrative district level during the period 1988-2011. Furthermore, we checked that no major changes in agency districts occurred during the period 1988-2011 – changes were few and insignificant.

For the regional Probit Analysis in stage zero we define 26 districts based on the German regional policy districts (*Regierungsbezirke*). We assign each grouped administrative district in the IABS data to the corresponding government district (see table AOA.9). Due to missing variation for the city districts Hamburg and Bremen, and small sample size, we group the initial 30 government districts into 26 regions.

Table AOA.9: Regional districts for Probit Analysis in Stage zero.

District	Description	#Obs.
1	Schleswig-Holstein	502
2	Lueneburg and Hamburg	632
3	Weser-Ems and Bremen	772
4	Hannover	469
5	Braunschweig	372
6	Muenster	586
7	Detmold	539
8	Duesseldorf	1044
9	Arnsberg	800
10	Koeln	785
11	Kassel	331
12	Giessen	211
13	Darmstadt	643
14	Koblenz	300
15	Trier and Saarland	311
16	Rheinhessen-Pfalz	359
17	Karlsruhe	544
18	Stuttgart	926
19	Tuebingen	391
20	Freiburg	507
21	Unterfranken	377
22	Oberfranken and Oberpfalz	648
23	Mittelfranken	452
24	Niederbayern	391
25	Schwaben	519
26	Oberbayern	823
		14234