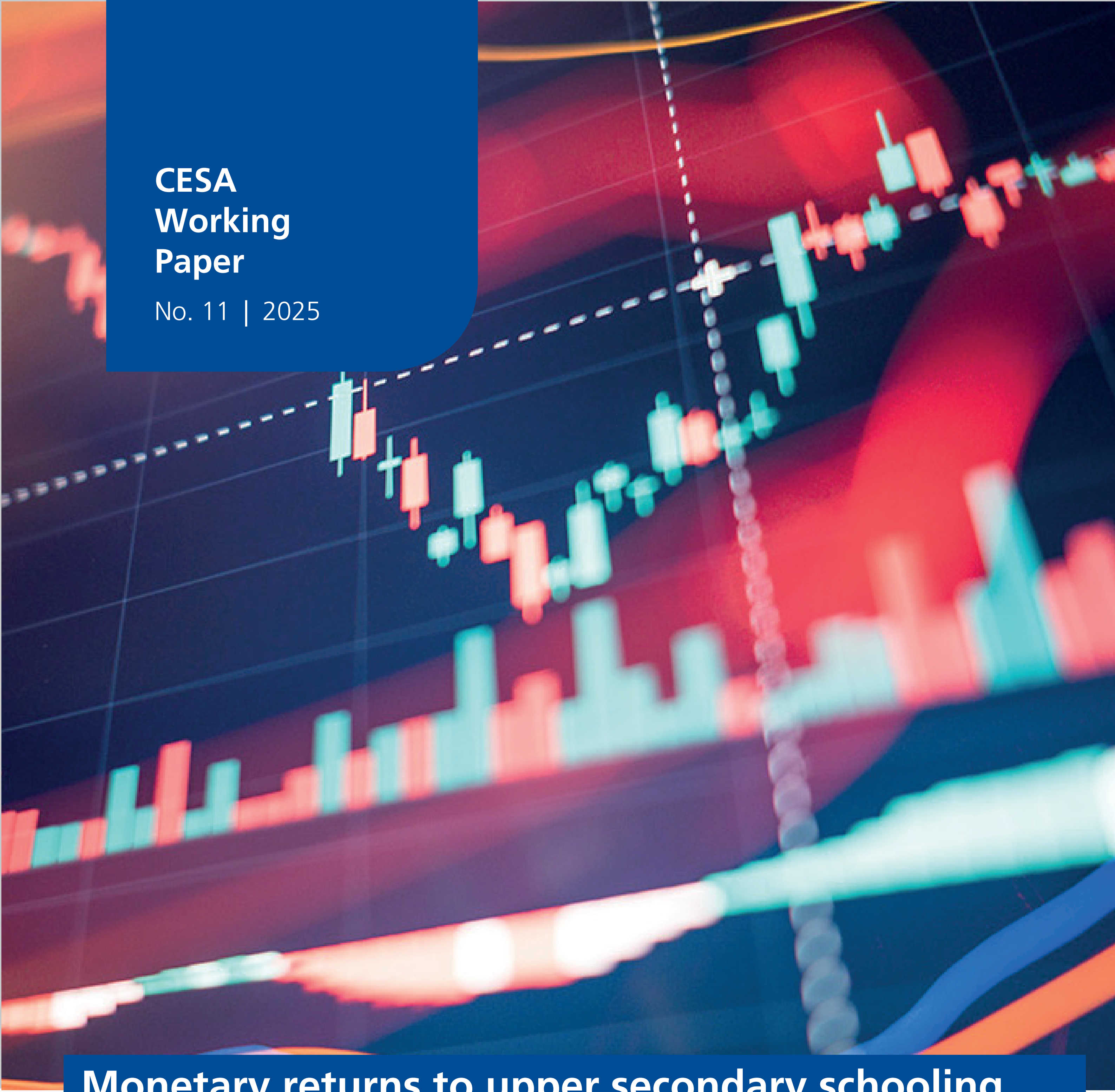


CESA Working Paper
No. 11 | 2025

Monetary returns to upper secondary schooling, the evolution of unobserved heterogeneity, and implications for employer learning

Anna Krumme[a,b] and Matthias Westphal[a,c]

a University of Hagen (FernUniversität in Hagen)

b TU Dortmund University

c RWI Essen

FernUniversität in Hagen

Monetary returns to upper secondary schooling, the evolution of unobserved heterogeneity, and implications for employer learning

Anna Krumme
FernUniversität in Hagen
TU Dortmund University

Matthias Westphal
FernUniversität in Hagen
RWI Essen

December 2024

Abstract

We study the evolution of monetary returns to high school education and their heterogeneity after the labor market entry using linked survey and administrative labor market data from Germany. By exploiting academic track school openings for cohorts from 1950–1985, we find sizeable monetary returns of 14–17% per year of additional schooling within the first 10 years of labor market experience. Whereas unobserved heterogeneity in the returns is initially uncorrelated with the schooling decision, the correlation starts evolving at higher levels of labor market experience. We interpret this finding considering employer learning – so far unconsidered in the literature.

Keywords: Returns to education, IV estimation, marginal treatment effects, unobserved heterogeneity, employer learning
JEL Classification: *I26, C26, J24*

Anna Krumme: FernUniversität in Hagen, Faculty of Business Administration and Economics, 58084 Hagen, Germany, E-mail: anna.krumme@fernuni-hagen.de.
Matthias Westphal: FernUniversität in Hagen, Faculty of Business Administration and Economics, 58084 Hagen, Germany, E-mail: matthias.westphal@fernuni-hagen.de.
Funding by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged (Project number WE 6751/2-1). We further thank Dennis Schlamann and Moritz Pickhardt for their excellent research assistance. We benefited from the valuable feedback from Damon Clark, Andreas Fischeneder, Hendrik Schmitz, and participants of the University of Hagen Workshop on Labor Economics, the Westphalian Research Workshop, the LIfBi Workshop on Economics of Education in Bamberg, the 14th International Workshop on Applied Economics of Education, the 39th meeting of the European Economic Association and the 76th European meeting of the Econometric Society, and the European Association for Labour Economists Conference 2024.

1 Introduction

Evidence on monetary returns to schooling has blossomed over the past two decades, crystallizing into the fact that ”human capital explains a substantial share of the variation in labor earnings within and across countries” (Deming, 2022). Beyond the fact that education is a worthwhile investment on average, many studies have shown that returns are individual-specific and may correlate with individual preferences and abilities for education (Card, 2001; Carneiro et al., 2011; Nybom, 2017; Westphal et al., 2022). However, less is known about how this heterogeneity is shaped throughout the working career: Is it that individual differences in the returns to education are determined already at the beginning of the career, with little change in the ranks and the variance of the returns thereafter – or does this heterogeneity need time to evolve? The answer to these questions may reveal at least two important insights into how a labor market functions. First, it may be that human capital complements labor market experience, such that education has increasing effects on marginal productivity. Second, it could be that human capital at labor market entry is unobserved, and employers may learn about the true earnings potential of their employees over their careers, such that differential wage increases reflect this learning process (Altonji and Pierret, 2001; Lange, 2007; Aryal et al., 2022).

In this paper, we approach these questions step by step by estimating monetary returns to academic track education, a German elite school type, which 32 percent of a cohort attended during our study. The academic track lasts 13 instead of 10 years, qualifying graduates for university studies or more abstract white-collar professions. To eliminate the confounding effects of selection into this school type, we exploit the massive opening of these schools throughout the German educational expansion with an instrumental variables approach. After estimating average returns, we go beyond the average in two important dimensions. First, we follow individuals over time, relative to their labor market entry, to assess when the returns are formed. Second, we unveil the evolution of the essential heterogeneity (Heckman et al., 2006) – a correlation between the unobserved returns and the propensity to take academic track education by estimating marginal treatment effects (MTEs). We then explore how the unobserved heterogeneity in returns to schooling changes after individuals enter the labor market and highlight the implications for employer learning.

We combine administrative data with survey data and can follow individuals from birth onwards over their careers in the labor market. While the survey data from the National Educational Panel Study (NEPS) allows us to analyze the residential history (such as the place of birth) and a rich set of background characteristics, the administrative labor market data from the Institute for Employment Research (IAB) provides us with detailed yearly information on daily wage and employment from the individuals' first appearance in the labor market onward. We merge self-collected information on all academic track schools in Germany. Overall, we have a data set of 3,469 male individuals.

Our paper contributes to several strands of the literature. First, we add evidence to the literature on monetary returns to secondary schooling. While many papers exist on the returns for compulsory schooling reforms (e.g., Bhuller et al., 2017), evidence on secondary schooling beyond the compulsory level is scarce because many secondary education systems are comprehensive, i.e., do not allow for general education choices until tertiary education starts. Exceptions include Clark and Del Bono (2016), who study labor market effects for elite schools, and Birkelund and van de Werfhorst (2022) and Matthewes and Ventura (2022), who study tracking decisions at the age of 16. We study long-run returns to tracking decisions at the age of 10 in Germany, thereby complementing Dustmann et al. (2017), who study the returns of students who are, in terms of academic capability, at the margin between two tracks but attend the higher track. They document that these students are downgraded throughout secondary schooling (relative to the academically marginal students attending a lower track), explaining their null results. In this dimension, our paper also explores the persistence of initial tracking decisions for a more general fraction of students whose tracking decision is affected by school availability. Second, we test for selection into gains, i.e., whether the unobserved propensity to attend a higher track correlates with future earnings returns. Most evidence on this unobserved heterogeneity in the returns to education focuses on college education (Kamhöfer et al., 2019; Carneiro et al., 2011; Nybom, 2017). In contrast, we want to add to Carneiro et al. (2017) and provide evidence on heterogeneous returns to upper secondary schooling. Third, we investigate the development of this heterogeneity over time, which, to our knowledge, remains unexplored. As the evolution of this heterogeneity may be shaped by employers who learn about the true earnings potential of their employees over time, we are (besides Aryal et al., 2022) the second to rely on instrumental variable approaches in assessing employer learning and the first to propose using marginal treatment effects for this endeavor.

We find substantial positive monetary returns of 14 to 17% per year of academic-track education within the first ten years in the labor market, documenting that initial track enrollment is essential. Yet, these effects do not exist directly after labor market entry but rise after the first two years of working experience. Our results further suggest that the returns are heterogeneous for higher experience levels, revealing clear selection into gains. Taken at face value, our results imply that individuals with the lowest desire for academic track schooling may exhibit negative or zero returns later in their careers. Therefore, a further expansion of academic track education that most likely affects these individuals would arguably not pay off. We document that essential heterogeneity evolves substantially with experience. We then show that two productivity-related mechanisms might drive this pattern. First, individuals with higher resistance to academic-track education might have lower productivity growth over experience than those more prone to choose this high educational path. Second, employers might learn about innate abilities as one component of productivity over time and adjust wages according to their beliefs on true productivity. Finally, we discuss the conditions under which we can prove the occurrence of employer learning.

This paper proceeds as follows. Section 2 presents the institutional background, the data set, and the baseline empirical strategy. We present the baseline results in Section 3. Section 4 introduces the marginal treatment effect estimation and the interpretation of potential heterogeneity in this context before we present and discuss the results. Section 5 concludes.

2 Institutional Background, Data, and Baseline Identification Strategy

2.1 Institutional Setup

The (West-)German Secondary Schooling System

After four years of elementary school and at age 10, the elementary school teacher recommends children to one of three different secondary school types based on their (perceived) performance. Although elementary school teachers give recommendations for the track choice in all states, the parents are the final decision-makers in most of the West German states.[1] We refer to these tracks as the basic, intermediate, and academic.[2] The basic and intermediate track education aims at preparing students for apprenticeships in blue- or white-collar jobs and ends typically after 5 or 6 years, respectively.[3] The academic track education at a *Gymnasium* prepares students for tertiary education at colleges and universities. It lasts nine years for all West German students (before the graduating class of 2007). After completing 13 years of schooling and passing final exams, the graduates achieve the *Abitur*, i.e., the highest schooling degree and university-entrance diploma.

The different school types also show content-related differences. Students in the academic track have more weekly teaching hours, and priorities are set on second or third foreign

[1] Before 2010, in 2 of the 10 West German states (excluding Berlin), the recommendation of the elementary school was widely binding; deviations require an official procedure but are generally possible. For more information, see `https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf`.

[2] Since 1971, comprehensive schools have been founded that accommodate all students. However, they have played a minor role as only a small percentage of students attended this comprehensive school type. Until 1990, the share of students at a comprehensive school out of all students at general schools never exceeded the 10% limit, and less than 3 % of all graduates with Abitur received their degree at a comprehensive school (Köhler and Lundgreen, 2014).

[3] Compulsory schooling reforms between 1956 and 1969 increased the basic track duration from five to six years.

languages and natural sciences rather than social sciences, sports, or vocational preparation (Dustmann et al., 2017). In the core subjects of mathematics and German, the contents are typically more advanced the higher the track (for more detail on the differences between tracks, see Dustmann et al., 2017).

Academic Track School Openings during the Educational Expansion

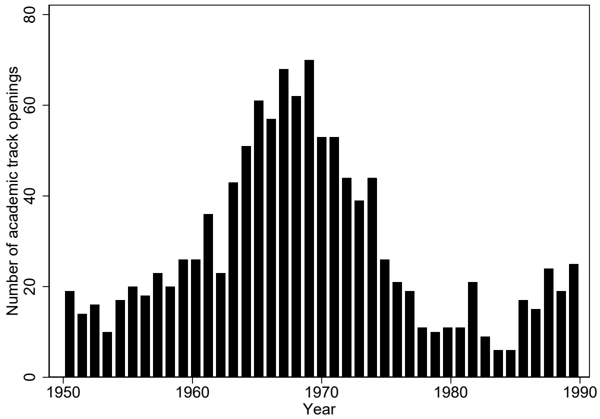
In the 1950s, educational opportunities were scarce. For example, in 1952, only 12.4% of all grade 8 students attended the academic track, and not more than 3.8% of all school-leavers were academic track graduates (Köhler and Lundgreen, 2014). These low numbers are caused by the supply rather than the demand side. In 1950, only 1,823 academic track schools (Franzmann, 2006) and 33 universities (Kamhöfer et al., 2019) existed in the entire Federal Republic of Germany territory. However, public and political opinion changed, and geographical or economic barriers should no longer restrict access to education (Becker, 2006). In the next decades, Germany's educational infrastructure changed substantially. This is referred to as the educational expansion, which we use to identify the monetary returns in this paper.

Economic and sociopolitical arguments drove the educational expansion in the early 60s. Picht (1964) proclaimed an education crisis that drastically impacted the country's economic situation and demanded public investments in higher education to ensure ongoing economic growth. At the same time, attention shifted more and more to equal educational opportunities, e.g., Dahrendorf (1965) requested "education as a civil right." Therefore, one of the main goals of educational reforms since the 1960s was to facilitate access to education. While the educational expansion also affected tertiary education,[4] the focus was on expanding higher education, as, for example, Picht (1964) called for doubling the number of graduates with Abitur.

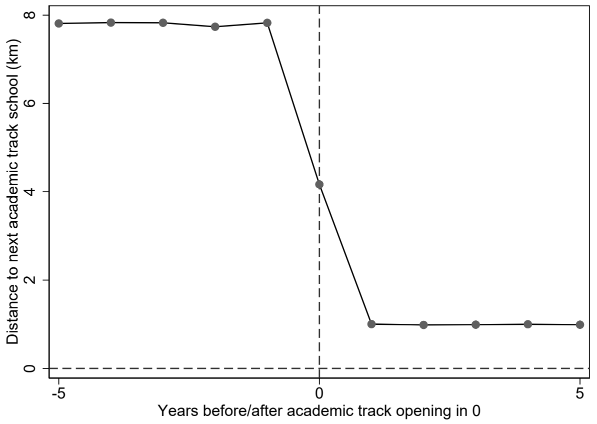
Between 1960 (the earliest secondary schooling entry cohort in our sample) and 1990, yearly public expenditures for academic track schools increased from 1,130 to 11,559 million Deutschmark. This increase led to the foundation of 796 academic track schools (relative to 1,396 existing schools), and the number of students nearly doubled from around 850 thousand to 1.6 million within this period (Franzmann, 2006). Figure 1 visualizes the distribution of academic track openings over time (Panel a) and space (Panel c) and presents the effect of the openings on the distance to the closest school (Panel b). We see at least six openings annually between 1960 and 1990 (our cohorts' main track decision years), with most openings, i.e., more than 40 per year, concentrated in the decade following the mid-sixties. This benefits our analysis because our panel is balanced for these cohorts. Panel (c) shows the spatial distribution of academic track schools. The black municipalities

[4] Between 1950 and 1990, the number of colleges doubled in West Germany (see, for example, Kamhöfer et al., 2019 for information on college openings in Germany).

Figure 1: Distribution of academic track school openings and effects on the average distance



(a) Over time



(b) Relative to a school opening

(c) Across West Germany

Notes: Own illustration based on self-collected school information. Panel (b) is restricted to municipalities with a first opening. In opening year 0, typically half of the students (based on school year start and birth months) already have access to the newly opened school.

already had an academic track school in 1950. The red municipalities experienced the first opening of an academic track school between 1950 and 1990. This demonstrates that the political goal of the educational expansion seems to be fulfilled: more remote areas disproportionately benefit from improved educational opportunities. To visualize that this improvement also affected neighboring municipalities, we plot a 10km radius around the new academic track in light red. Finally, we quantify the increased accessibility of academic track schools in Panel (b), where we plot the average distance to the next academic track school relative to the opening year for municipalities with an opening. It documents that students had to travel about 8 km on average to the closest academic track school before one opened in their municipality (measured from municipality centroid to school location).

2.2 Data

Our main data set is the NEPS-SC6-ADIAB (Bachbauer et al., 2022), which links administrative records from the IAB to survey data on educational paths from the Leibniz Institute for Educational Trajectories (LIfBi). This data covers adults born between 1944 and 1989 and living in Germany.

The NEPS-SC6 data includes 17,140 individuals whose educational trajectories we know and their approximate residence history at a very local level. The latter is crucial information for our project since we need information on residence before the individuals choose a secondary school track to correctly assign the data on geographical access to academic track schools. We use the municipality of residence in the last year of primary school whenever available and fill this information with the municipality of births else (around one-half of the cases). We include a dummy for the origin of the residence information in all regressions to control for systematic differences.

Most of the NEPS can be linked to the high-quality administrative labor market data provided by the IAB covering 1975 to 2019. A share of 74% of the surveyed individuals consented to link their responses to and were identified in the longitudinal labor market data of the IAB. These individuals appear in the administrative data if they ever had insurable employment subject to social security or marginal part-time employment (since 1990), have registered as jobseekers, or have participated in a labor market policy measure of the Federal Labor Office. Therefore, except for self-employed and civil servants, we have nearly complete labor market trajectories on individuals in the workforce.[5]

The second data set is a purpose-built collection of academic track school openings, presented in Figure 1. Starting with a list of all existing schools in 2010, we manually gathered their opening years. With additional information on the addresses, we end up with the precise geolocation of all schools. We then built a panel with information on geographical access to an academic track school for individuals from a particular municipality and birth cohort. Since there were differences in the starting times of school years, we took the varying school entry cutoffs into account (for details, see Koebe and Marcus, 2022, and Dustmann et al., 2017).[6] Additionally, we distinguished between boys- and girls-only schools. Any distance measures rely on the geodetic distance from the

[5] More detailed information on the linked data products are provided by Bachbauer and Wolf (2022).

[6] Precise information on the school year starting times by federal state and birth cohorts can be found in Appendix B in Koebe and Marcus (2020) for cohorts born between 1935 and 1965. There were no further changes for cohorts born after 1965, and school years started uniformly in August in all West German states (Dustmann et al., 2017).

center point of the municipality to a school because we have no access to residential information that is more local than the municipality. As long as the distribution of the residences around the centers is random, this does not lead to any problems, except that the access measures might be slightly less precise. However, the level of instrument assignment is more local than in most previous analyses using geographical access to school education. The IAB and the LIfBi research data centers merged our data set with the NEPS-SC6-ADIAB based on the municipality of childhood residence, gender, birth year, and month.

We restrict our sample to males born between 1950 and 1985 in Germany and follow them in the labor market from 1975 to 2019. Hence, we have a balanced panel for individuals aged 25 to 34, i.e., the decisive wage-formation years after labor-market entry. We exclude females to rule out fertility decisions and avoid academic-track-induced selection into employment and hours worked, which may be pronounced for females in the studied cohorts (see Westphal et al., 2022, for a discussion about university education). To ensure comparability in the educational system, which was different in East Germany before the reunification, we restrict the data set further to individuals with West German municipalities of residence in childhood, excluding Berlin. We additionally exclude males with missing information on schooling degrees, childhood residence (municipality and district level), or geographical access to academic track schools.

The final yearly panel includes 3,469 males for whom we observe earnings for one up to 10 years of labor market experience. We use actual work experience taken from the labor market trajectories of the IAB data, i.e., the number of complete years worked in full- and part-time at the end of the year. Thus, an experience of one indicates the year the individual entered the labor market for the first time, and an experience of 10 indicates being in the 10th year with nine fully completed years in the labor market, respectively. Excluded are any employment episodes before entering the labor market regularly, such as vocational training, student side job, or internship episodes. Since the experience only increases by one if one additional full year of employment is completed by the end of the observation year, individuals might be observed more than once at some experience levels. This is a result of any unemployment episode. Furthermore, not all individuals are observed when completing ten years of employment. Hence, our panel of 33,929 person-year observations is unbalanced across the experience.

Our treatment variable, "Abitur," indicates a university-entrance diploma (ordinarily attained by academic track graduation). Notably, Abitur excludes the vocational baccalaureate diploma (*Fachabitur*), which restricts one to studies of specific subjects at universities of applied sciences (*Fachhochschulen*). With 1,125 of 3,469 individuals in the sample, around one-third have Abitur.[7]

Dependent Variable

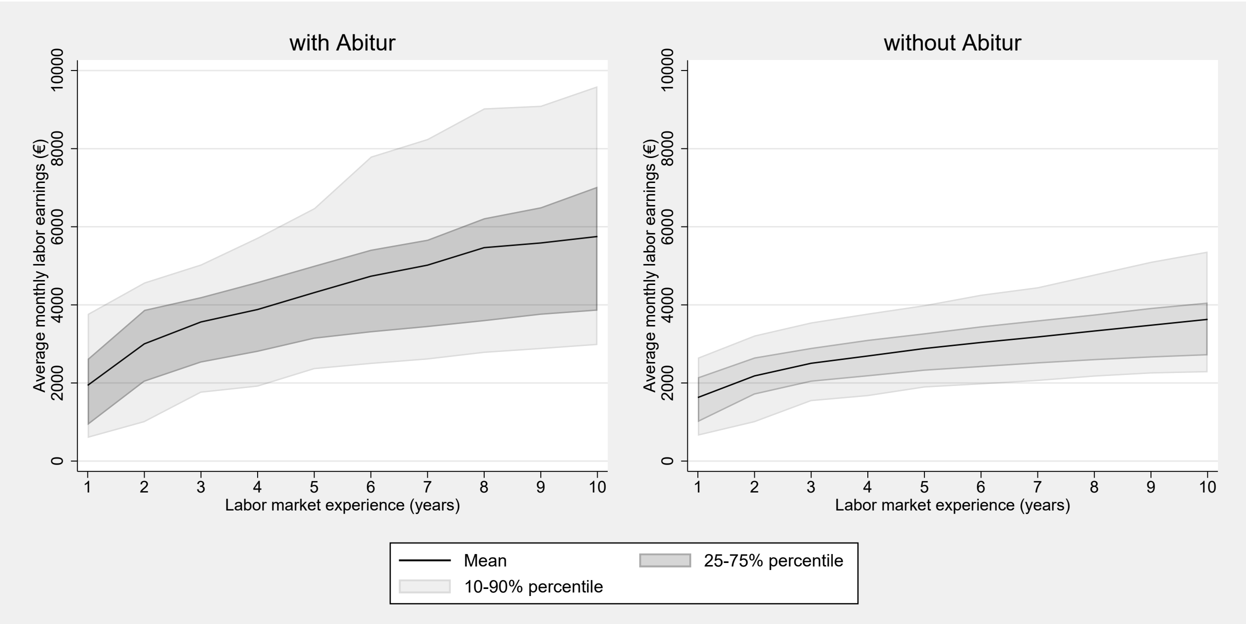
The gross daily wages for each job episode the employer obligatorily reports are the most crucial information for our analysis. We transform this information first into monthly labor earnings and then collapse it into a yearly panel using months with employment only. We further restrict our analysis to years with full-time employment.[8] We cannot use hourly wages because we do not have administrative information on hours worked. Nonetheless, because males' hours worked are inelastic in Germany (as they predominantly work full-time), we expect this measure to proxy hourly wages well. In preparing the earnings variable, we follow Dauth and Eppelsheimer (2020) and deflate daily wages with the consumer price index with the base year 2015. We also use the suggested procedure to identify and impute wages above an assessment ceiling. This is necessary because wages are only reported up to an upper limit for the statutory pension scheme, which varies with location and time. This step is particularly relevant for our paper, as we want to analyze the whole range of heterogeneity in monetary returns.

Figure 2 reports the average monthly labor earnings per year (deflated to 2015 prices) and the 10, 25, 75, and 90 percent quantiles for individuals with and without Abitur for experience levels from 1 to 10 years. After entering the labor market, the average monthly gross earnings for individuals with Abitur amounts to 1,940.02€ . These earnings increase substantially in the first two years, with an annual growth rate of 55 and 19 percent, respectively. Afterward, these earnings still grow steadily (reaching 5,751.54€ in the 10th year), but the growth rate declines from 9 to about 3 percent. Albeit on a lower level, earnings without Abitur also increase markedly across experience (from 1,628.25€ to 3,628.96€). However, not only the absolute but also the relative gap increases. While the relative Abitur earnings differential is 19 percent in the entering year, the gap constantly widens to 64 percent in the 8th year of experience. After that, the relative difference seems to remain at around 60 percent. The quantile trajectories additionally reveal an increased dispersion for academic track graduates, whereas the wage disparity remains relatively constant for individuals without Abitur. In sum, this figure reveals the importance of the early labor market years for wage formation, the Abitur earnings premium, and a potential of unobserved heterogeneity affecting this premium. We will explore both of these in greater detail in the succeeding analyses.

[7]The share of individuals that attended the academic track once in their educational trajectories is slightly higher with 38%, which highlights that initial tracking decisions are highly persistent, even though up- and downgrading is possible.

[8]Full-time work in the IAB data depends on the ratio of own to the establishment-specific common working hours. Missing values in this variable are replaced with information from the NEPS survey data.

Figure 2: Descriptive earnings trajectories over experience by academic track degree



Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure shows means for individuals with and without the academic track degree (Abitur) over labor market experience.

2.3 Baseline Empirical Strategy

Because individuals and their parents may choose the secondary school track, academic track education is endogenous, and a simple regression of earnings on Abitur (plus further observable characteristics) will not yield the causal effect of Abitur. We try to fix this problem using local changes in the academic track choice set as a quasi-experimental variation. Specifically, we use openings of academic track schools, measured by the variable Z_{it} (specified below), which we assume to be exogenous to individual decisions. With such a variable, we can recover the causal effect through the following two-stage least squares (2SLS) regression model

$$\begin{aligned} D_i &= \pi_0 + \pi_1 Z_i + X_i'\gamma + \nu_i \\ Y_{it} &= \alpha_t + \beta_t \widehat{D}_i + X_i'\delta_t + \varepsilon_{it}. \end{aligned} \tag{1}$$

Because the outcome variable changes for different values of experience t, we estimate these two equations separately for each experience level. In the first regression (the first stage), we regress the Abitur indicator of individual i (D_i) on the instrument Z_i and a vector of control variables X_i. Under the assumptions provided below, π_1 captures the share of individuals who only attend an academic track school because of changes in Z_i (i.e., the opening of an academic track school in the surrounding), but not without. This

An individual is assumed to work full-time if the self-reported contractual working hours are greater or equal to 39.

is the group of compliers. For this group, the second stage estimates causal effects on monthly earnings Y_{it} by labor market experience t. To this end, we regress Y_{it} on the fitted values from the first stage $\widehat{D}_i$ plus the same set of controls. Our interest lies in the coefficient β_t, the absolute effect of Abitur on labor earnings with experience t. We cluster our standard errors at the municipality level in all our regressions.

We need to assume that, conditional on the controls X, the instrument is unrelated to unobserved factors that correlate with education (i.e., $\text{Cov}(\nu_i, Z_i) = 0$) and wages $\text{Cov}(\varepsilon_{it}, Z_i) = 0$). To make this assumption credible, we include a large set of fixed effects in X_i, which we specify below. We additionally assume that other municipality-level changes do not coincide with the timing of the academic track openings (exclusion restriction). This assumption seems plausible given a rich set of control variables that absorb general trends. However, it may well be that due to general equilibrium effects, the quality of newly opened schools and teachers differ. We attribute these to the potential mechanisms behind our results. Finally, we assume that the openings did not deter individuals from attending the academic track (monotonicity). While we cannot test this on the individual level, we regard this as unlikely, given that the overall accessibility of academic track education improves through the openings.

Control Variables

The most important control variables are cohort and district fixed effects to absorb differential trends, wage levels, education, and academic track availability between districts. As the federal states may have implemented the educational expansion differently, we also account for federal-state-specific linear trends. Moreover, we include fixed effects for the birth month. In a second specification, we additionally include a dummy variable that indicates if the individual's birth municipality already had at least one academic track school in the decision year of the oldest individual in our sample. By absorbing any time-constant differences across municipalities, this dummy ensures that only openings after that can affect our estimates. Therefore, the specification without this "pre-existing academic track dummy" uses any variation in the geographical access to academic track schools across different municipalities in the same district. Table A.1 in the Appendix provides information on all covariates of the main specifications and variables used in any additional regression.

To make the independence assumption most credible in our setting, besides school entry cohort dummies, our ideal specification would include municipality-fixed effects to control for any time-constant differences that might affect school openings at the local level. Unfortunately, it is not feasible with our sample, which has fewer observations than municipalities. Hence, we are restricted to district-fixed effects, leading to the identifying assumption that the location and timing of academic track school openings are exogenous within districts. Since the educational expansion aimed at improving the comprehensive supply of (higher) schools to reduce social inequality, academic track schools should not only be opened but also in municipalities with specific inhabitant characteristics that also affect the later earnings of the children. The most substantial expectable differences between municipalities are most likely driven by bigger cities that might be systematically distinct from smaller and more rural municipalities [9]. We mainly control for these differences by including district-fixed effects, as urban districts are not subdivided into municipalities in Germany. This is also the case for the three German city-states. Furthermore, according to Henz and Maas (1995), already for the birth cohorts 1959–61, there were no urban-rural differences in the probability of attending a Gymnasium anymore. This might justify the assumption that municipality-specific inhabitant characteristics affecting later wages did not promote academic track school openings in certain municipalities within districts. Therefore, in both specifications, the location and timing of the openings might be random. However, the differences between municipalities that already had academic track schools before, those where academic track schools were opened, and those that never established any might be problematic. We expect some municipalities with existing academic track schools to be predominantly large cities. Using the same argument as before, district-fixed effects thus already partially control for the problem. Nevertheless, we explicitly control for the differences between these groups in our second specification, including the dummy for pre-existing academic track schools in the birth municipalities.

Instruments

We use two different measures as the instrument Z_i. The most intuitive measure for academic track school availability is a dummy variable indicating at least one existing school in the municipality ("academic track in municipality"). We use this binary instrument in our main specification to estimate the LATE. We additionally use a continuous instrument, especially for our subsequent MTE estimation, and construct an academic track availability index like Kamhöfer et al. (2019) did for universities. We consider the distances to the three closest academic track schools, compute an inverse distance weight, and aggregate these weights for each municipality and year. Specifically, the index is defined as $Z^I_{m,\tau} = \sum_{l=1}^{3} K\left(\frac{x_{m,l,\tau}}{5}\right)$ with $K(\cdot)$ denoting the Gaussian Kernel and $x_{m,l,\tau}$ being the distance in km from municipality m to its first, second and third ($l = 1, 2, 3$) closest academic track school in year τ (to distinguish it from the labor market experience t). The denominator, or bandwidth, determines how quickly weights decay in the distance. We use a bandwidth of 5 km. For example, schools within a 1 km distance enter the equation with 0.39, whereas schools 5 km away are valued lower with 0.24. Schools 10 or more km

[9] For example, universities, labor market conditions, or any infrastructural advantages can attract significantly higher educated families with children that might have better probabilities for having high wages in their later lives.

away enter the equation with only small values of 0.05 or less. Thus, individuals with more schools nearby have a greater index than those living farther away from the next three academic track schools.

Table 1 provides summary statistics for the dependent, treatment, and instrumental variables. While we have already discussed the monthly labor earnings and Abitur variables, we now discuss the instruments. The academic track indicator reveals that nearly two-thirds of our sample have an academic track school in their municipality. The continuous index additionally documents a similar range, mean and standard deviation. The background information on the number of schools and distance to the nearest school, which determine the index, shows that some individuals in our sample are close to the maximum possible value (1.197), which results if the three nearest schools lie on the municipality's centroid. In contrast, the lowest value is zero. The mean value of the index is 0.7, which would be implied if the three nearest academic track schools were 5.15 km away. For instance, the average distance from the school to the municipality centroid is 3.5 km, ranging from zero to 35. Likewise, the number of schools varies from zero to 65 (in large cities like Hamburg and Munich), with a mean number of about 5.

Table 1: Descriptive statistics of earnings, academic track education, and instruments with background information

	Statistics			
	Mean	SD	Min	Max
Labor market outcome:				
Monthly labor earnings	3,197.365	1900.851	13.078	33,290.110
Educational outcomes:				
Abitur	0.324	0.468	0	1
Academic track attendance	0.380	0.485	0	1
Instruments:				
Academic track in municipality	0.632	0.482	0	1
Index (Z^1)	0.702	0.426	0	1.194
Background information on instruments:				
Distance to 1st nearest school	3.428	4.340	0.1	35.3
Number of schools in municipality	5.198	10.285	0	65

Notes: Own calculations based on NEPS-SC6-ADIAB data.

3 Results I – Baseline IV

First stage and Instrument Balancing

Panel A of Table 2 shows the first stage effects of the instrument on attaining the Abitur (academic track certificate), the main treatment variable, in our sample of 3,469 males. We contrast this with effects on academic track attendance (graduates plus dropouts).

The coefficients' signs are positive and of similar magnitude for both specifications and variables. This documents our first finding. The tracking decision for our academic was highly persistent. The availability of an academic track at age 10 is highly predictive of attaining the academic track certificate, demonstrating that the educational expansion leveraged the academic potential of the population. The estimates indicate that having the corresponding school type upon completing elementary school significantly increases the probability of graduating from the track with the final degree by 8.25 or, respectively, 6.73 percentage points (relative to an overall mean of 32 percent). The results, however, vary in significance, revealing a weaker first stage for the second specification (2), which exploits only variation in the instrument across municipalities caused by school openings after 1940. The first stage is significant in both specifications. Yet, the conventional rule of thumb suggested by Staiger and Stock (1997), an F statistic exceeding 10, is not fulfilled in specification (2). New econometric insights suggest, however, that this rule of thumb is unnecessarily high in just-identified settings with one endogenous regressor and one instrument, as in our case. For instance, a recent study by Angrist and Kolesár (2024) concludes that conventional IV estimates and t-tests are compromised little by a weak instrument in the just-identified case as long as the endogeneity is not extraordinarily high.[10] We report the continuous instrument's similar first stage and 2SLS results and separate first-stage F-Statistics for all experience levels in the Appendix for comparison and robustness checks.[11]

Besides being relevant, the instrument needs to be unrelated to unobserved factors that correlate with education, conditional on the controls. To test for any observable predetermined differences between individuals with and without Abitur left after conditioning on the control variables of our two specifications, we provide balancing tests in Panel B of Table 2. Therefore, we regress different individual characteristics that should have been determined before but are potentially related to track choice on our instrument and

[10] They show that with a non-substantial degree of endogeneity (roughly defined by a correlation between the first stage and structural residual) $|\rho|$ not exceeding 0.565, the rejection rate of a conventional 5% t-test for a significant IV estimate is below 5% regardless of the population first stage F-statistic. As the estimated endogeneity parameters related to the first stage results of Panel A in Table 2 are far below this threshold for both specifications, we neglect further worries about weak instruments at this point. Following the approach of Angrist and Kolesár (2024), the respective endogeneity estimates $\hat{\rho}$ that allow for heteroskedastic or cluster-dependent errors are -0.026 and -0.007.

[11] Panel A of Table A.2 shows that a one-unit higher index, which indicates that academic track schools are relatively close, increases the probability of attaining Abitur by 8 to 11 percentage points. Again, standard errors are higher for specification (2), while the results are still significant at the 10%-level. Both F-statistics are lower than in Table 2, indicating lower first stages for the continuous than for the binary instrument. The first-stage estimates further vary with the experience level since the panel is not balanced over experience. Figure A.1 in the appendix shows that the F-statistic of the first specification with the binary instrument is above 10 for most experience levels. Again, this does not hold for the second specification or the continuous instrument.

Table 2: Regression results for first stage estimation and balancing checks

	Abitur		Academic track attendance	
	(1)	(2)	(1)	(2)
***Panel A:* First Stage**				
Acad. track in municipality	0.0828***	0.0673**	0.1104***	0.0787***
	(0.0241)	(0.0263)	(0.0247)	(0.0272)
F-statistic (instrument)	11.85	6.57	19.98	8.41
Observations	3,469		3,466	

***Panel B:* Balancing checks**	Explanatory variable: academic track in municipality			
	(1)		(2)	
	Coef	S.E.	Coef	S.E.
Dependent variables:				
Repeated year in primary school	−0.0022	(0.033)	−0.0041	(0.0037)
Firstborn	−0.0151	(0.0242)	−0.0222	(0.0266)
No. of older siblings	0.1530	(0.1400)	0.1662	(0.1619)
No. siblings	−0.0432	(0.0936)	0.0083	(0.1165)
Father's years of education	0.3233***	(0.1061)	0.0937	(0.1186)
Father born in Germany	−0.0209	(0.0145)	−0.0156	(0.0159)
Raised by single parent	0.0060	(0.0132)	−0.0067	(0.0142)
Raised by patchwork family	0.0016	(0.0097)	−0.0069	(0.0114)
Pre-existing acad. track dummy	no		yes	

Notes: Own calculations based on NEPS-SC6-ADIAB data. Panel A shows results from first stage estimations on two different dummy variables. "Abitur" indicates the hold of a university-entrance diploma, and "Academic track attendance" denotes whether an individual has ever attended an academic track school. In Panel B, the coefficients of the dummy variable "Academic track in municipality" are shown for different outcomes. Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. Standard errors in parentheses are clustered on the municipality level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

all covariates.[12] We see no significant relation to the instrument for all but one variable for both specifications. The significantly positive coefficient related to the fathers' years of education for specification (1) is likely to be driven by time-constant differences left between the municipalities in the same district that had academic track schools from early on and those that did not, as the coefficient turns insignificant for the specification

[12] Note that the two variables regarding the family constellation are referred to the age of 15. Still, any differences in the family situation between ages 10 to 15 are not likely related to access to academic track schools.

(2). Limiting the exploited variation across municipalities by including the dummy for a pre-existing academic track school in the birth municipality, thus, seems to be crucial for the credibility of the instrument's exogeneity. Based on these results, we believe that, most likely, the exogeneity assumption holds at least for our second specification. Nonetheless, we compare both specifications and show that neither a weaker first stage nor the risk of an endogenous instrument driven by the father's education alone drives our results.

Overall OLS and 2SLS

Next, we report OLS estimates of Abitur on average earnings within the first ten years of labor market experience in Panel A of Table 3 as a benchmark. The coefficients for the two specifications are nearly the same in magnitude. Both are highly significant, while the standard errors are slightly smaller in (1). Individuals with Abitur have, on average, about 1,312€ higher monthly labor earnings after conditioning on covariates. Compared to a baseline of 2,791€ on average for people without Abitur, this results in a relative difference of 47%. But, due to the potential endogeneity, these OLS results are not causally interpretable. We report the 2SLS estimates in Panel B for the causal effect for the subgroup of compliers. The coefficients are, again, very similar, and standard errors are more pronounced for specification (2). While the result is significant at the 1%-level for (1), specification (2) results are still significant at the 10%-level. In summary, over the total years of experience considered here, i.e., the first ten years, having Abitur increases the monthly gross labor earnings on average by around 1,900 €. Compared to people without Abitur, the higher degree increases earnings by around 74%.

Table A.3 in the Appendix reveals for both specifications that having Abitur goes along with a higher age when individuals enter the labor market, a higher probability of having a university degree, and more completed years of education for the subpopulation of compliers. The considerable effect of 74%, therefore, includes the effect of being older (and potentially better at bargaining wages) and the additional effect of having a university degree. The latter is included per definition because Abitur, a university entrance diploma, is the precondition for acceptance. Typically, the Abitur is associated with three additional years of schooling. Surprisingly, it increases the total years of education, including vocational and tertiary education, by only around 3.5 to 4.2 years for our compliers. Thus, the annualized average monetary returns to academic track education range between 14% and 17% per additional year of education ($1.74^{\frac{1}{4.2}}/1.74^{\frac{1}{3.5}}$). Panel B of A.2 in the Appendix shows that the IV results are robust to using the continuous instrument.[13] As standard in the literature on returns to education, our IV coefficients of both instruments are larger than our OLS estimates. However, this depends on the instrument and, accordingly, on the subgroup of compliers.

[13] Besides the larger standard errors, the 2SLS estimates are similar to those based on the binary instrument.

Table 3: Regression results for OLS and IV estimations

	Dependent variable: Monthly labor earnings [∅ in first 10 years]			
	(1)		(2)	
	Coef	S.E.	Coef	S.E.
***Panel A:* OLS**				
Abitur	$1,312.85^{***}$	(60.88)	$1,1311.10^{***}$	(70.00)
***Panel B:* 2SLS**				
Abitur	$1,927.52^{**}$	(779.80)	$1,857.39^{*}$	(1017.53)
Pre-existing acad. track dummy	no		yes	
Baseline w/o Abitur	2,791.06		2,791.06	
Observations	3,469		3,469	

Notes: Own calculations based on NEPS-SC6-ADIAB data. Panel A shows OLS estimates and Panel B IV estimates of Abitur on monthly labor earnings with the "Academic track in municipality" dummy as the instrument. Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. Standard errors in parentheses are clustered at the municipality level. Baseline: Average monthly earnings for individuals without Abitur. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

Our general IV results align with previous studies finding positive effects of high secondary schooling. Our LATE results are similar to those from Carneiro et al. (2017) for upper secondary schooling in Indonesia. While the total effect of upper secondary schooling is slightly higher, the annualized returns are slightly lower than the effect we find for Germany. Matthewes and Ventura (2022) find negative effects of attending vocational education compared to academic education in England. These results also point in the same direction as ours but are much smaller in magnitude. For Scotland, Clark and Del Bono (2016) even find no effects of attending an elite school on male wages. One reason for smaller or no effects might be the big difference in the educational system, with no or small differences in years of schooling when attending the academic or elite track. However, Dustmann et al. (2017) also do not find any effects of initially attending a more advanced track for marginal students on wages in Germany, suggesting a substantial difference between attending and completing the track with the final degree. At least in Germany, this can be driven by up- and downgrading between tracks, which is the main argument of Dustmann et al. (2017) for their zero effect on wages.[14] As we do not find this

[14] Cygan-Rehm and Westphal (2024) confirm negligible lifetime earnings returns for males. However, they document small but statistically significant effects on academic track graduation. They explain small but significant effects on lifetime earnings for females by deferred fertility in career phases with steep earnings trajectories.

up or downgrading matters quantitatively in our setting, our compliers are academically not at the margin between two tracks (instead, they may be at the availability margin).

The Evolution of LATEs

To go a step further, we also report the evolution of the LATEs on earnings over labor market experience in Figure 3.[15] As before, the results of the two specifications only differ slightly, with the LATEs of specification (2) mainly being marginally below those of specification (1). Also, the standard errors for specification (2) are higher throughout all experience levels, as shown by the larger confidence intervals. Shortly after entering, Abitur's earnings premium is small and statistically insignificant. A potential explanation is that the group without Abitur includes more individuals with vocational education within a firm. After completing their vocational training, they often officially enter the labor market with more practical experience than those with Abitur. At the beginning of the working careers, the advantage of practical experience might offset the signaling effect of higher education. Generally, the returns to education tend to increase over experience, even though the LATE estimate is not statistically significant for every experience level (also depending on the considered specification). Comparing the effects to the average labor market earnings for males without Abitur displayed in Figure 2, we get relative effects between -1 and 111%. In the Appendix, we report similar effects when using the continuous instrument (Figure A.3).

The findings on monetary returns over experience are mainly comparable to results on the evolution of returns for different ages. Bhuller et al. (2017) use Norwegian data and find that additional years of schooling increase earnings as soon as individuals reach the age of 25. This is most likely when higher-educated individuals (for example, people with academic-track schooling) enter the labor market. After that point, monetary returns increase sharply, at least until the age of 45. This matches our generally increasing returns to Abitur in the first 10 years of labor market experience.

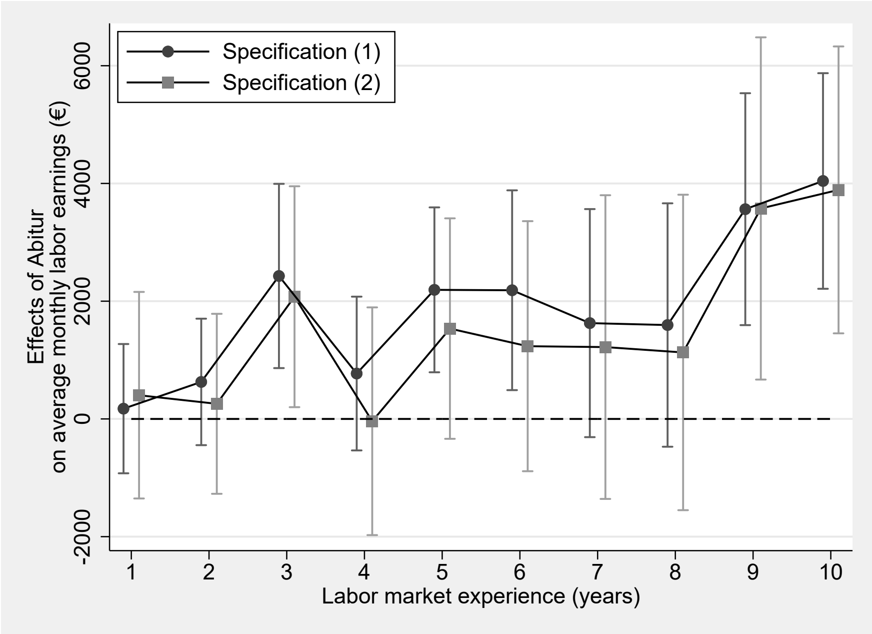
4 The Evolution of Unobserved Heterogeneity

4.1 MTE Framework

We now want to explore how unobserved heterogeneity in the Abitur earnings premium evolves for individuals who are indifferent between Abitur and the next best track by esti-

[15] OLS estimates are only included as a benchmark in Figure A.2 in the Appendix and depicted by the dashed lines, whereas the solid lines represent the experience-specific LATE estimates. For some experience levels the OLS estimates are slightly larger, for others the OLS estimates are clearly smaller than the IV estimates.

Figure 3: Local average treatment effects of Abitur on earnings over experience



Notes: Own illustration based on NEPS-SC6-ADIAB data. The graph reports regression results for IV estimations (with the "Academic track in municipality" dummy as the instrument) of Abitur on monthly labor earnings over labor market experience. Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. The vertical bars denote the 90% confidence intervals based on standard errors clustered at the municipality level.

mating marginal treatment effects (MTEs, see Heckman and Vytlacil, 2005). We therefore introduce a potential outcome framework, where we model these potential outcomes as

$$Y_{it}^{j} = \eta_t^j + X_i'\mu_t^j + U_{it}^j, \quad j \in \{0,1\}, \tag{2}$$

where X_i is a vector of demeaned (time-invariant) controls, such that η^j gives the mean earnings in the treatment state j (with 1 and 0 referring to individuals with and without Abitur, respectively) at time t. Equation 2 implies that the individual treatment effect is given by $Y_{it}^1 - Y_{it}^0 = (\eta_t^1 - \eta_t^0) + X_i'(\mu_t^1 - \mu_t^1) + (U_{it}^1 - U_{it}^0)$ and consists of observable (X_i) and unobservable factors (U_{it}^1, U_{it}^0). Both can potentially drive heterogeneity in the returns.

To generally allow for a correlation between the individual returns with the Abitur propensity underlying D_i, we set up a Roy model, which derives U_D, a value on the unit interval, which defines a hypothetical value, at which the individual is indifferent about Abitur. As forward-looking individuals should base their education decision on lifetime earnings, we use $\widetilde{Y}_i^j = \sum_{t=0}^{T} \frac{Y_{it}^j}{1+r}$ to denote discounted lifetime earnings (with interest rate r) accruing across the labor market career from 0 to experience level T. With these considerations, we can be more explicit about the underlying tracking decision. We start by relating the lifetime Abitur benefit $\widetilde{Y}_i^1 - \widetilde{Y}_i^0$ to potential costs $C(X, Z, U^C)$, which we assume to be additively separable in the instrument Z, the covariates X, and the unobserved costs U^C:

$$\begin{aligned} D_i &= \mathbb{1}\left\{(\widetilde{Y}_i^1 - \widetilde{Y}_i^0) \geq C(X_i, Z_i, U_i^C)\right\} \\ &= \mathbb{1}\left\{\pi_0 + \pi_1 Z_i + X_i'\gamma \geq V_i\right\} \\ &= \mathbb{1}\left\{F_V(\pi_0 + \pi_1 Z_i + X_i'\gamma) \geq F_V(V_i)\right\} \\ &= \mathbb{1}\left\{P(Z_i, X_i) \geq U_{D,i}\right\} \end{aligned} \tag{3}$$

Using the definition of lifetime earnings and the individual treatment effect from Equation 2, the second step separates all the observable components on the left- and all the unobserved components ($V_i := U_i^C - (\widetilde{U}_i^1 - \widetilde{U}_i^0)$) on the right-hand side of the inequality. Note that coefficients equal (up to scale depending on the variance of V_i) those from the first stage if both are estimated by a probit model. To clarify this, we apply the cumulative distribution function on V_i (a rank-preserving monotonic transformation) to both sides. The last step defines the notation of the resulting quantities. First, the propensity score $P(Z_i, X_i)$ maps the observables to the unit interval by assigning a probability to receive Abitur based on Z_i and X_i. Second, the unobserved value $U_{D,i}$ – being the quantiles of V_i – can be interpreted as the unobserved resistance to Abitur. Note that from now on, we suppress the i on this index. Thus, the MTE is now defined as $E(Y_i^1 - Y_i^0 | U_D = u_D)$, i.e., the causal effect of D at different quantiles u_D of V_i. For every higher value of the observables induced by Z_i (holding X_i constant), more people select an academic track education. Individuals who change D due to a marginal change in Z_i are indifferent, and $P(Z_i, X_i)$ equals U_D. For these additional marginal individuals being at the u_D-th quantile of the distribution ($P(Z_i, X_i) = u_D$), we can evaluate the marginal treatment effect at any observed value of the propensity score or, equivalently, any quantile U_D. This is generally not restricted to stronger assumptions than those required for the LATE estimation, as shown by Vytlacil (2002).

4.2 MTE Estimation

For estimating the MTE, we primarily need the conditional expectation $E[Y_i | X_i = x, P(Z_i, X_i) = p]$. Its derivative with respect to the propensity score results in the MTE:

$$MTE(X_i = x, U_{Di} = p) = \frac{\partial E[Y_{it} | X_i = x, P(Z_i, X_i) = p]}{\partial p} \tag{4}$$

It is the effect of an increase in the propensity score on the outcome in experience year t. An estimable expression of the conditional expectation $E[Y_{it}|X_i = x, P(Z_i, X_i) = p]$ is given by

$$E[Y_{it}|X_i = x, P(Z_i, X_i) = p] = \eta_0 + X_i\mu_0 + (\eta_1 - \eta_0)p + X_i(\mu_1 - \mu_0)p + K(p) \quad (5)$$

as shown by Heckman and Vytlacil (2005), $K(p)$ is a function of the propensity score and is generally not further specified. Hence, we must first estimate the propensity score to estimate Equation 5.[16]

We estimate a linear MTE, i.e. $K(p) = \lambda p^2$. This is relatively inflexible but condenses the heterogeneity into one parameter – the coefficient for the slope. Although the degree of heterogeneity is overestimated due to extrapolation, the change of the linear MTE slopes over time can still report an increase or decrease in heterogeneity. However, a corresponding parametric model of order two can approximate a nonparametric fit well using the specification test of Härdle and Mammen (1993). The two specifications' corresponding p-values of 0.25 and 0.46 exceed 0.1, suggesting that the parametric model can be used here. In the preferred MTE specification, we estimate the propensity score using the continuous instrument. However, we show comparable results based on the binary instrument in the Appendix. Moreover, we only use observations within common support, i.e., the overlap of $P(Z)$ for $D = 1$ and $D = 0$. We report standard errors clustered at the municipality level based on bootstrapping using 200 repetitions.

4.3 Results II: Marginal Treatment Effects

Detecting a Pattern in Unobserved Heterogeneity: Selection into Gains

Table 4 includes first-stage results and MTE results for the first (Panel A) and last (Panel B) observed experience year, i.e., average effects for experience levels from the first and tenth year. Columns 1 and 2 display the average marginal effects of the instrument on the probability of having Abitur from probit estimations of the selection equation in Equation 3 for each specification. For both points in time, the coefficient is slightly higher in (1), while standard errors are smaller than for (2). As expected, the effects are significantly

[16] To forego further assumptions, we follow the approach of Kamhöfer et al. (2019) and estimate MTEs that only vary over the unobservables. Hence, we restrict μ_1 to equal μ_0 in Equation 5 , such that

$$E[Y_{it}|X_i = x, P(Z_i, X_i) = p] = \eta_0 + X_i\mu_0 + (\eta_1 - \eta_0)p + K(p). \quad (6)$$

Thus, we restrict the covariates to have the same effect in both potential outcome equations. This might seem a harsh restriction, but it is equivalent to common IV approaches (such as LATE) where the treatment indicator is typically not interacted with other covariates. Even if the true effects vary over X, the restriction only affects the intercept of the MTE as the derivative of Equation 5 concerning the p is constant in X. Since we are mainly interested in the heterogeneity driven by differences in unobservables for the whole population, any of the restrictions' shortcomings are irrelevant.

positive for the continuous instrument in both specifications and experience years since a higher index indicates that academic track schools are relatively close.[17] The table also shows the χ^2 test statistics on the significance of the excluded instrument, documenting significant results at the 5% level for both specifications in Panel A and B. We also use the binary instrument (also used for estimating the LATE) to estimate the linear MTE as a robustness check. The corresponding χ^2-statistics of the first stage selection models over experience are shown in the Appendix (Figure A.4). The propensity scores of the first stage estimation provide a sizeable common support for both specifications ranging between 0.03 and 0.88 (Figure A.5). We estimate the MTE as described above for observations within the common support of the particular specification.

Table 4: Regression results for selection equations and linear MTEs for experience year 1 and 10

	First Stage		**MTE**	
	Abitur		Monthly labor earnings	
	(1)	(2)	(1)	(2)
Panel A: Experience year 1				
Index (Z^I)	0.1470***	0.1034**		
	(0.0449)	(0.0487)		
χ^2-statistic (instrument)	10.69	4.54		
Intercept			−37.28	45.19
			(918.58)	(965.91)
Slope			−523.04	−289.92
			(1,441.78)	(1,308.12)
Panel B: Experience year 10				
Index (Z^I)	0.1555***	0.1214***		
	(0.0422)	(0.0461)		
χ^2-statistic (instrument)	13.62	6.92		
Intercept			7,574.66***	6,346.94**
			(2,320.81)	(2,771.07)
Slope			−10,855.16**	−8,563.76*
			(4,415.6)	(4,781.66)
Pre-existing academic track dummy	no	yes	no	yes

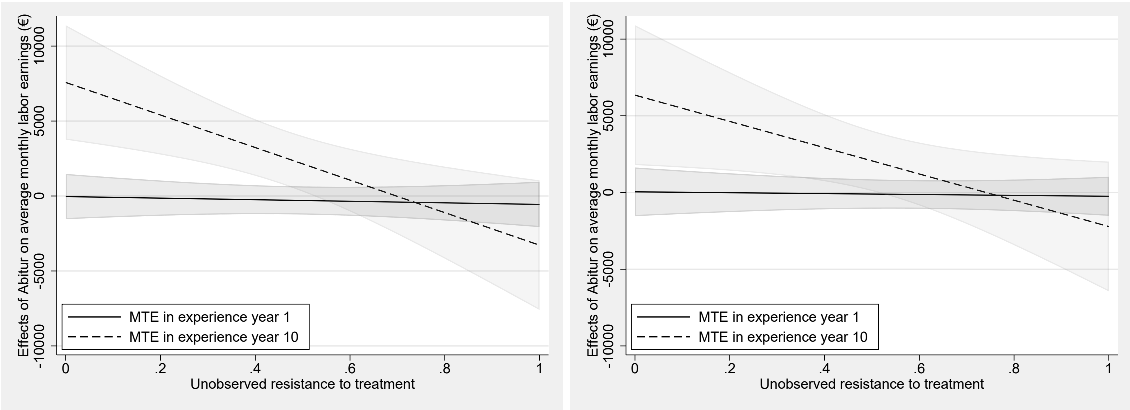
Notes: Own calculations based on NEPS-SC6-ADIAB data. Columns 1 and 2 report average marginal effects from a probit selection model, and columns 3 and 4 show parameters of linear MTEs for the first and 10th year of labor market experience with the continuous variable "index (Z^I)" as the instrument. Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. Standard errors in parentheses are clustered at the municipality level and in columns 3 and 4, bootstrapped with 200 repetitions. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

[17] Any deviations in the first stage results are attributed to differences in the composition of individuals observed in the distinct experience years.

Estimating the outcome equation in Equation 5 with $K(p) = \lambda p^2$ and taking the derivative with respect to the propensity score provides the parameters for the linear MTE. Columns 3 and 4 in Table 4 show the relevant parameters of the linear MTEs in experience year 1 (Panel A) and 10 (Panel B). $(\hat{\eta}_1 - \hat{\eta}_0)$ is the intercept, while $2\hat{\lambda}$ specifies the slope parameter of the linear MTE. The latter indicates the degree of heterogeneity. The estimated intercept and slope parameters of specifications (1) and (2) are generally similar in both panels. Visible, instead, is the difference in the results between both panels. We see no significant intercept or slope parameters in the first year in the labor market. Thus, we cannot document any evidence for significant returns to Abitur for any individual along U_D, i.e., the resistance to academic track education, when people first enter the labor market. This, however, looks considerably different for experience year 10 in Panel B. Here, we have significantly positive intercepts of the MTE in both specifications and, most importantly, clear and significant negative slopes. For this time in the individuals' careers, we find clear evidence for heterogeneity in unobservables for the returns to Abitur. The negative slope and its relevant magnitude further detect a pattern of selection into gains: Individuals with the highest returns to Abitur are the first to attend the newly opened academic track schools. This suggests that individuals (or their parents and teachers) choose their educational path based on expected gains.

The marginal effects along U_D in Figure 4 show the results of Table 4 graphically. The difference between the highest and lowest return on an MTE curve corresponds to the estimated slope parameter. Hence, a steeper slope parameter of an MTE indicates greater unobserved heterogeneity in the returns. The flat solid lines in (a) and (b) correspond to the MTE curves of experience year 1. They visualize the lack of heterogeneity in the returns already shown by the insignificant slope parameters. 90% confidence intervals further show that returns are insignificant for any specification at every U_D. The dashed lines indicate the MTE curves for experience year 10. As expected from the parameters in Panel B of Table 4, they are very similar for both specifications and sharply decreasing. For individuals with low quantiles of U_D, i.e., with low resistance to treatment, we find the highest monetary returns to Abitur. For people with low resistance to academic track education, already low values of $P(Z)$ exceed the U_D, making them very likely to choose academic track education. For those the returns are significantly positive irrespective of the specification. We document no significant returns to academic track education for the upper half of the U_D distribution. Interestingly, the returns even turn negative for high levels of U_D. Nonetheless, returns derived from the linear MTE at the outer margins of the resistance distribution potentially rely heavily on extrapolation. Besides the lack of significance at this point, they can only be interpreted with caution. Comparing the results with those from a binary instrument shows similar linear MTEs in every subfigure, i.e., in experience years 1 and 10 for both specifications and instruments (Figure A.6).

Figure 4: Comparison of linear MTE results in year 1 and 10



(a) Specification (1) (b) Specification (2)

Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure presents the regression results of linear MTE estimations of Abitur on monthly labor earnings for the first and 10th year of labor market experience with the continuous variable "index (Z^I)" as the instrument. Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. The shaded areas denote 90% confidence intervals based on standard errors clustered at the municipality level, which are bootstrapped with 200 repetitions.

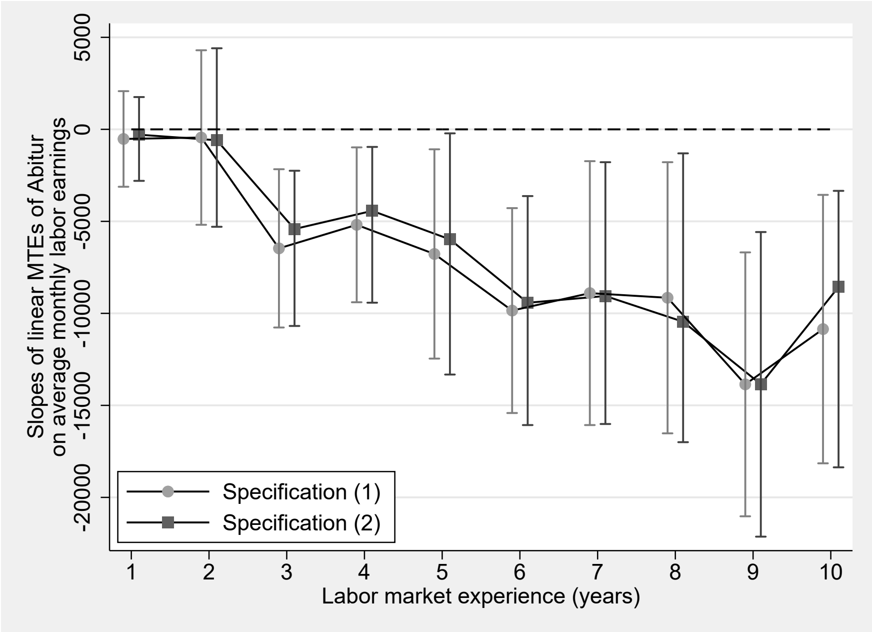
Our linear MTEs match the MTE results from Carneiro et al. (2017) for Indonesia, where returns are also negative for the higher end of the U_D distribution. MTE results on the returns to college education in Germany from Kamhöfer et al. (2019) also align with our finding of a negative slope, suggesting that theoretically compelling selection into gains for monetary returns to education is a stylized fact also empirically.

The Evolution of Selection into Gains

Figure 5 shows the slope parameters of the linear MTEs for every year of working experience. The slopes are nearly zero in the year of entering the labor market and the following year. The 90% confidence intervals show that the heterogeneity in unobservables is also not statistically significant at that time. However, from experience year two onwards, we observe significantly negative slopes, revealing a clear selection pattern into gains. The generally decreasing slope parameters also detect increasing heterogeneity in monetary returns regarding resistance to treatment. The longer the individuals are in the labor market, the greater the differences in returns between individuals from different quantiles of U_D. This interpretation is similar when using the binary instrument (A.8 in the Appendix). Figure A.7 in the Appendix illustrates these differences by adding the level – it shows the effect trajectory for individuals with the lowest and highest resistance to Abitur.[18] Noticeably, with the inserting heterogeneity in year 2, the returns for the individuals with high resistance to academic track education become negative. Note that U_Ds near 0 or

[18]Note that the slope parameter is $2\hat{\gamma}$ and since U_D and p are bounded within the 0/1-interval, the marginal individual with the highest potential resistance to treatment have a propensity score of 1. Accordingly, the returns for the highest resistance to treatment are given by $(\hat{\eta}_1 - \hat{\eta}_0) + 2\hat{\lambda} \cdot 1$. The slope parameters in Figure 5 determine the distance between the returns for low- and high-resistance individuals in Figure A.7.

Figure 5: Slope parameters of linear MTEs over experience



Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure presents the estimated slope parameters of linear MTE estimations of Abitur on monthly labor earnings over labor market experience with the continuous variable "index (Z^I)" as the instrument. Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. 90% confidence intervals are based on bootstrapped standard errors with 200 repetitions clustered at the municipality level.

1 (outside the common support) belong to hypothetical individuals that might not exist. Thus, we interpret these results cautiously, even though negative monetary returns to education for a part of the distribution have also been found in the previous literature (Carneiro et al., 2011, 2017; Nybom, 2017) and we also find significantly negative returns at, for instance, $U_D = 0.8$, i.e., inside the common support.

4.4 Interpreting the Evolution of Unobserved Heterogeneity in Light of Employer Learning

The evolution of MTEs over time can additionally give insights into wage development. For simplicity, we assume that labor markets are competitive and that all employers can access the same information about employees. Appendix B discusses deviations from these assumptions. We document that U_D negatively correlates with earnings returns starting two years after labor market entry. An expected result if individuals base their education decision on expected gains.[19] Initially, it is fair to assume that these expected gains, captured by U_D, are private information, mainly unknown to employers. Yet, over time, employees reveal their true productivity, and employers notice, learn, and may act on

[19] Such a pattern of effect heterogeneity is a well-established finding also in other settings (see, for example, Carneiro et al., 2011, Carneiro et al., 2017 or Kamhöfer et al., 2019).

it. If wages reflect productivity (as in perfectly competitive labor markets), the correlation between U_D and the returns to education increase as a by-product of this learning. Put differently: employers learn about time-invariant productivity factors (which correlate with U_D) over time and adopt wages according to the expected productivity given all previous signals on job performance (Ablay and Lange, 2023).[20] Hence, the information asymmetry on productivity decreases over time, and the education signal per se as an easily observable signal becomes less determining for wages relative to idiosyncratic aspects (captured by U_D). Consequently, the MTE slopes must become steeper (i.e., more negative) with more labor market experience, which we also documented.[21]

Because general labor market experience may also directly explain increasing effect heterogeneity, more negative MTE slopes over labor market experience are insufficient to prove employer learning. To distinguish between the direct effects of experience and employer learning more formally, we focus on the unobserved component that affects the potential outcome in state j, which we model as $U^j_{it} = \theta^j_{it} A^j_{it}(U_D, t) + \omega^j_{it}$. The endowment A^j_{it} is unobserved to us as researchers, but individuals may know about it. Importantly, because they (or their teachers and parents) can anticipate A^j_{it} at the time of the tracking decision, A^j_{it} correlates with U_D – individuals have a low U_D partially because of a high A^1_{it} relative to A^0_{it}. Additionally, these skills may form over time, for instance, through on-the-job training or learning by doing (the direct effects of experience). Thus, we model A^j_{it} as a function of U_D and t. However, these skills may only translate into higher wages if the employer knows about them. Hence, employer learning may be reflected in the parameters θ_{it}, which are implicitly a function of U_D and t as well. Over time t, individuals reveal more about their U_D. The term ω^j_{it} contains everything uncorrelated to A^j_{it} (like unanticipated or transitory wage shocks). Additionally, we assume it to be uncorrelated with U_D for simplicity. Therefore, focusing on the first part is instructive for employer learning and how MTEs may reveal it. Here, it is important to distinguish between individual endowment (the variable A^j_{it}) and how this endowment is valued in the labor market (through the coefficient θ^j_{it}). Both are multiplicatively linked. Using the definition of U^j_{it}, the unobserved heterogeneity underlying the MTE is captured by:

$$E(\Delta U_{it} | U_D = u_D) = \Delta\theta_{it} A^1_{it} + \theta^0_{it} \Delta A_{it},$$

[20] The standard employer learning model introduced by Farber and Gibbons (1996) and further developed by Altonji and Pierret (2001) assumes competitive labor markets with asymmetric information on productivity between employers and employees. Additionally, all employers can access the same information (symmetric employer learning).

[21] Even if the relevance of employer learning differs with the educational degree, as found by Arcidiacono et al. (2010) for college versus high school graduates, increasing essential heterogeneity over experience in the returns still indicates employer learning for at least one of the groups.

where we use Δ to denote differences in potential outcomes.[22] Note that $\Delta\omega_{it}$ disappears, as we assume it does not correlate with U_D and follow the convention that the overall mean of U_{it}^j is zero.

The derivative of $E(\Delta U_{it}|U_D = u_D)$ with respect to t demonstrates how time in the labor market forms the unobserved heterogeneity and is key to gaining insight into employer learning. It reads:

$$\frac{\partial E(\Delta U_{it}|U_D = u_D)}{\partial t} = \underbrace{\frac{\partial \Delta A_{it}}{\partial t}\theta_{it}^0 + \frac{\partial A_{it}^1}{\partial t}\Delta\theta_{it}}_{\text{Productivity growth}} + \underbrace{\frac{\partial \Delta\theta_{it}}{\partial t}A_{it}^1 + \frac{\partial \theta_{it}^0}{\partial t}\Delta A_{it}}_{\text{Employer learning}} \tag{7}$$

Hence, the partial derivatives of θ_{it}^j and A_{it}^j determine the two mechanisms that drive the evaluation of the MTEs along t:

(1) Productivity growth (determined by derivatives of A_{it}^j): Over time t, individuals may acquire more skills A_{it} and the effect of D_i on A_{it} may grow.

(2) Employer learning (determined by derivatives of θ_{it}^j): Time t in the labor market reveals private information to employers, which is important for wage setting. This is reflected by changes in $\Delta\theta_{it}$ and θ_{it}^0 over time.

The important task is distinguishing between the two mechanisms outlined above. We want to isolate (2). Hence, we have to fix (1). The learning model introduced by Farber and Gibbons (1996), which has become the standard in the empirical employer learning literature, assumes that productivity growth does not depend on unobserved ability. Except for Aryal et al. (2022), who explicitly allow for varying returns to skill with experience, studies on employer learning rely on this assumption. Following the majority, we could attribute changing MTE slopes only to (2).

Assessing slope changes of MTEs across experience levels can, in principle, also identify employer learning without this assumption. Equation (7) demonstrates that decreasing monetary returns for individuals with a higher U_D must be driven by employer learning under two relatively mild assumptions. First, we need to assume a non-negative productivity growth channel (1): irrespective of U_D, productivity can only grow with experience (at least at the beginning of the career where we assess the returns). Second, Abitur is neither supposed to deteriorate skills (i.e., we assume $\Delta A_{it} \geq 0$) or be uninformative to employers (assuming $\Delta\theta_{it} \geq 0$). This implies that the first part of the equation, i.e., the productivity growth component, cannot be negative – even for individuals with the highest U_D ($= \overline{u_D}$).

This differs for the employer learning component. The employer will reward and increase wages for employees on which she updates her beliefs that the individual is of the high-productive type at the expense of individuals she believes are of the low-productive type.

[22] We get there by adding and subtracting $\theta_{it}^0 A_{it}^1$ before rewriting the term in differences.

The employer learning effect alone must be negative for individuals with a high U_D. If we detect declining returns relative to labor market entry for these individuals, the employer learning effect must overcompensate the productivity growth channel. Hence, if the effect for $E(\Delta Y_{it}|U_D = \overline{u_D}, t)$ decreases along t, employer learning must (at least partly) drive increasing heterogeneity detected by the MTE slopes over time.

If we put any concerns regarding the linear MTE results for people with very high resistance to academic track education aside for a moment, this is what we observe. Figure A.9 in the Appendix shows that for both specifications, the returns of individuals with high resistance to treatment have insignificant zero returns when entering the labor market that generally decrease to (partly) significant and negative returns later in their careers. Interpreted in this light, employers seem to learn relatively fast (particularly since productivity growth only allows the detection of a lower bound). After two years of labor market experience, there is a clear and sharp decrease in the slope parameter of the MTE. It decreases further but at a much smaller rate. Our results indicating fast learning match the previous literature analyzing the speed of employer learning (Lange, 2007; Aryal et al., 2022).

5 Conclusion

We estimate monetary returns to academic track education by exploiting variations in the availability of those schools induced by the educational expansion in West Germany using a high-quality data set linking representative survey data to administrative labor market data (NEPS-ADIAB). To identify the returns, we rely on academic track school openings that substantially changed geographical access to higher secondary schooling to instrument the choice of academic track education. We analyze how returns evolve with increasing labor market experience and uncover heterogeneities in the effects along the resistance to education – an important dimension of unobserved effect heterogeneity. We are first to analyze the evolution of unobserved heterogeneity in returns to education after individuals enter the labor market and provide an interpretation in light of employer learning.

We find average returns to the highest German schooling degree (Abitur) of over 70% within the first 10 years after labor market entry, corresponding to returns of 14% to 17% per additional year of education. These positive returns first appear after two years of working experience. Lacking positive effects directly after labor market entry reveals a weak signaling value of this schooling degree. We also document substantial heterogeneity in the returns to Abitur along unobserved distaste to academic track education from year two onwards. Clear selection into gains after ten years of experience suggests that individuals most likely to take academic track education benefit the most. However, this education may not pay off for everybody. With increasing experience, our results imply very low to negative returns for some individuals. Although average effects are positive for most experience levels, further incentives for academic track education are less likely to pay off, as they will attract people with lower returns.

The evolution of heterogeneity along characteristics unobservable to employers gives further insights into employers' wage setting. With growing experience, the returns for individuals with the highest and lowest resistances to academic track education get more heterogeneous. This is consistent with employers adjusting wages over time depending on productivity expectations. Employers might learn about innate abilities as a time-constant component of productivity. People with unobserved characteristics reducing resistance to academic track schooling might simultaneously have higher productivity returns over experience. We generally conclude that a combination of both most likely drives our findings on increasing heterogeneity. Yet, suppose we assume labor market experience to increase earnings for everyone and take our negative earnings results for high-resistance individuals at face value, our results even prove the occurrence of employer learning.

References

Ablay, M. and Lange, F. (2023). Approaches to Learn About Employer Learning. *Canadian Journal of Economics/Revue canadienne d'économique*, 56(2):343–356.

Altonji, J. G. and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Angrist, J. and Kolesár, M. (2024). One Instrument to Rule Them All: The Bias and Coverage of Just-ID IV. *Journal of Econometrics*, 240(2):105398.

Arcidiacono, P., Bayer, P., and Hizmo, A. (2010). Beyond Signaling and Human Capital: Education and the Revelation of Ability. *American Economic Journal: Applied Economics*, 2(4):76–104.

Aryal, G., Bhuller, M., and Lange, F. (2022). Signaling and Employer Learning with Instruments. *American Economic Review*, 112(5):1669–1702.

Bachbauer, N., Wolf, C., Graf, T., Grießemer, S., Kaimer, S., Köhler, M., Lehnert, C., Oertel, M., Schneider, A., and NEPS-Netzwerk (LIfBi) (2022). National Educational Panel Study (NEPS), Starting Cohort 6 (SC6) Survey Data Linked to Administrative Data of the IAB (NEPS-SC6-ADIAB) – Version 7520 v1. Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB).

Bachbauer, N. and Wolf, C. B. (2022). NEPS-SC6 Survey Data Linked to Administrative Data of the IAB (NEPS-SC6-ADIAB). FDZ-Datenreport, 01/2022 (en), Nuremberg.

Becker, R. (2006). Dauerhafte Bildungsungleichheiten als Unerwartete Folge der Bildungsexpansion? In Hadjar, A. and Becker, R., editors, *Die Bildungsexpansion: Erwartete und unerwartete Folgen*, pages 27–61. VS Verlag für Sozialwissenschaften, Wiesbaden.

Bhuller, M., Mogstad, M., and Salvanes, K. G. (2017). Life-Cycle Earnings, Education Premiums, and Internal Rates of Return. *Journal of Labor Economics*, 35(4):993–1030.

Birkelund, J. F. and van de Werfhorst, H. G. (2022). Long-term Labor Market Returns to Upper Secondary School Track Choice: Leveraging Idiosyncratic Variation in Peers' Choices. *Social Science Research*, 102:102629.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69(5):1127–1160.

Carneiro, P., Heckman, J. J., and Vytlacil, E. (2011). Estimating Marginal Returns to Education. *American Economic Review*, 101(6):2754–2781.

Carneiro, P., Lokshin, M., and Umapathi, N. (2017). Average and Marginal Returns to Upper Secondary Schooling in Indonesia. *Journal of Applied Econometrics*, 32(1):16–36.

Clark, D. and Del Bono, E. (2016). The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom. *American Economic Journal: Applied Economics*, 8(1):150–176.

Cygan-Rehm, K. and Westphal, M. (2024). School Starting Age and the Gender Pay Gap over the Life Cycle. *Ruhr Economic Papers*, 1115.

Dahrendorf, R. (1965). *Bildung ist Bürgerrecht*. Wegener, Hamburg.

Dauth, W. and Eppelsheimer, J. (2020). Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide. *Journal for Labour Market Research*, 54(1):1–14.

Deming, D. J. (2022). Four Facts about Human Capital. *Journal of Economic Perspectives*, 36(3):75–102.

Dustmann, C., Puhani, P. A., and Schönberg, U. (2017). The Long–Term Effects of Early Track Choice. *The Economic Journal*, 127(603):1348–1380.

Farber, H. S. and Gibbons, R. (1996). Learning and Wage Dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047.

Franzmann, G. (2006). The Development of the Educational System in Germany: Comprehensive Schools 1960–2000. GESIS Data Archive, Cologne. ZA8220 Data file Version 1.0.0.

Ge, S., Moro, A., and Zhu, B. (2021). Testing for Asymmetric Employer Learning and Statistical Discrimination. *Applied Economics*, 53(12):1361–1377.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Heckman, J. J. and Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation1. *Econometrica*, 73(3):669–738.

Henz, U. and Maas, I. (1995). Chancengleichheit durch die Bildungsexpansion. In Sahner, H. and Schwendtner, S., editors, *27. Kongreß der Deutschen Gesellschaft für Soziologie - Gesellschaften im Umbruch: Sektionen und Arbeitsgruppen*, pages 389–393. Westdt. Verl., Opladen.

Härdle, W. and Mammen, E. (1993). Comparing Nonparametric Versus Parametric Regression Fits. *The Annals of Statistics*, 21(4):1926–1947.

Kahn, L. B. (2013). Asymmetric Information between Employers. *American Economic Journal: Applied Economics*, 5(4):165–205.

Kamhöfer, D. A., Schmitz, H., and Westphal, M. (2019). Heterogeneity in Marginal Non-Monetary Returns to Higher Education. *Journal of the European Economic Association*, 17(1):205–244.

Koebe, J. and Marcus, J. (2020). The Impact of the Length of Schooling on the Timing of Family Formation. DIW Berlin Discussion Paper No. 1896.

Koebe, J. and Marcus, J. (2022). The Length of Schooling and the Timing of Family Formation. *CESifo Economic Studies*, 68(1):1–45.

Köhler, H. and Lundgreen, P. (2014). *Allgemeinbildende Schulen in der Bundesrepublik Deutschland: 1949-2010: ZA8570 Datenfile Version 1.0.0*. GESIS, Köln.

Lange, F. (2007). The Speed of Employer Learning. *Journal of Labor Economics*, 25(1):1–35.

Matthewes, S. H. and Ventura, G. (2022). On Track to Success? CEPA Discussion Papers (58).

Nybom, M. (2017). The Distribution of Lifetime Earnings Returns to College. *Journal of Labor Economics*, 35(4):903–952.

Picht, G. (1964). *Die deutsche Bildungskatastrophe: Analyse und Dokumentation*. Walter.

Schönberg, U. (2007). Testing for Asymmetric Employer Learning. *Journal of Labor Economics*, 25(4):651–691.

Staiger, D. and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557–586. Publisher: Wiley, Econometric Society.

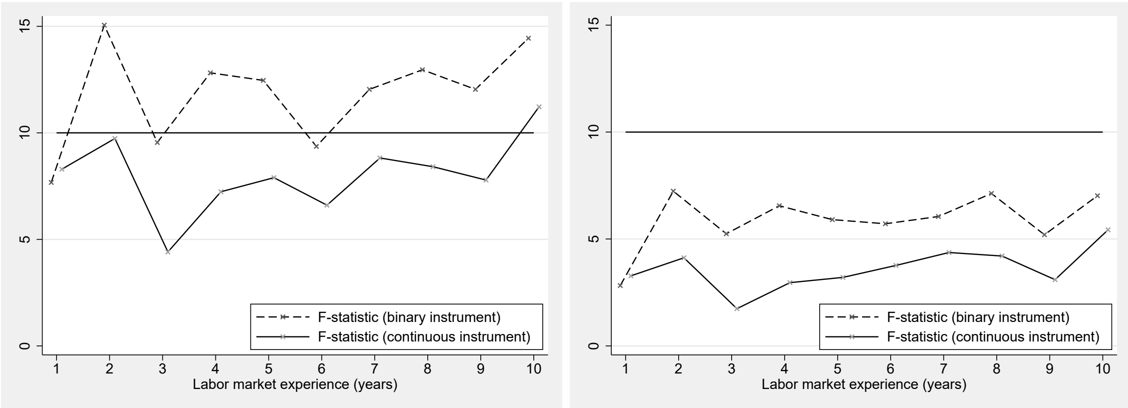
Vytlacil, E. (2002). Independence, Monotonicity, and Latent Index Models: An Equivalence Result. *Econometrica*, 70(1):331–341.

Westphal, M., Kamhöfer, D. A., and Schmitz, H. (2022). Marginal College Wage Premiums Under Selection Into Employment. *The Economic Journal*, 132(646):2231–2272.

Appendix

A Additional tables and figures

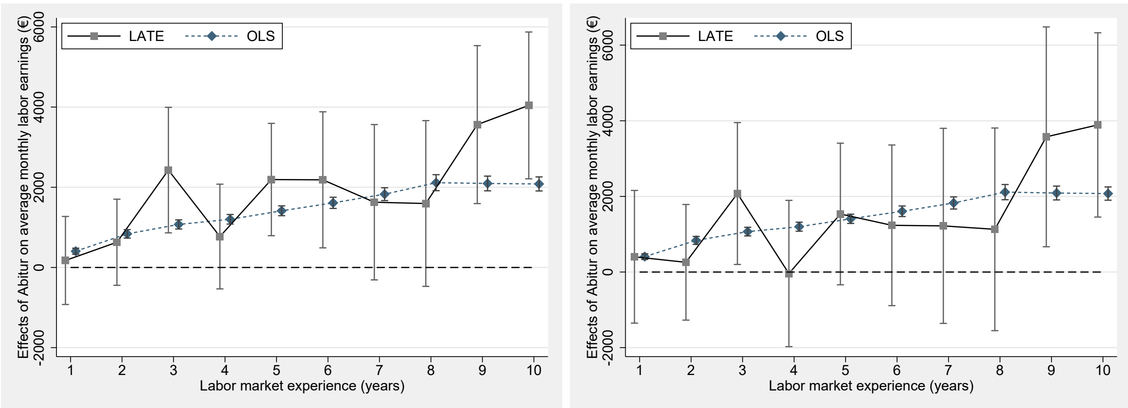
Figure A.1: First stage F-statistic over experience



Specification (1) Specification (2)

Notes: Own illustrations based on NEPS-SC6-ADIAB data. The figure reports the F-statistic from a linear first-stage regression of the instruments on the treatment (Abitur) over labor market experience. The binary instrument is the "Academic track in the municipality" dummy; the continuous instrument is "index (Z^I)". Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940.

Figure A.2: Local average treatment effects compared to OLS estimates over experience



Specification (1) Specification (2)

Notes: Own illustration based on NEPS-SC6-ADIAB data. The graph reports regression results for OLS and IV estimations of Abitur on monthly labor earnings over labor market experience with the "Academic track in the municipality" dummy as the instrument. Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. The vertical bars denote the 90% confidence interval based on standard errors clustered at the municipality level.

Table A.1: Variables and means by academic track degree

	Definition	Without Abitur	With Abitur
Academic track attendance	=1 if ever attended an academic track school	0.139	0.882
Secondary school entry	Year individual entered secondary school	1975	1978
Month of birth	Month the individual was born	6	6
Pre-existing academic track dummy	=1 if born in municipality with academic track school before 1940	0.579	0.687
Repeated year in primary school	=1 if individual repeated at least one year in primary school	0.010	.
Years of education father	Fathers years of education	12.081	13.889
Father born in Germany	=1 if father was born in Germany	0.923	0.916
Raised by single parent	=1 if raised by a single parent (from birth to age 15)	0.065	0.044
Raised by patchwork family	=1 if raised in a patchwork family (from birth to age 15)	0.052	0.021
Firstborn	=1 if individual was the firstborn child in the family	0.296	0.342
Nr older siblings	Number of older siblings	1.430	1.178
Nr siblings	Number of siblings	1.995	1.458
Age start working	Age (years) when individual first entered the labor market (without vocational training)	20.798	24.396
University degree	=1 if individual has a degree from a university	0.018	0.489
Years of education		13.103	16.535
Number of observations		2344	1125

Notes: Own calculations based on NEPS-SC6-ADIAB data. The table shows means of the variables for individuals with and without the academic track degree (Abitur). Values might be missing due to data protection rules.

Table A.2: Regression results for First Stage and IV estimations with continuous instrument

	Dependent variable:			
	Abitur		Monthly labor earnings [∅ in first 10 years]	
	(1)	(2)	(1)	(2)
Panel A: First Stage				
Index (Z^I)	0.1084***	0.0816*		
	(0.0387)	(0.0419)		
Panel B: 2SLS				
Abitur			$1,904.43^{**}$	$1,814.87$
			(894.28)	(1258.56)
Pre-existing acad. track dummy	no	yes	no	yes
F-statistic (instrument)	7.84	3.80		
Baseline w/o Abitur			2,791.06	
Observations	3,469		3,469	

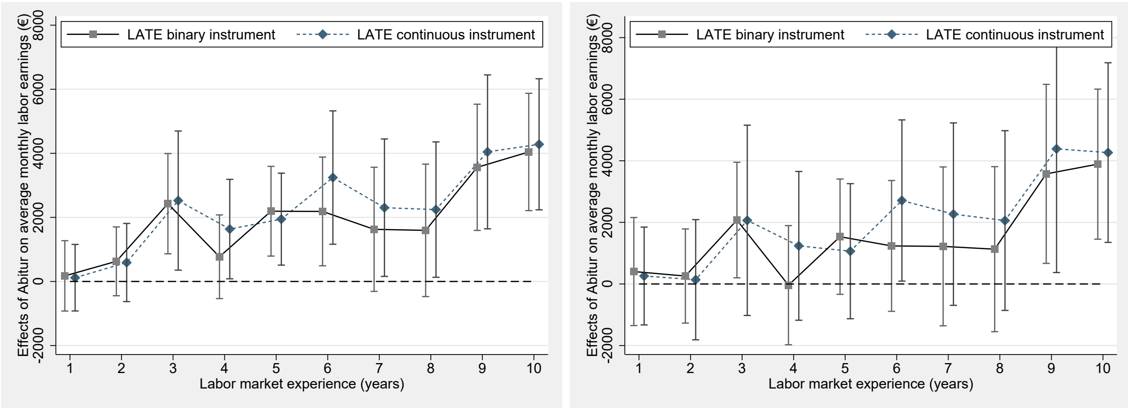
Notes: Own calculations based on NEPS-SC6-ADIAB data. Panel A shows first stage estimates and Panel B IV estimates of Abitur on monthly labor earnings with the continuous variable "index (Z^I)" as the instrument. Specification (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. Standard errors in parentheses are clustered at the municipality level. Baseline: Average monthly earnings for individuals without Abitur. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

Table A.3: Regression results for IV estimations with other dependent variables

	Dependent variable:					
	Age start working		University degree		Years of education	
	(1)	(2)	(1)	(2)	(1)	(2)
Abitur	5.80***	6.51**	0.55***	0.46*	4.17***	3.53***
	(2.01)	(2.71)	(0.18)	(0.24)	(0.99)	(1.32)
Long-establ. acad. track dummy	no	yes	no	yes	no	yes
Baseline outcome w/o Abitur	20.80		0.02		13.10	
N	3,469	3,469	3,467	3,463	3,446	3,446

Notes: Own calculations based on NEPS-SC6-ADIAB data. The graph reports IV estimations of Abitur on three different outcomes for two specifications with the "Academic track in the municipality" dummy as the instrument. Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. Standard errors in parentheses are clustered on the municipality level. Baseline: Average outcome for individuals without Abitur. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

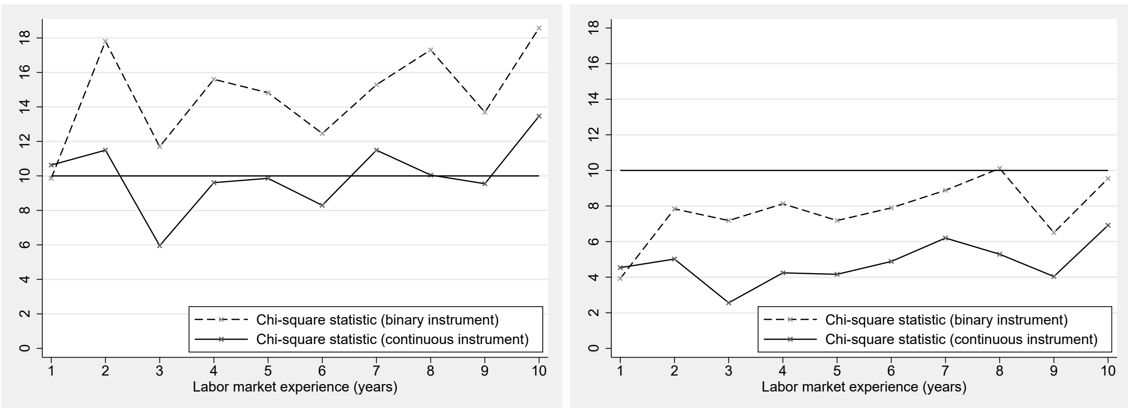
Figure A.3: Comparison of local average treatment effects over experience for different instruments



Specification (1) Specification (2)

Notes: Own illustrations based on NEPS-SC6-ADIAB data. The graphs report regression results for local average treatment effects of Abitur on monthly labor earnings over labor market experience for different instruments. The binary instrument is the "Academic track in the municipality" dummy; the continuous instrument is "index (Z^1)". Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. The vertical bars denote the he vertical bars denote the 90% confidence intervals based on standard errors clustered at the municipality level.

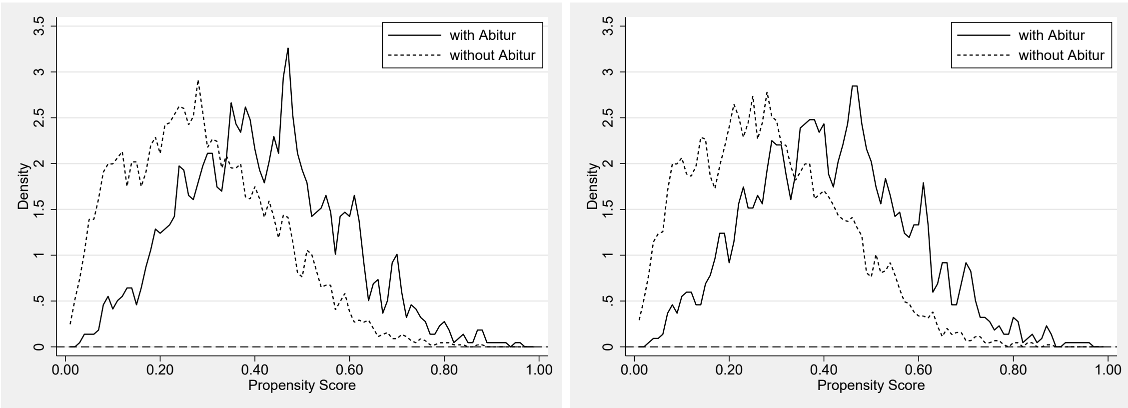
Figure A.4: Chi-squared statistics of a probit selection model over experience



Specification (1) Specification (2)

Notes: Own illustrations based on NEPS-SC6-ADIAB data. The figure reports the chi-squared statistics from a probit regression of the instruments on the treatment (Abitur) over labor market experience. The binary instrument is the "Academic track in the municipality" dummy; the continuous instrument is "index (Z^1)". Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940.

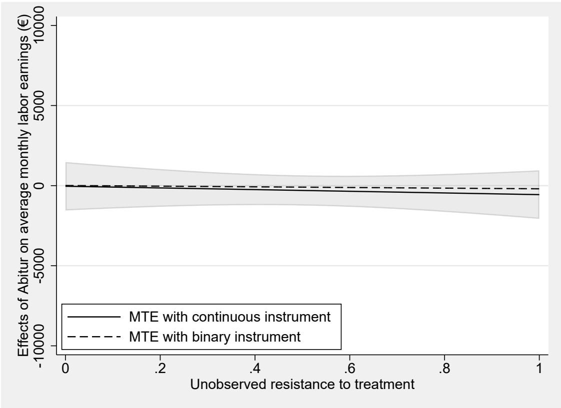
Figure A.5: Support by treatment status



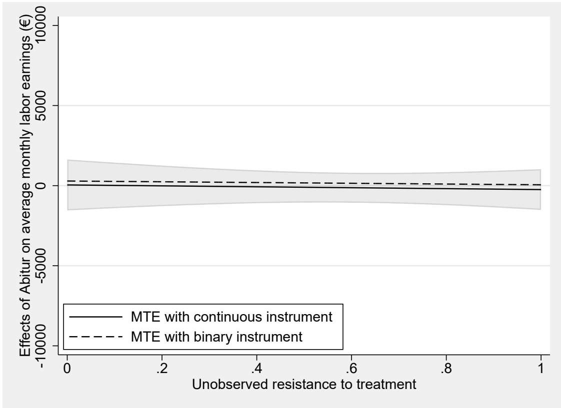
Specification (1) Specification (2)

Notes: Own illustration based on NEPS-SC6-ADIAB data. This graphs depict for each specification the estimated density of the propensity score separately for individuals with and without Abitur. Propensity scores are predicted from probit estimations of the selection equation in Equation 3 with all individuals (in a cross-section). Specifications (1) and (2) include district and entry cohort fixed-effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940.

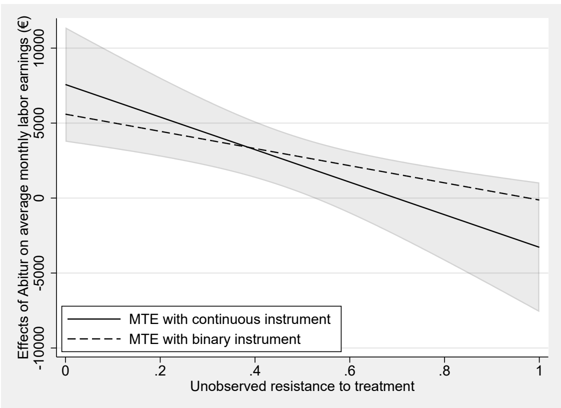
Figure A.6: Comparison of linear MTEs results for experience year 1 and 10 with different instruments



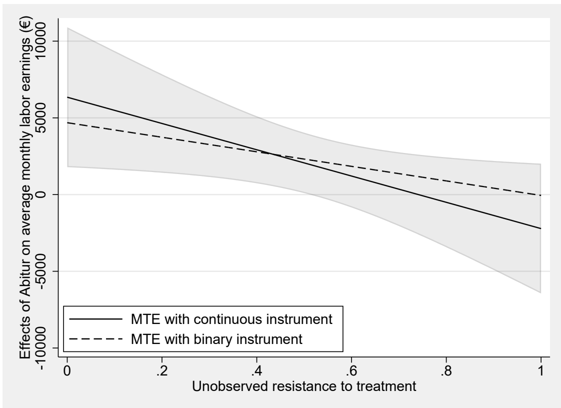
Year 1 – Specification (1)



Year 1 – Specification (2)



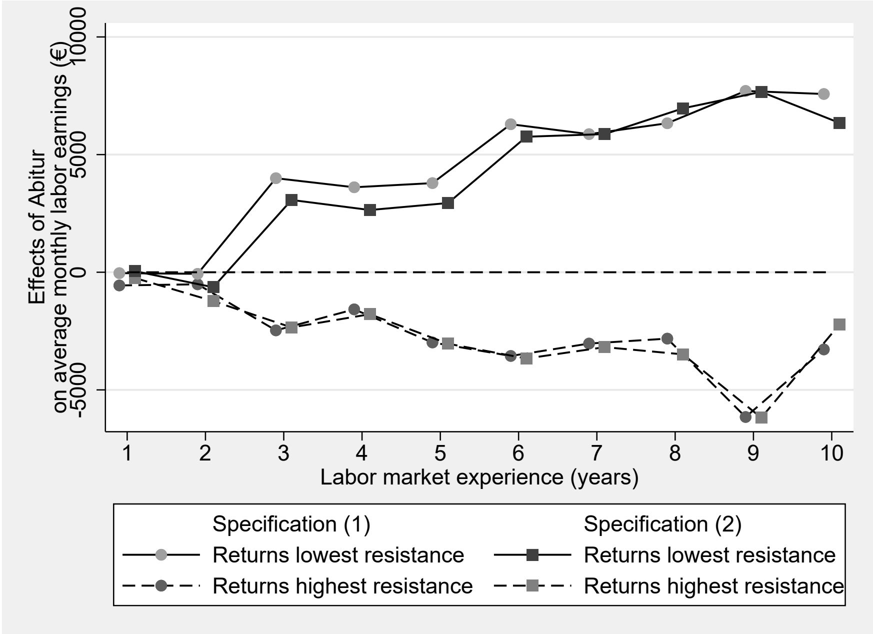
Year 10 – Specification (1)



Year 20 – Specification (2)

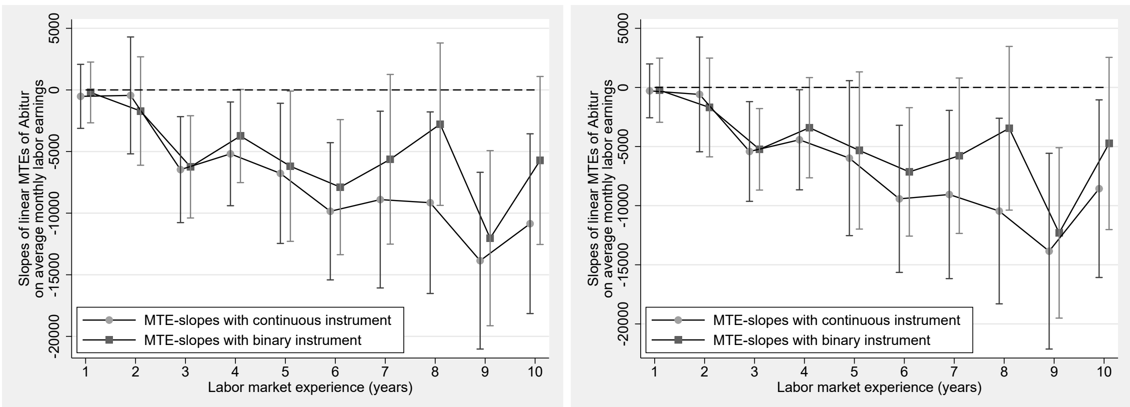
Notes: Own illustration based on NEPS-SC6-ADIAB data. The figures compare the regression results of linear MTE estimations of Abitur on monthly labor earnings for two different instruments. The binary instrument is the "Academic track in the municipality" dummy; the continuous instrument is "index (Z^I)". There is one subfigure for every combination of experience year (1 and 10) and specification ((1) and (2)). Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. The shaded areas denote 90% confidence intervals based on standard errors clustered at the municipality level, which are bootstrapped with 200 repetitions.

Figure A.7: Effects of Abitur on earnings over experience by resistance to treatment



Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure shows monetary returns to academic track education for individuals with different resistances to treatment (from MTE estimation). Individuals with the highest resistance to academic track education have a propensity score of 0, and those with the lowest have a propensity score of 1. Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940.

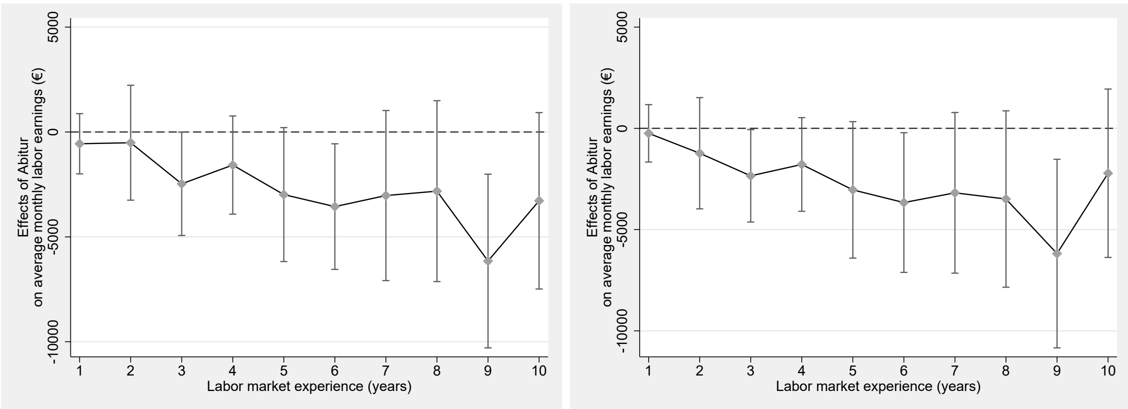
Figure A.8: Comparison of slope parameters of linear MTEs over working experience with different instruments



Specification (1) Specification (2)

Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure presents the estimated slope parameters of linear MTE estimations of Abitur on monthly labor earnings over labor market experience for different instruments. The binary instrument is the "Academic track in the municipality" dummy; the continuous instrument is "index (Z^I)". Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. 90% confidence intervals are based on bootstrapped standard errors with 200 repetitions clustered at the municipality level.

Figure A.9: Effects of Abitur on earnings over experience for high resistance to treatment



Specification (1) Specification (2)

Notes: Own illustration based on NEPS-SC6-ADIAB data. The figure shows monetary returns to academic track education for individuals with high resistance to treatment, i.e., with a propensity score of 1. Specifications (1) and (2) include district and entry cohort fixed effects, birth month dummies, and state-specific trends. Specification (2) additionally includes a dummy indicating if there was any academic track school in the birth municipality before 1940. 90% confidence intervals are based on bootstrapped standard errors with 200 repetitions clustered at the municipality level.

B Implications of Restricted Wage Adjustment, Asymmetric Learning, and Imperfectly Competitive Labor Markets

Restricted Wage Adjustment and Asymmetric Learning

This employer-learning interpretation is based on the standard or symmetric employer-learning model, which means that every employer has the same information and sets wages equal to expected productivity at any time. This rules out any long-term contracts and collective payment agreements. In Germany, most employees have unlimited contracts, typically including a payment scheme negotiated ex-ante. A collective bargaining system also regulates wages according to pay scales bargained, for example, by trade unions for whole industries. The assignment of employees to these pay scales is typically not directly based on any productivity measures. Taken together, this substantially hampers setting wages equal to expected productivity. However, productivity can still affect promotions into higher positions, which allows for wage adjustments based on productivity. Observing increasing heterogeneity in monetary returns along innate abilities is already a clear sign that there are any productivity-related payment mechanisms. Thus, the German circumstances do not rule out employer learning in a broader sense and do not change the interpretation of our results.

Moreover, learning is assumed to be symmetric, meaning all employers have the same information on employees' productivity. It might be reasonable to assume that current employers have better information than future employers. There has partly been empirical support for asymmetric employer learning, especially for higher-educated employees (for example, in Schönberg, 2007; Kahn, 2013; Ge et al., 2021). However, job references or information about previous employers, tasks, and wages still allow information to flow across employers. Thus, asymmetric learning does not rule out common learning, but without the possibility of job changes, symmetric employer learning could be even faster. An employer-learning interpretation is still valid for asymmetric learning as long as employers and employees make no strategic decisions based on asymmetric learning.

Imperfectly Competitive Labor Markets

Last, the standard model assumes competitive labor markets. If, contrarily, the employers have any market power, wages might fall below marginal productivity. As long as potential mark-downs are proportionally the same for all employees, the employer-learning interpretation would not change. Differences in mark-downs might result from varying elasticities of labor supply. Aryal et al. (2022) argue that education supports specialization and the supply of specialized labor might be less elastic. However, potential differences in mark-downs driven by education do not restrict the general interpretation of our findings. If the deviation between wages and productivity is smaller for individuals with than without Abitur, but this deviation is constant over time, the increase in heterogeneity still results in the same interpretation. Indeed, this is not true when interpreting the overall extent of heterogeneity.

CESA Working Paper

Series

CESA Working Paper No. 1 | 2023
Andreas Hauptmann, Benjamin Schwanebeck and Hans-Jörg Schmerer
Plant-level adjustments to imports and exports at the extensive margin

CESA Working Paper No. 2 | 2023
Joscha Beckmann, Marco Kerkemeier and Robinson Kruse-Becher
Regime-specific exchange rate predictability

CESA Working Paper No. 3 | 2023
Joscha Beckmann, Timo Heinrich and Jennifer Rogmann
Inflation expectations and cognitive uncertainty

CESA Working Paper No. 4 | 2024
Daniel Monsees and Matthias Westphal
Disruptions in Primary Care: Can Resigning GPs Cause Persistently Negative Health Effects?

CESA Working Paper No. 5 | 2024
Giulio Callegaro, Mario Lackner and Hendrik Sonnabend
The Napoleon complex revisited: New evidence from professional soccer

CESA Working Paper No. 6 | 2024
Mario Lackner and Hendrik Sonnabend
When performance melts away: Heat causes mental errors in high-stakes competitions

CESA Working Paper No. 7 | 2024
Johannes Hollenbach, Hendrik Schmitz and Matthias Westphal
Gene-environment interactions with essential heterogeneity

CESA Working Paper No. 8 | 2024
Kamila Cygan-Rehm and Matthias Westphal
School Starting Age and the Gender Pay Gap over the Life Cycle

CESA
Working
Paper
Series

CESA Working Paper No. 9 | 2024
Marco Kerkemeier
(Co-)Explosiveness of corporate credit spreads

CESA Working Paper No. 10 | 2024
Christoph Wegener, Robinson Kruse-Becher and Tony Klein
EU ETS Market Expectations and Rational Bubbles

CESA Working Paper No. 11 | 2025
Anna Krumme and Matthias Westphal
Monetary returns to upper secondary schooling, the evolution of unobserved heterogeneity, and implications for employer learning

FernUniversität in Hagen