RUHR
ECONOMIC PAPERS

# Is Academic-Track Schooling Worthwhile Without College? Decomposing Monetary Returns to Education

Anna Krumme, and Matthias Westphal

technische universität dortmund
rwi
#1197

# Imprint

**Ruhr Economic Papers #1197**

Responsible Editor: Ronald Bachmann
RWI – Leibniz-Institut für Wirtschaftsforschung e.V.
Hohenzollernstraße 1–3 | 45128 Essen, Germany
Fon: +49 201 8149-0 | email: rwi@rwi-essen.de
www.rwi-essen.de

The Institute has the legal form of a registered association;
Vereinsregister, Amtsgericht Essen VR 1784

**Published by**

RWI – Leibniz-Institut für Wirtschaftsforschung e.V.
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Bergische Universität Wuppertal, Schumpeter School of Business and Economics
Gaußstraße 20, 42119 Wuppertal



RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.


ISSN 1864-4872 (online)
ISBN 978-3-96973-382-0
DOI https://dx.doi.org/10.4419/96973382

# Is Academic-Track Schooling Worthwhile Without College? Decomposing Monetary Returns to Education

Anna Krumme
FernUniversität in Hagen
TU Dortmund

Matthias Westphal
FernUniversität in Hagen
RWI Essen

January 2026

**Abstract**

We estimate monetary wage returns to academic-track education, Germany's elite secondary school type granting university entrance. Because academic-track attendance and subsequent university education are institutionally linked, we disentangle their contributions using a causal mediation analysis. Leveraging quasi-experimental variation from the educational expansion – independent openings of schools and universities – we identify (i) the direct effect of academic-track education holding university attendance constant and (ii) the indirect effect operating through university education. We find total monetary returns of 118%, with about 60 percentage points attributable to the indirect effect of additional university education with prior academic-track schooling, and the remaining 40 points to academic-track education alone.

**Keywords:** Returns to education, IV estimation, causal mediation analysis
**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 Ronald Bachmann and the participants in the internal RWI health research seminar for their helpful comments and discussions.

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Adults, 10.5157/NEPS:SC6:12.1.0. From 2008 to 2013, NEPS data were collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is conducted by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in collaboration with a nationwide network.

# 1 Introduction

Education is one of the key determinants of individual earnings and productivity, yet the interaction between different levels of education in shaping labor market outcomes remains poorly understood (Deming, 2022). While a vast literature has estimated the returns to secondary and tertiary education separately (see Gunderson and Oreopolous, 2020, and Patrinos and Psacharopoulos, 2020, for overviews), much less is known about how these education stages complement each other in generating wage returns. In Germany – as in all other tracked education systems – this question is particularly relevant for policy: the early tracking of students into different school types based on observed ability may set individuals on distinct educational and career trajectories that may be challenging to alter later in life (see Betts, 2011 for a general overview, and Dustmann et al., 2017 and Cygan-Rehm and Westphal, 2024 for evidence on Germany).

In this paper, we address these questions by (i) estimating the causal monetary returns to academic-track schooling – Germany's elite secondary track attended by about 35% of students – and (ii) disentangling the contribution of subsequent university education to these returns. While both margins have been analyzed in isolation in Dustmann et al. (2017), Krumme and Westphal (2024), Kamhöfer et al. (2019), and Westphal et al. (2022), investigating the interaction between both margins is particularly important because, traditionally, the academic-track degree, *Abitur*, has been the main prerequisite for university entry. Conceptually, this allows us to decompose the total treatment effect into two components: the causal effect of Abitur (our treatment) on earnings, holding university education (our mediator) constant (i.e., the direct effect of academic-track education), and the causal effect of university education, holding the academic-track degree constant (that is, the indirect effect through university education). To identify these returns, we build on Schmitz and Westphal (2025) and exploit two instruments – school and university openings – that jointly alleviate supply-side constraints in educational choices over the substantial societal change of the educational expansion.

To illustrate the intuition behind our approach, we begin by conditioning on the compliers of the school openings – i.e., a group for which we estimate the overall returns to academic-track schooling. Within this group, we compare earnings across two combinations of academic-track and university attainment: academic-track graduates with and without subsequent college education. Thereby, we explicitly allow for unobserved heterogeneity in the university education decision. Accounting for this margin of heterogeneity is crucial, as it represents a key confounder when comparing causal effects (see, e.g., Hollenbach et al., 2024) and may render subsequent university enrollment endogenous. We address this concern using the marginal treatment effect (MTE) framework (Heckman and Vytlacil, 2005), which enables us to compare earnings of academic-track compliers who are indifferent between enrolling in university and pursuing the next-best alternative (most likely vocational training). In this way, the causal mediation framework simultaneously addresses two sources of endogeneity – selection into both academic-track and university education.

We combine representative survey data with purpose-built datasets on academic-track school and university openings. The survey data from the National Educational Panel Study (NEPS) enables us to analyze individuals' educational and residential histories, along with their earnings information later in life. We merge self-collected information on all academic-track schools and colleges[1] in West Germany. We focus on males to minimize the interference of earnings with fertility-related labor supply decisions that are more prevalent for females (Westphal et al., 2022). Overall, we have a data set of 3,726 male individuals who lived in West Germany at the time of their educational choices.

Our paper contributes to the literature by bridging the gap between two stylized facts: upper secondary schooling and tertiary education have positive effects on individual earnings. We are the first to conduct a causal mediation analysis on the monetary returns to schooling, aiming to shed light on the interaction between secondary and tertiary education, thereby disentangling the additional effect of a college degree for individuals with an academic-track degree. Closely related questions on sequential educational choices have been examined by Zamarro (2010) for Spain – methodologically the closest counterpart to our work – and by Biewen and Thiele (2020) for Germany. Zamarro (2010) distinguishes between low, medium, and high educational attainment and develops a sequential multinomial framework within the marginal treatment effect (MTE) approach using two instruments. Biewen and Thiele (2020) analyze all possible transitions from secondary to tertiary education using a latent factor model that imposes structure on both the dimensionality and the functional form of unobserved heterogeneity. In contrast, our study leverages quasi-experimental variation and instrumental-variables strategies. In comparison to Zamarro (2010), who focuses on developing an econometric framework without presenting clearly interpretable wage effects, we adopt a causal mediation framework that explicitly decomposes the return to the academic-track. This approach allows us to avoid parametric restrictions on unobserved heterogeneity. The downside is twofold. First, we can only decompose the local effect for the population affected by school openings. Second, to keep the problem tractable, we need to focus on the transition from academic-track to university education and its next-best alternative, i.e., other higher educational options (*Fachhochschulabschluss*) or vocational training.

Our results show that, on average, compliers with academic-track availability experience substantial returns to academic-track education of approximately 118%. This effect partly reflects the returns to college education, as obtaining the Abitur increases the probability

[1]We use the words university and college as synonyms to refer to German *Universitäten* and closely related institutions like institutes of technology (*Technische Universitäten*/*Technische Hochschulen*), and universities of the armed forces (*Bundeswehruniversitäten*/*Bundeswehrhochschulen*).

of earning a college degree for this group by about 37 percentage points. We then examine how both components interact. The indirect effect is statistically significant at the 5% level and accounts for roughly 60 percentage points of the total return, indicating additional wage gains from college education among individuals with Abitur. The direct effect of Abitur on earnings is not statistically significant at conventional levels but remains economically meaningful, explaining the remaining 40 percentage points of total returns.

This paper is structured as follows: Section 2 provides an overview of the institutional background. Section 3 outlines the causal mediation approach and its assumptions by focusing on the non-technical intuition. In Section 4, the employed data and the empirical strategy for the baseline models are described. Section 5 presents the baseline effect of academic-track education on earnings and reports the main results on the mediation analysis. Finally, Section 6 concludes.

## 2 Institutional Background

**The (West-)German Secondary Schooling and Higher Educational System**

In the German education system, students are assigned to one of three secondary school tracks following four years of elementary education, typically at age 10.[2] This assignment is based on the student's (perceived) academic performance and is initially recommended by the elementary school teacher. While such recommendations are standard across all federal states, in most West German states, the final decision regarding track placement rests with the parents. The three secondary school tracks are commonly referred to as the basic track (*Hauptschule*), the intermediate track (*Realschule*), and the upper track or academic-track (*Gymnasium*).[3] The basic and intermediate tracks are primarily designed to prepare students for vocational training in blue- or white-collar professions and typically end after 5 or 6 years, respectively.[4] In contrast, the academic-track at a *Gymnasium* offers a more comprehensive curriculum to prepare students for tertiary education at universities or colleges. This track traditionally spans nine years and culminates in the *Abitur*, the highest secondary school leaving certificate and a prerequisite for university admission. This structure was in place for all West German students up to the graduating class of 2007.

[2]Section 2 draws on text from Krumme and Westphal (2024) and Kamhöfer et al. (2019).

[3]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).

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

Despite the early onset of tracking, the German system allows for flexibility as students may transfer between tracks at various stages of their education. Moreover, individuals who complete the basic or intermediate tracks may subsequently enroll in academic programs that enable them to attain the Abitur and thereby qualify for higher education. For more detailed information about the tracking system and the different school types, see Dustmann et al. (2017) and Krumme and Westphal (2024).

Upon completion of secondary education, adolescents in Germany typically pursue one of two pathways: enrollment in higher education or entry into the vocational training system through an apprenticeship. Apprenticeships follow a dual system that combines part-time, occupation-specific classroom instruction with part-time, on-the-job training. These programs generally span three years, after which participants frequently transition into full-time employment, either within the training firm or another company in the same sector. As mentioned, university enrollment requires the Abitur, which was awarded almost exclusively by academic-track schools with 13 years of schooling in the years under review (school graduates from 1960-1990).

Germany's higher education landscape is primarily composed of two types of institutions: traditional universities/colleges and universities of applied sciences (*Fachhochschulen*). Universities are typically large institutions that offer a broad range of academic disciplines and emphasize theoretical and research-oriented education. In contrast, universities of applied sciences are generally smaller, often specialize in particular fields (e.g., business or engineering), and offer practice-oriented curricula with pedagogical approaches more closely resembling those of secondary education. Higher education institutions in Germany generally do not impose tuition fees. However, students are responsible for their living expenses. In contrast, individuals in vocational training programs receive a modest monthly wage. These differences in direct and indirect costs – such as relocation expenses or opportunity costs associated with forgone earnings – may represent significant financial barriers. They can influence an individual's decision to either pursue higher education or vocational alternatives.

**Academic-Track School and College Openings during the Educational Expansion**

In the early 1950s, educational opportunities in the Federal Republic of Germany were minimal. For instance, in 1952, only 12.4% of students in grade eight were enrolled in the academic-track, and merely 3.8% of all school leavers graduated from this track (Köhler and Lundgreen, 2014). These low participation rates were primarily driven by constraints on the supply side rather than a lack of demand. In 1950, there were only 1,823 academic-track schools across the entire territory of West Germany (Franzmann, 2006). Over time, however, both public sentiment and political priorities shifted, with increasing recognition that educational access should not be restricted by economic or geographical barriers

(Becker, 2006). This led to a substantial change in the educational infrastructure – referred to as educational expansion – which we use in this study to assess the economic returns to education.

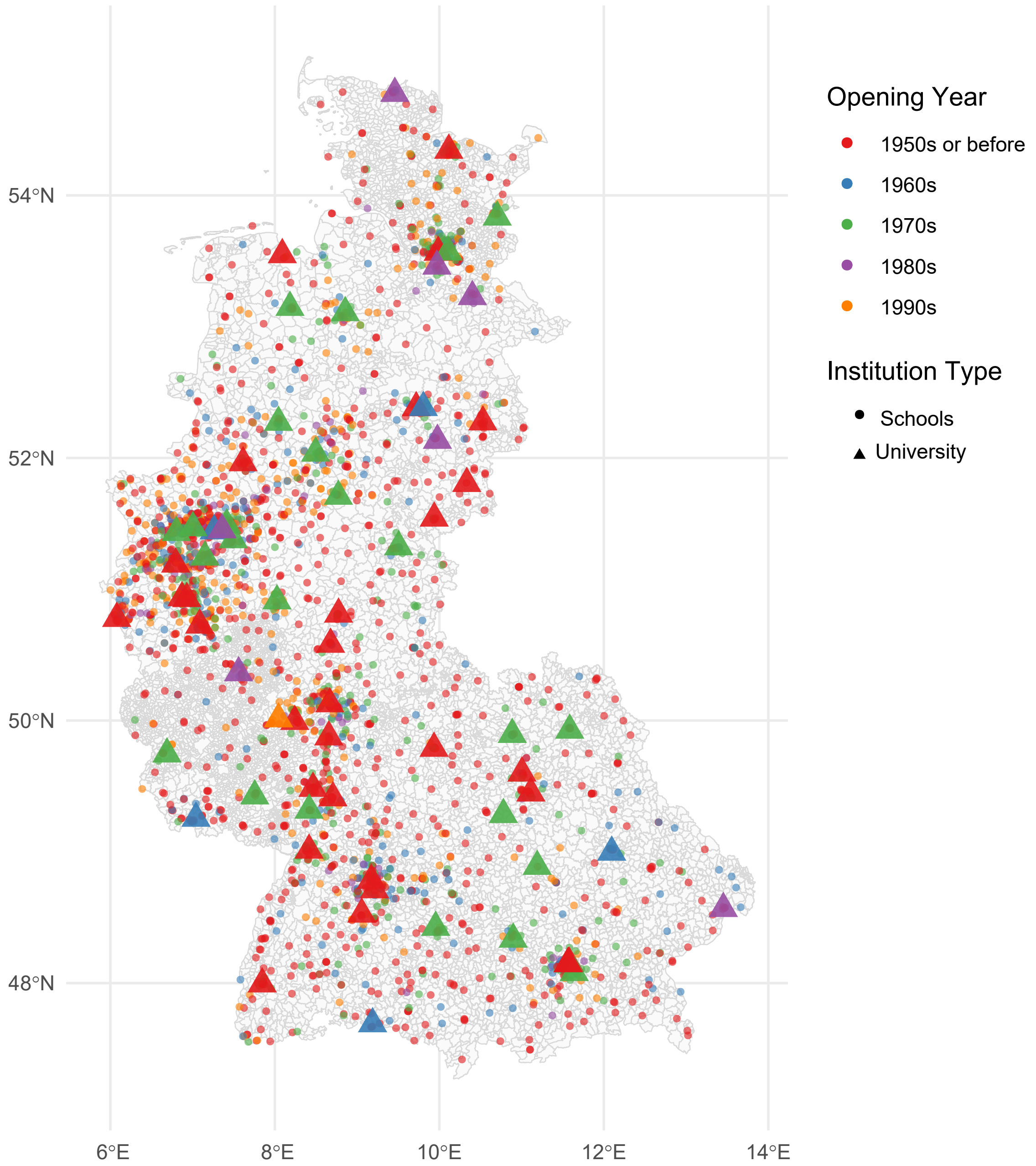
Both economic and sociopolitical considerations drove the educational expansion of the early 1960s. Picht (1964) identified an "education crisis," arguing that insufficient investment in higher education threatened the economic future and calling for substantial public funding to sustain growth. Concurrently, the discourse increasingly emphasized equal access to education; for instance, Dahrendorf (1965) advocated for "education as a civil right." As a result, a central objective of educational reform from the 1960s onward was to broaden educational access, with a particular emphasis on higher education.

Between 1960 and 1990, public expenditure on academic-track schools rose from 1,130 million to 11,559 million Deutschmarks. This substantial investment facilitated the establishment of 796 new academic-track schools, representing a significant increase relative to the 1,396 existing schools. Over the same period, student enrollment in these schools nearly doubled, growing from approximately 850,000 to 1.6 million (Franzmann, 2006).

The educational expansion also significantly improved access to tertiary education. Between 1958 and 1990, the number of colleges in Germany doubled from 33 to 66, reducing the average travel distance to the closest college by approximately 50 km in districts where a college was established. In addition to the founding of new institutions, existing colleges expanded significantly in capacity. The average number of students per college rose from 5,013 in 1958 to 15,438 in 1990. Among the original 33 institutions, 30 remained operational in 1990, with an average enrollment of 23,099 students. Overall, the total number of tertiary students increased from 155,000 in 1958 to approximately one million by 1990. We define extended college availability as both the establishment of new institutions and the expansion of existing ones. More information about the educational expansion in tertiary education, with a focus on the exogeneity of college openings, can be found in Kamhöfer et al. (2019).

Figure 1 illustrates the spatial distribution of academic-track school openings (dots) and universities (triangles) from the 1950s to the 1990s. It visualizes the growing accessibility of academic-track school and college education during the educational expansion, as evidenced by the number of newly opened institutions, which were present in all federal states across West Germany.

Figure 1: Spatial distribution of universities and academic-track schools by opening decade in Germany




*Notes:* Own illustration based on self-collected school and university information. The figure shows academic-track school and university openings across West Germany by decade.

# 3 Causal mediation analysis

## 3.1 Target parameters

Our target parameter – commonly referred to as the total treatment effect (TTE) in the causal mediation literature (see Huber, 2020 for an overview) – is the causal effect of academic-track education, denoted by the treatment $D$, on wages $Y$. In potential outcome notation, $Y^1$ represents a wage an individual would earn with academic-track education, while $Y^0$ represents the wage without it. By the structure of the German educational system, academic-track education directly affects university education, denoted by the mediator $M$. Accordingly, we define $M^1$ as the potential university education with academic-track education, and $M^0$ as the potential university education without it. With this notation, we can define the total treatment effect as

$$TTE = E(Y^1 - Y^0) = E(Y^{1M^1}) - E(Y^{0M^0}). \tag{1}$$

The term on the right-hand side introduces a second potential outcome dimension – potential university attendance $M^0$ and $M^1$ – and emphasizes that the total treatment effect does not condition on this dimension. Specifically, $Y^{1M^1}$ denotes the potential wage under academic-track education and the corresponding university enrollment that naturally follows from it, while $Y^{0M^0}$ represents the analogous natural potential wage without academic-track education. University enrollment for individuals with $M^1 = M^0$ is unaffected by academic-track education. Consequently, differential university education cannot directly explain individual returns to academic–track education. We say, $M$ has no mediating role for these individuals. However, if $M^1 \neq M^0$, the situation changes: a causal effect of $D$ on $M$ exists, which in turn may generate downstream effects on wages. To quantify the role of the mediator in the total treatment effect, we need to separate the potential treatment state from the potential mediator state. Adding and subtracting the never naturally observed quantities $E(Y^{dM^l})$ to Eq. (1), where $d \neq l \in \{0, 1\}$, we can hold the direct treatment effect fixed while allowing the mediator channel to vary. This allows us to rewrite the total treatment effect as a sum of a direct and an indirect effect of the treatment, holding the other component fixed at a particular state $TTE = DTE(1) + ITE(0)$ and $TTE = DTE(0) + ITE(1)$. Specifically, the indirect treatment effect fixes the direct treatment dimension in the potential outcome:

$$ITE(d) = E(Y^{dM^1} - Y^{dM^0})$$

It measures the causal effect of changing university education from $M^0$ to $M^1$ for individuals with an academic-track state of $d$. Note that for individuals with $M^1 = M^0$, this effect is zero. Likewise, the direct treatment effect is

$$DTE(l) = E(Y^{1M^l} - Y^{0M^l})$$

This effect measures the hypothetical direct effect of academic-track education when the university education is fixed at $M^l$. In our setting, where academic-track education $D$ is a prerequisite for university education $M$, such that $M^0 = 0$, we can not identify $Y^{01}$. Hence, we focus on $TTE = DTE(0) + ITE(1)$, that is, we decompose the wage return to the academic-track into the causal effect of $D$ in absence of additional university education

($DTE(0)$) and the causal effect of university education among individuals who attended the academic-track ($ITE(1)$).

## 3.2 Identification

In the following, we outline the approach to causal mediation analysis as developed by Schmitz and Westphal (2025). Frölich and Huber (2017) propose a related framework based on a control function approach. Other methods from the literature address endogeneity in either treatment or mediator, but not both simultaneously (see Dippel et al., 2020).

Assume that $D$ is randomly assigned to individuals. In this case, the total treatment effect can be identified as a simple difference in means. However, identifying direct and indirect effects is more complex, since the mediator is endogenous, implying that $M^0 \neq M^1$ for whatever reason. This endogenous response of the mediator is precisely what makes a causal mediation analysis both necessary and informative, yet also challenging.

Now we assume an individual-specific control variable $U_M$ exists, such that, conditional on $U_M$, the choice of the mediator $M$ is as good as random. Conditioning on $U_M$ thus addresses the endogeneity problem while we can assess the treatment- and mediator-specific heterogeneity along $U_M$ to learn about the essential factors of direct and indirect treatment effects.

With two valid instruments, both assumptions can hold in practice – even in a quasi-experimental setting where neither the treatment nor the mediator can be directly manipulated. We now sketch the approach using the binary instrument $Z_D$ to instrument the treatment $D$ and the continuous instrument $Z_M$ to instrument the mediator $M$ (plus control variables $X$, which we may include in all that follows, but omit them from the notation for simplicity). First, we construct a propensity score for the mediator $M$ using the continuous instrument $Z_M$ (plus $D$ and $Z_D$). Note that we demean all controls to ensure that the interpretation of the outcome level is preserved. In the next step, we construct means of the outcome, conditional on the three binary variables – $D$, $M$, and $Z_D$ – and the continuous propensity score $P(Z_M)$: This yields conditional expectation functions $m_{dz_dm}(p) := E(Y \mid D = d, Z_D = z_d, M = m, P(Z_M) = p)$ for each combination of $d$, $z_d$, and $m$ along the propensity score for $M$. If all possible combinations exist in the data and there is variation in $M$, this would result in eight distinct functions. In our case, we observe six different combinations, of which only four – those with $D = 1$ – exhibit variation in $M$ (as only individuals with Abitur can go to university). For combinations with $D = 0$, the propensity score is always zero, since $D$ perfectly predicts $M$. Accordingly, only four conditional expectation functions $m_{dz_dm}(p)$ vary with respect to the propensity score $p$.

We can now apply the framework of Imbens and Rubin (1997) to construct conditional expectation functions for each state of $M$, corresponding to those for the compliers to

$Z_D$ with and without academic-track education (indicated by $C$). This yields two different conditional expectation functions in our case, denoted by $m^C_{1m}(p)$. This approach addresses endogeneity of $D$ because, by definition, conditional on being a complier, the treatment $D$ is as good as random. The next step is to examine how these two conditional expectation functions vary along the propensity score $p$. Since we have conditioned on all other variables, the remaining variation must be caused by the instrument $Z_M$. Building on this insight, Heckman and Vytlacil (2005) show that the derivative of $m^C_{dm}(p)$ with respect to $p$ corresponds to a local IV estimator that unravels causal effects along an important univariate dimension of unobserved heterogeneity. They refer to this index as essential heterogeneity – the source of both endogeneity and variation in treatment effects that determines external validity of the results. Formally, Heckman and Vytlacil (2005) demonstrate that $m^C_{dm}(p)' = E(Y^{dm} \mid U_M = p, C)$, implying that this derivative identifies $U_M$, the essential variable that renders the mediator choice conditionally exogenous.

We now have two marginal treatment response (MTR) functions already informative for the mediation analysis: $E(Y^{1m} \mid U_M = p, C)$ for $m \in \{0, 1\}$. Averaging these functions over the support of $U_M$ (from 0 to 1), yields $E(Y^{1m} \mid C)$ for $m \in \{0, 1\}$. This allows us to compute the controlled ITE(1), i.e., differences across mediator states while holding the treatment state fixed: $E(Y^{11} - Y^{10} \mid C)$. The remaining step to obtain direct and indirect treatment effects as defined above is to move from controlled mediator states ($m \in \{0, 1\}$) to natural mediator states ($M^0$ or $M^1$) that correspond to each treatment state. We do so by weighting the conditional expectations $E(Y^{11} \mid U_M = p, C)$ and $E(Y^{10} \mid U_M = p, C)$ with the corresponding mediator probabilities:[5]

$$\begin{aligned} E(Y^{1M^l} \mid U_M = p, C) \quad = \quad & E(Y^{11} \mid U_M = p, C) E(M^l \mid U_M = p, C) \\ & + E(Y^{10} \mid U_M = p, C) E(1 - M^l \mid U_M = p, C) \end{aligned}$$

Note that $E(M^0 \mid U_M = p, C)$ is zero foreach $p$, such that $E(Y^{1M^1} \mid U_M = p, C) = E(Y^{11} \mid U_M = p, C)$. Finally, we average these expectations across $U_M$ to integrate out unobserved heterogeneity. The resulting two quantities allow us to identify the natural ITE(1). As we can simply identify the TTE with the typical two-stage least squares approach, we can derive the corresponding direct effect $DTE(0) = TTE - ITE(1)$. For further information on identification, see Appendix B, where the conditional expectation functions and all estimation steps are explained in more detail.

[5] For $D = 0$ the conditional expectation is only identified for $M = 0$ and is constant over $p$. For this case, the expectation for the controlled and natural mediation state is equal: $E(Y^{0M^0} \mid C) = E(Y^{00} \mid C)$.

### 3.3 Assumptions

Our approach relies on three assumptions. The first two assumptions are the general LATE assumptions for the instruments and not specific to mediation analysis:

***Assumption 1a—$Z_D$* is a valid instrument for *$D$*:** $Z_D$ is correlated with $D$, conditionally independent of the potential outcomes $Y^{dm}$, affecting the outcomes only through the treatment (exclusion restriction), without defiers existing (monotonicity).

***Assumption 1b—$Z_M$* is a valid instrument for *$M$*:** $Z_M$ is correlated with $M$ (relevance condition), conditionally independent of the potential outcomes $Y^{dm}$, affecting the outcome only through the mediator (exclusion restriction), without defiers existing (monotonicity).

The last assumption is crucial for the mediation analysis to distinguish between returns for the academic-track and university education.

***Assumption 2:*** $Z_D$ and $Z_M$ are conditionally independent.

# 4 Data and Baseline Empirical Strategy

### 4.1 Data

The primary data source for this study is the German National Educational Panel Study (NEPS), see Blossfeld and Roßbach (2019), which provides representative individual-level data on the educational trajectories of over 60,000 participants. NEPS employs a multicohort sequence design encompassing six "starting cohorts": newborns and their parents, preschool-aged children, students in grades 5 and 9, first-year university students, and adults. To examine the long-term effects of education, our analysis focuses on the adult cohort. In addition to detailed educational histories, the data include information on residential location at a fine geographic scale, which is essential for our analysis. Specifically, we require residential information before individuals select their secondary school track or pursue higher education to accurately assign measures of local access to academic-track schools or colleges. Because individuals face limited mobility at the time of secondary track choice, access to academic-track schooling is determined at the municipal level – using the municipality of residence in the final year of primary school, or, where unavailable, the municipality of birth (in approximately half of cases). For access to higher education, we rely on the district (*Kreis*) information.

We combine this data with purpose-built collections of academic-track school and college openings, presented in Figure 1. With manually coded opening years and precise geolocations of all academic-track schools existing in 2010, we built a panel with information on geographical access to an academic-track school for individuals from a particular municipality and birth cohort.[6] Thereby, we took varying starting times of school years as well as boys- and girls-only schools into account. For more information on the academic-track opening data, see Krumme and Westphal (2024). Information on colleges are the same as in Kamhöfer et al. (2019) and taken from the German Statistical Yearbooks 1959–1991 (German Federal Statistical Office, 1991).[7] Distance to college measures in km are calculated as the Euclidean distance between district centroids.

The NEPS-SC6 dataset comprises information on 17,140 individuals. We restrict our sample to respondents residing in West Germany (excluding Berlin) at the time of their secondary school track and college decisions, as the East German educational system differed substantially before reunification. For 8,944 individuals, we can match both academic-track school and university data based on residential history at the municipality and district levels. The NEPS-SC6 additionally provides detailed individual-level employment information, including monthly gross earnings and weekly working hours. Labor market outcomes over time are available for most respondents (8,887 individuals). To avoid capturing short-run effects arising from delayed labor market entry among highly educated individuals at younger ages, we focus on income observations between ages 35 and 65. This restriction excludes 152 individuals without income information in this age range. Furthermore, we limit our analysis to males to abstract from, for instance, fertility-related labor supply decisions (potentially causing selection into employment, see Westphal et al., 2022), resulting in a sample of 4,375 men. Finally, we exclude individuals with missing observations in any of the key variables used for estimation. Our final sample consists of 3,726 individuals.

**Academic-track and College Education**

The first educational variable, ”Abitur,” represents the general university entrance qualification, typically obtained upon graduation from an academic secondary school. Notably, this category excludes the *Fachabitur* (vocational baccalaureate diploma), which permits access only to specific fields of study at universities of applied sciences (*Fachhochschulen*). In our sample of 3,726 individuals, 1,180 (approximately one-third) hold an Abitur.

The second educational variable, “college degree,” is a binary indicator equal to 1 if an individual holds a degree from a traditional university, and 0 otherwise. This excludes

[6] Distances from municipalities to academic-track school provided as additional information are geodetic distances from the center point of the municipality to a school.

[7] Note that only colleges are used. Administrative data on openings and student numbers are unavailable for other higher educational institutions.

degrees from universities of applied sciences, thereby rules out individuals who hold a college degree without having obtained the Abitur.[8] Among the sample, 685 individuals have attained a university degree. As shown in Table 1, these are 57% of those who completed the academic-track leading to the Abitur.

**Dependent Variable**

The NEPS employment data provide information on gross monthly earnings and weekly working hours. Our primary outcome variable is gross monthly earnings in €, which exhibits the fewest missing observations and carries a lower risk of misreporting compared to other labor market indicators. When corresponding information on working hours is available[9], we additionally compute gross hourly wages by dividing reported monthly earnings by total working hours and the average number of weeks per month, approximated as 4.3. Periods of unemployment are assigned zero earnings and wages. Both outcomes are initially aggregated at the monthly level. We collapse these data to obtain annual averages of gross monthly earnings and gross hourly wages, covering the years 1975 to 2019. Annual earnings and wages are adjusted for inflation using the consumer price index (base year 2015). To mitigate potential biases from reporting errors or outliers, we cap earnings and wages at their 95th percentiles. Finally, we aggregate the average outcomes across all observed years per individual, yielding a cross-sectional dataset comprising 3,726 observations.

Table 1 presents descriptive statistics and reveals substantial differences in average earnings and wages across educational attainment levels. Individuals holding an Abitur exhibit higher average incomes and have access to higher education at universities. Among those with an Abitur, average earnings are even higher for individuals who have additionally obtained a university degree. As hourly wages are employed in our robustness analyses, descriptive statistics on wages are also reported in Table 1.

**Instruments**

We utilize information on access to academic-track schools at the municipal level and universities at the district level to instrument track and university choices. For the former, we use a dummy variable indicating the presence of at least one existing school in the municipality ("academic-track in municipality", $Z_D$).

[8] We adjust the Abitur dummy from 0 to 1 for 127 individuals with a degree from a university, as they likely have the academic-track or an equivalent secondary education.

[9] Information on actual working hours are partly missing. If available, contractual hours are used instead.

Table 1: Descriptive statistics of earnings wages by educational degrees

| | Abitur ($D = 1$) | | no Abitur ($D = 0$) |
|---|---|---|---|
| | college degree ($M = 1$) | no college degree ($M = 0$) | no college degree ($M = 0$) |
| *Monthly earnings:* | | | |
| Mean | 5,858.09 | 4,924.56 | 3,573.02 |
| Min | 0 | 0 | 0 |
| Max | 10,786.08 | 10,786.08 | 10,786.08 |
| N | 685 | 495 | 2,546 |
| *Hourly wage:* | | | |
| Mean | 33.81 | 27.92 | 20.55 |
| Min | 0 | 0 | 0 |
| Max | 62.38 | 62.38 | 62.38 |
| N | 676 | 488 | 2,515 |

*Notes:* Own calculations based on NEPS-SC6 data. Gross earnings and gross hourly wages are truncated above the 95% percentile.

Our instrument for university education is the continuous university availability index constructed by Kamhöfer et al. (2019), which combines information on new college openings and the expansion of existing ones. The continuous index is defined by

$$Z_{M,it} = \sum_{j}^{326} K(dist_{ij}) \times \left( \frac{\#students_{jt}}{\#inhabitants_{jt}} \right). \tag{2}$$

It captures the local supply of higher education opportunities for individual $i$ in year $t$ by measuring the number of college spots (proxied by enrolled students) per inhabitant in each district $j$, weighted by the distance from individual $i$'s home district. To account for regional variation in college size, the student count is normalized by district population. These values are then distance-weighted using a Gaussian kernel based on the distance between the centroids of the home and the district $j$. Hence, the instrument for college education is assigned at the district level and is less locally defined, as college students are less restricted to their (parents') residence when deciding to enroll in a college. It includes the sum of all district-specific college availabilities within a 250 km bandwidth, applying kernel weights that give higher importance to nearby colleges. For instance, colleges in the same district receive a weight of 0.40, those 100 km away receive 0.37, and the weight drops to 0.24 at 250 km. Colleges 500 km away receive only the minimal weight (0.05), reflecting their limited relevance for local access. We use an alternative dichotomous instrument $\tilde{Z}_M$ for robustness checks, which equals one if at least one college is within a 30 km radius, and zero otherwise.

Table 2 provides summary statistics for both instrumental variables together with additional background information.

Table 2: Descriptive statistics of instruments with background information

| | Statistics | | | |
|---|---|---|---|---|
| | Mean | SD | Min | Max |
| *Instruments:* | | | | |
| Academic-track in municipality ($Z_D$) | 0.619 | 0.486 | 0 | 1 |
| College availability index ($Z_M$) | 0.486 | 0.261 | 0.051 | 1.128 |
| $\geq$ 1 college within 30 km ($\tilde{Z}_M$) | 0.602 | 0.490 | 0 | 1 |
| *Background information on instruments:* | | | | |
| Distance to nearest academic-track school | 3.702 | 4.637 | 0 | 35.3 |
| Distance to nearest college | 27.854 | 26.383 | 0 | 171.355 |
| Observations | 3,726 | | | |

*Notes:* Own calculations based on NEPS-SC6 data. Distances to academic-track schools rely on geodetic distances from the center of the municipality to a school, and distances to colleges measure distances between district centroids.

**Control Variables**

The choice of control variables is similar to the specifications used in Krumme and Westphal (2024). Most importantly, we include cohort and district fixed effects to account for differential trends, wage levels, education, and academic-track availability across districts. As regions may have implemented the educational expansion differently, we also account for district-specific linear trends. Moreover, we include the distances (in 10km steps) from the residential municipality to the next academic-track school before 1940 and binary indicators for the father's degree. Missing values for paternal education are coded as zero; however, an additional indicator for a missing value in the father's degree is also included. Table A.1 in the Appendix provides definitions and mean values by schooling degree for all covariates included in the main specification and for variables used in supplementary regressions.

The biggest threats to identification are any differences between municipalities within districts. Unfortunately, the inclusion of municipality fixed effects is not feasible with our sample, which consists of only a single observation for many municipalities. However, the most substantial expected differences between municipalities are likely driven by larger cities that may be systematically distinct from smaller and more rural municipalities. We primarily control for these differences by including district-fixed effects, as urban districts in Germany are not subdivided into municipalities. The additional control variables further mitigate remaining heterogeneity related to parental education and initial access to schooling between municipalities (within districts).

## 4.2 Baseline strategy

Our analysis centers on the three essential variables: earnings ($Y$), the study indicator ($M$), and the Abitur indicator ($D$). We first focus on the simple overall causal effects of $D$ on $M$

and $Y$, before the mediation analysis (detailed in the appendix) disentangles the impact of $M$ on $Y$ from the effect of $D$ on $Y$.

Since individuals (or their parents) may self-select into a particular secondary school track, a simple regression of $Y$ (or $M$) on the academic-track indicator $D$ is endogenous, and the academic-track coefficient does not measure its average wage return. Consequently, we utilize openings of academic-track schools, measured by the variable $Z_{D,i}$, which is assumed to be exogenous to individual decisions. With this quasi-experimental variation, we can estimate wage returns by two-stage least squares (2SLS). The regression model reads

$$\begin{aligned} D_i &= \pi_0 + \pi_1 Z_{D,i} + X_i'\pi_C + \nu_i \\ M_i &= \alpha_0 + \alpha_1 \widehat{D}_i \quad + X_i'\alpha_C + u_i \\ Y_i &= \beta_0 + \beta_1 \widehat{D}_i \quad + X_i'\beta_C + \varepsilon_i \end{aligned}$$

In the first regression (the first stage), we regress the Abitur indicator of individual $i$ ($D_i$) on the instrument $Z_{D,i}$ and a vector of control variables $X_i$. Under the conventional IV validity assumption (Assumption 1a), $\pi_1$ captures the complier share – individuals who only attend an academic-track school because of changes in $Z_{D,i}$. We have two second stages where we account for endogeneity by using the first stage predicted values that normalize the effects of $Z_{D,i}$ on $M_i$ and $Y_i$ to a change in $D_i$.

The corresponding coefficient $\alpha_1$ captures the causal effect of obtaining the Abitur on the probability of holding a college degree for compliers with respect to $Z_D$. Similarly, $\beta_1$ measures the causal effect of Abitur attainment on gross earnings $Y_i$ for the same group of compliers and corresponds to the total treatment effect within the mediation framework. We then estimate direct and indirect treatment effects using the variables above, along with the instrument $Z_M$, following the approach outlined in Section 3 and detailed in Appendix B. We cluster our standard errors at the district level in all our regressions (including the mediation analysis).

Consistent with Assumption 1a, we need to assume that, conditional on the controls $X_i$, the instrument is unrelated to unobserved factors that correlate with education and wages (i.e., $\nu_i, u_i, \varepsilon_i \perp\!\!\!\perp Z_{D,i}$). To make this assumption credible, we include a large set of fixed effects in $X_i$, specified above. We additionally assume that the exclusion restriction holds, i.e., that other municipality-level changes do not coincide with the timing of the academic-track openings. Given a rich set of control variables that also absorb district-specific linear trends, this assumption seems plausible. 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.

Whenever analytical standard errors are unavailable (i.e., for mediation analysis), we use the Bayesian bootstrap, which accounts for the high dimensionality and associated multicollinearity problems with resampled data, including many fixed effects (Rubin, 1981).[10]

# 5 Results

## 5.1 Baseline regressions

**Effect of academic-track completion on earnings**

Table 3 reports OLS association between earnings and Abitur, as well as reduced form effects of the instrument on earnings, Abitur (the first stage), and 2SLS results of Abitur on earnings. The OLS coefficient in column (1) indicates a significant positive correlation between Abitur and wages. On average, individuals with Abitur have 1,576 € higher earnings than those without. However, due to self-selection mechanisms into higher secondary schooling, we cannot interpret this as a causal effect. Therefore, we use quasi-experimental variation and complement the table by the reduced form result in column (4). The coefficient indicates an effect of academic-track ability on earnings of around 380 € , which is significant at the 5% level.

The first stage results in column (2) document that access to an academic-track school at the end of elementary school increases the probability of obtaining the Abitur by 9 percentage points, relative to an unconditional mean of 32 percent. This effect is statistically significant at the 1% level, and the corresponding F-statistic of 14 (above the conventional threshold of 10 suggested by Staiger and Stock, 1997) dispels concerns about weak instrument bias. Finally, column (3) shows the LATE, i.e., the complier-specific causal effect of Abitur on earnings. The result of 4,235.5 € corresponds to a significant increase in earnings of around 118% due to Abitur for compliers (compared to average earnings without Abitur). This substantial effect is statistically significant at the 5% level and comparable to the impact of upper secondary schooling on wages in Indonesia reported by Carneiro et al. (2017).

In contrast, Krumme and Westphal (2024) find a smaller effect of obtaining the Abitur on earnings during the first ten years after labor market entry. Including unemployed individuals and those out of the labor force in our sample may partly account for the larger estimated returns, as their earnings are recorded as zero, and they are disproportionately

[10] The Bayesian bootstrap renders the same expected variation in each bootstrap draw, but instead of resampling with replacement (or reweighting with discrete weights drawn from a multinomial distribution that include many zeros), Bayesian Bootstrap reweights with the corresponding (i.e., conjugate prior distribution from the binomial distribution) continuous distribution (the Dirichlet distribution). These weights are never zero, thereby preserving the same collinearity structure of the data.

concentrated among individuals without an Abitur. Additionally, the relatively high average age of individuals in our sample (46 years) may contribute to these larger effects, as Bhuller et al. (2017) document increasing returns to education over the life cycle. Somewhat surprisingly, the LATE parameter is more than twice the size of the OLS estimate. This, however, is a common finding in the literature on monetary returns to education, which is based on instruments regarding the supply side. Compliers identified by supply-side variation tend to face higher educational costs or barriers rather than lower ability (Card, 2001).

Table 3: Regression results: Effects of Abitur on earnings

| | (1) OLS | (2) First Stage | (3) 2SLS | (4) Reduced Form |
|---|---|---|---|---|
| Abitur (D) | $1,576.20^{***}$ | | $4,235.50^{**}$ | |
| | (125.20) | | (1,775.80) | |
| Acad. track in municipality ($Z^D$) | | $0.090^{***}$ | | $381.50^{**}$ |
| | | (0.024) | | (165.60) |
| F-statistic (instrument) | | 14.213 | | |
| Observations | 3,726 | 3,726 | 3,726 | 3,726 |

*Notes:* Own calculations based on NEPS-SC6 data. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940. Standard errors in parentheses clustered on district level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

**Potential Mechanism: Subsequent College Degree**

We want to differentiate the mediating effect of a college degree from the remaining monetary returns to Abitur. For $M$ to be a potential channel, there must be a causal impact of Abitur on college education and from college education on earnings. Results on the former relation are shown in Table 4. OLS and 2SLS regressions result in significantly positive effects at the 1% and 5% level, respectively. Abitur is associated with a nearly 55 percentage point higher chance of obtaining a university degree, as indicated by the OLS estimates. For the group of compliers, obtaining an Abitur leads to a 37 percentage point increase in the probability of obtaining a college degree. This large effect is unsurprising as Abitur certifies the successful completion of the academic-track, i.e., the track designed to prepare for tertiary education.

There is evidence in the literature for a positive effect of college education on earnings for different countries, including Germany (see, e.g., Carneiro et al., 2011; Nybom, 2017; Westphal et al., 2022). This also holds for our sample when analyzing the effect of college degrees on earnings, using our college availability instrument $Z_M$. The results are presented in Table 5. The first stage results in column (2) are highly significant and of the

Table 4: Regression results: effects of Abitur on college degree

| | (1) OLS | (2) 2SLS |
|---|---|---|
| Abitur (D) | 0.547*** | 0.371** |
| | (0.019) | (0.157) |
| Observations | 3,726 | 3,726 |

*Notes:* Own calculations based on NEPS-SC6 data. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940. Standard errors in parentheses clustered on district level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

expected sign, as a larger index indicates higher college availability.[11] The F-statistic of over 240 further eliminates any worries about weak instruments. On average, individuals with a college degree earn around 1,817 € per month more (after controlling for covariates). The LATE indicates a significantly positive effect of a college degree on earnings of around 3600 € for the compliers. Compared to average earnings for individuals without college education, this yields a relative effect of around 95%.

Table 5: Regression results: effects of college degree on earnings

| | (1) OLS | (2) First Stage | (3) 2SLS |
|---|---|---|---|
| College degree (M) | 1,817.30*** | | 3,597.2*** |
| | (137.90) | | (382.4) |
| College availability ($Z^M$) | | 2.218*** | |
| | | (0.143) | |
| F-statistic (instrument) | | 240.87 | |
| Observations | 3,726 | 3,726 | 3,726 |

*Notes:* Own calculations based on NEPS-SC6 data. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940. Standard errors in parentheses clustered on district level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

The final prerequisite for the mediation analysis is the propensity score. Figure A.1 in the Appendix shows its distribution by college degree ($M$) for individuals with Abitur ($D = 1$). Unsurprisingly, we see an accumulation of propensity scores at zero for those without a college degree, and vice versa, an accumulation at one for those with a college

[11] For an assessment of the first stage effect size, we refer to Kamhöfer et al. (2019), who give an intuition based on an exemplary city with a university opening.

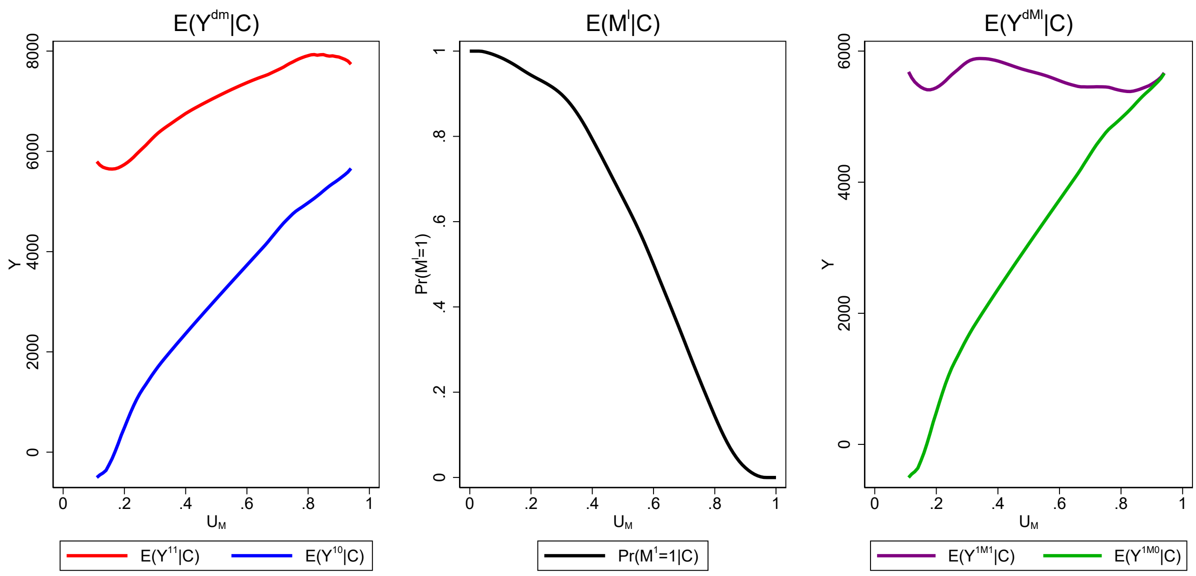
degree. Yet, the common support ranges from 0.107 to 0.943, which we use as truncation points for the propensity score in the subsequent outcome regressions.

### 5.2 Mediation Analysis

Figure 2 illustrates the three key components of our mediation analysis graphically, along the unobserved index $U_M$, conditional on which the choice of the mediator is as good as random. For clarity, we omit the confidence intervals here. In the left panel, we plot the two conditional expectation functions (controlling for academic-track and university choices, as indicated by the first and second potential outcome superscripts) for the $Z_D$-compliers along $U_M$. $U_M$ on the x-axis typically represents resistance to college education; it therefore inversely reflects unobserved characteristics that influence an individual's likelihood of attending college. The function $E(Y^{11} \mid U_M = p, C_D)$ increases slightly with higher values of $U_M$. Hence, low resistance to college education does not necessarily correspond to higher wages among university graduates (a typical finding that is also reported in Carneiro and Lee, 2009 or Westphal et al., 2022). In contrast, $E(Y10 \mid U_M = p, C_D)$ increases more steeply with unobserved resistance to college education. Examining the difference between the two functions reveals a decreasing return to college education along $U_M$ for compliers with academic-track schooling, corresponding to a declining marginal treatment effect. $Z_D$-compliers who obtain the highest gross returns to college education are also those most likely to attend college. This pattern also aligns with findings in the literature on monetary returns to education, which consistently document selection into gains (e.g., Carneiro et al., 2011, Carneiro et al., 2017, Kamhöfer and Westphal, 2019, Nybom, 2017). Overall, the pattern corroborates the notion that $U_M$ measures a relative advantage for university education, but not an absolute advantage in the labor market. In the middle panel of Figure 2, the college degree probabilities for compliers to academic-track availability with Abitur are presented. The probability decreases with increasing $U_D$, as expected. The likelihood for a university degree for individuals with $D = 0$ is always zero, such that $Pr(M^0 = 1|C) = 0$ at each value of $U_M$, and, thus, excluded here.

The right panel presents the mediated outcomes, i.e., the two expected potential outcome functions by potential college education for compliers with Abitur weighted by the probability shown in the middle panel. For $Y^{1M^1}$, the complier-specific outcomes are observed, while $Y^{1M^0}$ indicates the counterfactual complier-specific outcomes if college education is fixed to the other state. The average difference over the $U_D$ interval between the resulting lines yields the natural indirect treatment effect, ITE(1). Weighting the $Y^{11}$ of the $Z_D$ compliers by their college degree probabilities based on $D$ is essential, as not every complier with Abitur subsequently obtains a college degree. The difference between the left and right panels also highlights the relevant distinction between the controlled (left panel) and natural (right panel) effects.

Figure 2: MTR functions for compliers to academic-track school access




*Notes:* Own calculations based on NEPS-SC6 data. Number of observations is 3,726. The applied bandwidth for the semiparametric MTR estimation is 0.25. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940.

The main results derived from the components shown in Figure 2 are presented in Table 6. We report confidence intervals instead of standard errors as the distribution of bootstrapped parameters is not symmetric around the estimate. We further assess statistical significance using a two-sided recentered test, testing $H_0 : TE = 0$ versus $H_1 : TE \neq 0$, and report the corresponding p-values.[12] Besides the total treatment effect (TTE), the indirect and direct treatment effects are given. As stated in Section 3, we only identify the indirect effect for individuals with Abitur, ITE(1), and the direct effect of Abitur for individuals without college education, DTE(0). For the ITE(1) we further distinguish between the controlled ($E(Y^{11} - Y^{10} \mid C)$) and the natural effect ($E(Y^{1M^1} - Y^{1M^0} \mid C)$). For the DTE(0), the controlled and the natural effect coincide.[13]

The total treatment effect in column (1) is, by definition, the same as the local average treatment effect of Abitur on earnings shown in Table 3. The given confidence interval excluding zero and the p-value of 0.052 indicate that it is significant at the 10% level. Therefore, the bootstrap-based inference implies slightly less precision than the analytic standard errors shown in Table 3.

Examining the main results for indirect and direct treatment effects, we find positive coefficients for both components. The ITE(1) quantifies whether – and to what extent – college education yields additional returns once an individual has already obtained the

[12]The p-value is calculated following $p = \frac{\#(|\hat{\theta}_b - \hat{\theta}_{obs}| \geq \hat{\theta}_{obs})}{B}$, where $\hat{\theta}_{obs}$ indicate the observed estimate and $\hat{\theta}_b$ the $B = 1000$ bootstrap estimates. The formula counts the number of bootstrap replicates that lie in the tails and divides this number by the total number of bootstrap repetitions.

[13]This results from the fact that $E(Y^{1M^1}|U_M = p, C) = E(Y^{11}|U_M = p, C)$ and $E(Y^{00}|C) = E(Y^{0M^0}|C)$.

Table 6: Total, direct, and indirect treatment effects

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Total treatment effect | Effect decomposition | | |
| | | Indirect TE | | Direct TE |
| | TTE=LATE | natural ITE(1) | controlled ITE(1) | DTE(0) |
| Treatment effect | 4,235.50 | 2,543.66 | 3,953.32 | 1,691.85 |
| 90% CI | [1539.09, 8272.76] | [1,035.25, 4,856.60] | [2,038.83, 7,060.14] | [-1,303.47, 5,391,98] |
| p-value | 0.052 | 0.044 | 0.028 | 0.382 |

*Notes:* Own calculations based on NEPS-SC6 data. Number of observations is 3,726. The applied bandwidth for the semiparametric MTR estimation is 0.25. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940. Confidence intervals and p-values are based on bootstrapping (1000 replications) with clustering at the district level (Baysian Bootstrap).

highest secondary schooling degree. For compliers with Abitur, this effect amounts to 2,544 €, representing approximately 60% of the total effect. Relative to average earnings among individuals with Abitur but without a college degree (3,793.02 €), this corresponds to a proportional increase of 67%. The effect is sizable in magnitude and statistically significant at the 5% level. By definition, the controlled ITE(1) in column (3) exceeds the natural ITE(1) from column (2). Fixing the mediator state at 1 – implying that all individuals with Abitur graduate from college – yields an estimated effect of subsequent college education of approximately 4,000 €. With a p-value below 0.05, this result is also statistically significant at the 5% level. The substantial difference between the two ITE(1) estimates underscores the relevance of distinguishing between controlled and natural effects. Overall, we interpret the results on the ITE(1) as evidence of a large and significant positive impact of additional college education on earnings among individuals already holding an academic-track degree.

The DTE(0) given in column (4) is defined as the difference between the total treatment effect (TTE) and the natural ITE(1). It captures the effect of obtaining the Abitur for compliers with respect to academic-track school availability when the probability of college attendance is fixed at $M^0 = 0$. In terms of magnitude, with a coefficient of 1,692 €, the relative direct effect for compliers amounts to approximately 47% of average earnings among individuals without an Abitur. However, the estimated DTE(0) is not statistically significant. Given this lack of precision, we cannot conclusively infer a positive direct effect of Abitur attainment on earnings in the absence of subsequent college education – even though the estimated magnitude is economically meaningful. Consequently, while returns to academic-track education for individuals pursuing vocational training or degrees from universities of applied sciences – rather than traditional university education – appear positive on average, they remain imprecisely estimated.

Our results are most closely related to those of Biewen and Thiele (2020), the only other study that implicitly decomposes wage returns. When aggregating all academic-track versus non–academic-track trajectories to obtain the total treatment effect (their ATE, identified using a latent factor model) and university versus non-university trajectories within the academic-track to obtain $ITE(1)$, weighted by the respective trajectory shares, they report $ATE = 0.216$ and $ITE(1) = 0.153$ (both in log points). This implies that $ITE(1)$ accounts for 71% of the total effect – a substantially higher relative contribution than what we find. Although their analysis is also based on NEPS data, it is important to emphasize that their target parameter is the average effect for the entire population, whereas our design identifies the local effect for compliers with school openings. Moreover, unlike our approach, their identification strategy does not rely on quasi-experimental variation and instead depends on the assumptions of a latent factor model.

## 5.3 Robustness checks

We present additional results to assess the robustness of our estimates with respect to changes in the outcome variable, the instrument for college education, and the estimation approach as well as sample composition. The first set of results in Table 7 reports mediation estimates using hourly wages as the outcome variable; the second set is based on the binary instrument for college education $\tilde{Z}_M$; and the third set relies on linear conditional expectation functions. The last row displays results for a sample restricted to employed men, thereby excluding the extensive margin of monetary returns to education. As before, we report 90% confidence intervals together with p-values from a two-tailed test.

The effects on hourly wages closely resemble the main results in both magnitude and statistical significance. In relative terms, ITE(1) is smaller while DTE(0) is correspondingly larger. Nevertheless, ITE(1) remains significant at the 5% level, whereas DTE(0) does not reach conventional levels of significance. The relatively smaller indirect effect suggests that a college degree raises total earnings partly through an increase in working hours; however, this mechanism accounts for only a limited portion of the overall impact, as the effect remains substantial and statistically significant even when controlling for hourly wages. Employing an alternative instrument likewise does not alter the interpretation of our main findings. When using the binary instrument for the mediator, ITE(1) becomes slightly smaller while DTE(0) increases marginally. These differences may reflect minor variations in the composition of the $Z_M$-complier group. Overall, our main conclusions remain robust to changes in both the outcome variable and the choice of instrument for college education. We additionally show results based on linear MTR functions for comparison. The estimated effects are similar in magnitude: ITE(1) is slightly smaller, whereas DTE(0) increases accordingly. At the same time, the results for the direct and indirect effects lose precision, such that the natural indirect effect is no longer significant at the 10% level. This loss of precision suggests that the relationship may be nonlinear and that the flexible functional form captures relevant variation that the purely linear model misses.

Table 7: Robustness checks

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Total treatment effect | Effect decomposition | | |
| | | Indirect TE | | Direct TE |
| | TTE=LATE | natural ITE(1) | controlled ITE(1) | DTE(0) |
| Outcome: wages | 24.632 | 12.909 | 20.440 | 11.723 |
| 90% CI | [10.245; 47.148] | [4.498; 23.838] | [8.827; 33.026] | [-3.942; 33.623] |
| p-value | 0.044 | 0.043 | 0.017 | 0.290 |
| Instrument: $\tilde{Z}_M$ | 4,235.50 | 2,826.32 | 4,102.50 | 1,409.18 |
| 90% CI | [1,539.09, 8,272.76] | [1,029.10, 5,122.00] | [2,005.43, 7,125.30] | [-1,449.18, 5,360.58] |
| p-value | 0.052 | 0.035 | 0.027 | 0.489 |
| Linear MTRs | 4,235.50 | 2,344.48 | 3,999.74 | 1,891.03 |
| 90% CI | [1,539.09; 8,272.76] | [-412.31; 4,928.05] | [60.96; 8,025.03] | [-1,534.52; 6,538.21] |
| p-value | 0.052 | 0.132 | 0.098 | 0.404 |
| Only working | 3,979.41 | 2,552.19 | 4,991.78 | 1,427.22 |
| 90% CI | [1,147.70; 7,779.98] | [348.04; 5,044.05] | [1,927.02; 7,883.04] | [-1,952.29; 5,700.94] |
| p-value | 0.061 | 0.078 | 0.041 | 0.481 |

*Notes:* Own calculations based on NEPS-SC6 data. The applied bandwidth for the semiparametric MTR estimations is 0.25. Row 1 shows results for the alternative outcome of hourly wages ($N = 3,697$), row 2 for the use of the alternative instrument $\tilde{Z}$ ($N = 3,726$), results in row 3 are based on linear MTR estimations, and row 4 shows the results of the main specification with the sample restricted to earnings above 0 ($N = 3,355$). Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940. Confidence intervals and p-values are based on bootstrapping (1000 replications) with clustering at the district level (Baysian Bootstrap).

Restricting the sample to employed males has a more pronounced impact on the mediation results. The total treatment effect decreases but remains statistically significant at the 10% level. Since selection into employment is captured in the main results but excluded here, a reduction in the estimated effect is expected. Precision of the indirect and direct effects is overall lower, but the result for the natural (controlled) ITE(1) remains significant at the 10% (5%) level. Examining coefficient magnitudes reveals slightly higher estimates for ITE(1) and lower estimates for DTE(0) relative to the TTE. This pattern suggests that part of the main results is driven by stronger effects of Abitur attainment for individuals without subsequent college education, compared to the effects of an additional college degree on labor market participation among individuals with Abitur. Since selection into employment based on educational attainment constitutes an integral part of the overall effect on labor market outcomes, we do not interpret the deviations observed in this restricted sample as evidence against the general validity of our main results.

The main results (all direct and indirect effects) of the mediation analysis are based on MTR functions estimated semiparametrically (see left panel of Figure 2) with a bandwidth of 0.25. We document the robustness of the estimated MTR curves to changes in the bandwidth size. Figure A.2 in the Appendix shows results for $E(Y^{11} \mid C)$ and $E(Y^{10} \mid C)$ along $U_M$ for different bandwidths varying from 0.15 to 0.5, with lighter colors indicating smaller bandwidths. Aside from some deviations at small bandwidths and higher levels of resistance to college education, the functions are overall compatible and correspond to similar indirect treatment effects (ITE(1)) ranging from 2,391 € to 2,624 €. Because the MTR functions vary only slightly, we conclude that our main results are robust to changes in the bandwidth and reflect genuine underlying relationships between educational attainment and earnings outcomes.

### 5.4 The validity of the two instruments and their independence

For a causal interpretation of our results, both instruments must be unrelated to unobserved factors correlated with educational attainment, conditional on the control variables (assumptions 1a and 1b). To examine potential observable pre-determined differences between individuals with varying values of $Z_D$ or $Z_M$ after conditioning on controls, we conduct balancing tests presented in Table 8. Accordingly, we regress a set of individual characteristics – determined before school track or college choice but potentially related to these decisions – on each instrument and all covariates.[14] For the regressions with $Z_M$, we additionally include $D$ (as for the propensity score estimation), compared to the regressions with $Z_D$. For each instrument, we first report results without additional controls and then include the full set of covariates specified in Section 4.

We find no significant relationships between either instrument and any predetermined characteristics, except for one coefficient in column (1). The positive association with the father's degree in this specification likely reflects residual variation at the municipality level that is not yet absorbed by the baseline controls. After including dummies for paternal education – thereby capturing systematic differences in parental educational composition across municipalities – the association disappears. In columns (2) and (4), all remaining coefficients are statistically insignificant, reinforcing the credibility of the identifying assumption. Based on these results, we conclude that the validity assumptions most likely hold when conditioning on our rich set of controls. Nevertheless, we cannot entirely rule out the presence of remaining unobserved confounders.

$Z_M$ must also be conditionally independent of $Z_D$ (assumption 2). This implies that $Z_D$ should not exhibit any predictive power for $Z_M$ once conditioning on the covariates. Since we employ $Z_M$ as an instrument only for individuals with Abitur, this independence condition must hold given $D = 1$. To verify this, we first regress $Z_M$ on $Z_D$, $D$, and $X$,

[14] The two variables describing family constellation refer to conditions at age 15. Any differences in family circumstances between ages 10 and 15 are unlikely to be related to access to academic-track schools.

Table 8: Balancing checks

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Explanatory variable: | | | |
| | $Z_D$ | | $Z_M$ | |
| Dependent variables: | | | | |
| Firstborn | −0.030 | −0.039 | −0.107 | −0.107 |
| | (0.029) | (0.030) | (0.175) | (0.175) |
| No. of older siblings | 0.192 | 0.247 | −0.237 | −0.206 |
| | (0.015) | (0.151) | (1.914) | (1.925) |
| No. siblings | −0.117 | −0.092 | 0.391 | 0.380 |
| | (0.105) | (0.112) | (0.727) | (0.728) |
| Father's degree | 0.246*** | 0.000 | 0.162 | 0.000 |
| | (0.065) | (0.000) | (0.488) | (0.00) |
| Father born in Germany | −0.018 | −0.020 | −0.049 | −0.033 |
| | (0.015) | (00.015) | (0.106) | (0.102) |
| Raised by single parent | −0.005 | −0.008 | −0.075 | −0.082 |
| | (0.013) | (0.014) | (0.090) | (0.091) |
| Raised by patchwork family | 0.010 | 0.009 | −0.007 | −0.011 |
| | (0.010) | (0.010) | (0.058) | (0.060) |
| Additional controls | no | yes | no | yes |

Notes: Own calculations based on NEPS-SC6 data. The number of observations varies with the dependent variable from 2992 to 3726 individuals. The table displays the coefficients of the $Z_D$-instrument, "Academic-track in municipality," and the $Z_M$-instrument, "College availability," for various outcomes. All specifications include district and entry cohort fixed effects and district-specific trends. In columns 2 and 4, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940 are additionally included. Results for $Z_M$ are given conditional on $D$ (Abitur). Standard errors in parentheses are clustered on the district level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

and test for statistical significance. In a second step, we restrict the sample to individuals with Abitur and estimate the same regression (excluding $D$). The results are presented in Table A.2. The very small coefficients, together with the absence of statistical significance, support the validity of this assumption.

# 6 Conclusion

This paper provides new evidence on how secondary and tertiary education interact in shaping individual earnings. Using quasi-experimental variation from school and university openings during Germany's educational expansion, we disentangle the direct and indirect earnings effects of academic-track schooling (Abitur) through subsequent college education. Applying a causal mediation framework within the marginal treatment effect setting allows us to jointly address endogeneity in both educational stages – school track choice and university enrollment – and thus to identify local treatment effects for compliers affected by these institutional changes.

Our results reveal substantial monetary returns to academic-track schooling. On average, compliers benefit from an increase in gross earnings of roughly 118%, which partly reflects higher probabilities of college attendance induced by access to the academic-track. The indirect effect of college education accounts for 60 percentage points of total returns and is statistically significant at the 5% level, whereas the direct effect of Abitur – holding college education constant – is smaller in magnitude and not statistically significant. These findings suggest that most long-run wage gains associated with completing the academic-track operate through increased access to tertiary education rather than through secondary schooling alone, although secondary education itself may improve prospects, for example, for high-quality vocational training.

Robustness analyses confirm that our main conclusions are stable across alternative specifications using hourly wages as outcomes, different instruments for college education, a linear MTR specification, and different bandwidths. Restricting the sample to employed men reduces overall magnitudes but does not challenge the general interpretation: selection into employment itself constitutes part of the overall labor market return to education. The results, therefore, underscore that both margins – extensive and intensive – jointly contribute to the significant indirect effect of subsequent college education for compliers with Abitur.

Taken together, our findings highlight that upper-secondary schooling plays a pivotal role not (only) as an independent driver of economic success but also as a gateway to higher education with substantial long-term benefits. From a policy perspective, this suggests that reforms targeting early tracking decisions can have far-reaching implications beyond secondary school completion, influencing access to tertiary education and labor market outcomes decades later.

# References

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.

Betts, J. R. (2011). The Economics of Tracking in Education. In *Handbook of the Economics of Education*, volume 3, pages 341–381. Elsevier.

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.

Biewen, M. and Thiele, M. (2020). Early tracking, academic vs. vocational training, and the value of 'second-chance' options. *Labour Economics*, 66:101900.

Blossfeld, H.-P. and Roßbach, H.-G., editors (2019). *Education as a lifelong process*. Springer Fachmedien Wiesbaden, Wiesbaden.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond LATE with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

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. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754–81.

Carneiro, P. and Lee, S. (2009). Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality. *Journal of Econometrics*, 149(2):191–208.

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.

Cygan-Rehm, K. and Westphal, M. (2024). School starting age and the gender pay gap over the life cycle. Working Paper 1115, RWI. Series: Ruhr Economic Papers.

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

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

Dippel, C., Ferrara, A., and Heblich, S. (2020). Causal mediation analysis in instrumental-variables regressions. *The Stata Journal*, 20(3):613–626. Publisher: SAGE Publications.

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

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.

Frölich, M. and Huber, M. (2017). Direct and indirect treatment effects-causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(5):1645–1666.

German Federal Statistical Office (various issues, 1959–1991). Statistisches Jahrbuch für die Bundesrepublik Deutschland. Technical report, German Federal Statistical Office (Statistisches Bundesamt), Wiesbaden.

Gunderson, M. and Oreopolous, P. (2020). Returns to education in developed countries. *The Economics of Education*, pages 39–51.

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

Hollenbach, J., Schmitz, H., and Westphal, M. (2024). *Gene-environment interactions with essential heterogeneity*. Number 1105. Ruhr Economic Papers.

Huber, M. (2020). Mediation Analysis. In *Handbook of Labor, Human Resources and Population Economics*, pages 1–38. Springer International Publishing, Cham.

Imbens, G. W. and Rubin, D. B. (1997). Estimating outcome distributions for compliers in instrumental variables models. *The Review of Economic Studies*, 64(4):555–574.

Kamhöfer, D. A. and Westphal, M. (2019). Fertility effects of college education: Evidence from the german educational expansion. In *DICE Working Paper*, volume 316. Düsseldorf.

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.

Krumme, A. and Westphal, M. (2024). Monetary Returns to Upper Secondary Schooling, the Evolution of Unobserved Heterogeneity, and Implications for Employer Learning. *Ruhr Economic Papers*, 1130.

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

Nybom, M. (2017). The distribution of lifetime earnings returns to college. *Journal of Labor Economics*, 35(4):903–952.

Patrinos, H. A. and Psacharopoulos, G. (2020). Returns to education in developing countries. In *The Economics of Education*, pages 53–64. Elsevier.

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

Rubin, D. B. (1981). The bayesian bootstrap. *The Annals of Statistics*, 9(1):130–134.

Schmitz, H. and Westphal, M. (2021). *Causal mediation analysis of the effect of education on cognitive abilities*. mimeo.

Schmitz, H. and Westphal, M. (2025). Early- and later-life stimulation: How retirement shapes the effect of education on old-age cognitive abilities. *Ruhr Economic Papers*.

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

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

Zamarro, G. (2010). Accounting for heterogeneous returns in sequential schooling decisions. *Journal of Econometrics*, 156(2):260–276.

# Appendix

## A Additional tables and figures

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

| | Definition | Without Abitur | With Abitur |
|---|---|---|---|
| Secondary school entry | Year individual entered secondary school | 1971 | 1971 |
| Start distance | Distance to closest academic-track school (in 10 km steps, rounded down) in 1940 | 0.251 | 0.166 |
| Father's degree | Highest general school-leaving qualification of father (categorial) | | |
| 1 | basic school-leaving qualification | 0.735 | 0.472 |
| 2 | intermediate school-leaving qualification | 0.092 | 0.163 |
| 3 | vocational baccalaureate diploma (Fachabitur) | 0.016 | 0.054 |
| 4 | University entrance qualification (Abitur) | 0.052 | 0.268 |
| 5 | school-leaving qualification of a special needs school | 0.000 | 0.000 |
| 6 | other qualification | 0.001 | 0.001 |
| *Background information:* | | | |
| Birth year | Year the individual was born | 1960 | 1961 |
| Start distance (continuous) | Distance to closest academic-track school in 1940 | 5.536 | 3.994 |
| Father born in Germany | =1 if father was born in Germany | 0.935 | 0.935 |
| Raised by single parent | =1 if raised by a single parent (from birth to age 15) | 0.062 | 0.042 |
| Raised by patchwork family | =1 if raised in a patchwork family (from birth to age 15) | 0.046 | 0.025 |
| Firstborn | =1 if individual was the firstborn child in the family | 0.291 | 0.336 |
| Nr older siblings | Number of older siblings | 1.444 | 1.148 |
| Nr siblings | Number of siblings | 2.068 | 1.595 |
| Number of observations | | 2,546 | 1,180 |

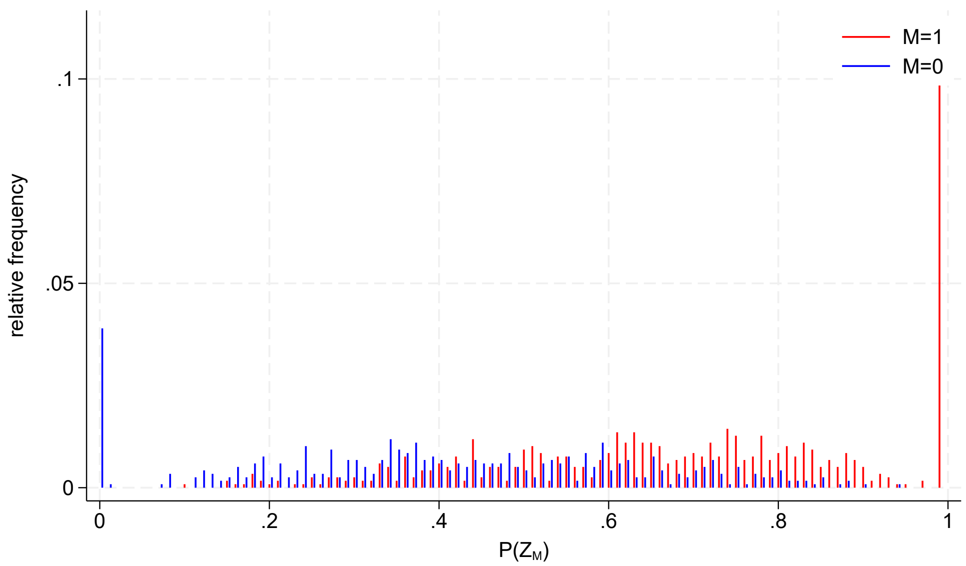
Notes: Own calculations based on NEPS-SC6 data The table shows means of the variables for individuals with and without the academic-track degree (Abitur).

Table A.2: Regression results: predicting $Z_M$ and $\tilde{Z}_M$ with $Z_D$ conditional on covariates

| | Dependent variable | |
|---|---|---|
| | $Z_M$ | $Z_M$ |
| $Z_D$ | $-0.0001$ | 0.0034 |
| | (0.0028) | (0.0056) |
| Observations | 3, 726 | 1,180 |

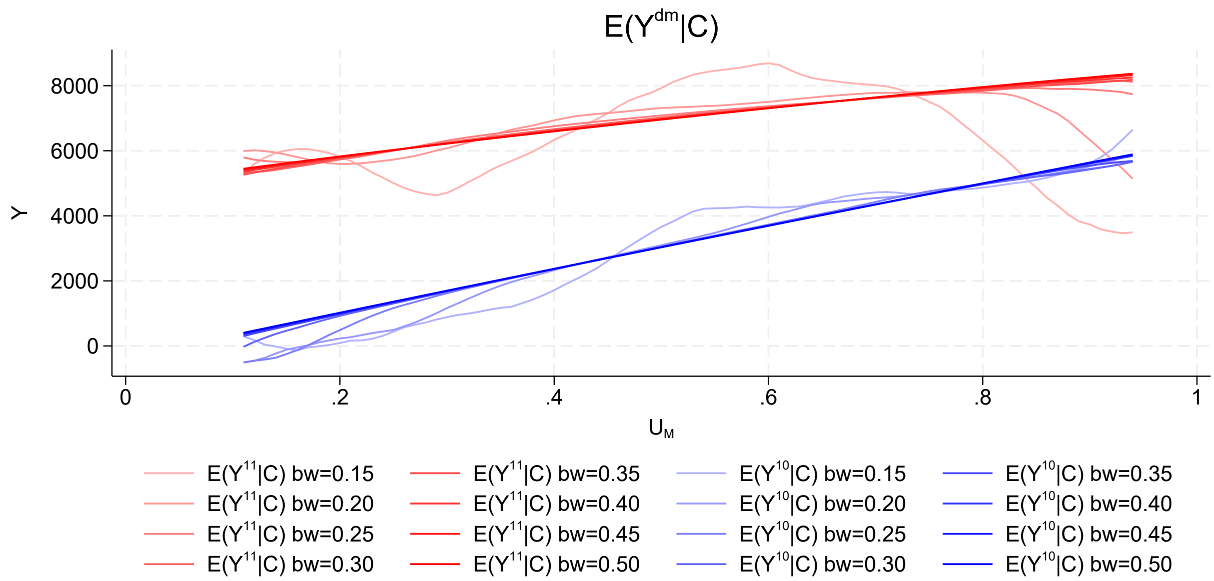
*Notes:* Own calculations based on NEPS-SC6 data. Included control variables are Abitur (only in column 1), district and entry cohort fixed effects, district-specific trends, dummies for father's degree, and distance (in 10 km steps) to the next academic-track school from the municipality before 1940. The sample in column 2 is restricted to individuals with $D = 1$. Standard errors in parentheses clustered on district level. * ($p < 0.1$), ** ($p < 0.05$), *** ($p < 0.01$)

Figure A.1: Support of $P(Z_M)$ by college education ($M$)




*Notes:* Own calculations based on NEPS-SC6 data. The graph plots the relative frequency of $P(Z_M)$ values in 0.01 bins of the propensity score by college degree for individuals with Abitur ($D = 1$).

Figure A.2: MTR functions for compliers to academic-track school access for different bandwidths




*Notes:* Own calculations based on NEPS-SC6 data. Number of observations is 3,726. Included control variables are district and entry cohort fixed effects, district-specific trends, dummies for fathers' degrees, and distance (in 10 km increments) to the next academic-track school from the municipality before 1940.

# B Details on the causal mediation analysis

## B.1 Conditional expectation and marginal treatment response functions

The conditional expectation functions are part of the concept of marginal treatment response functions (MTR) by Carneiro and Lee (2009), which are used to estimate potential outcome distributions normalized to the unit interval. The notation is based on a generalized Roy model. In the model, individuals choose $M = 1$ if their expected benefits are at least as high as their expected cost. Benefits and costs might vary with initial treatment assignment $D$. Hence, we write individual benefits as $Y^{d1} - Y^{d0} = (\mu^{d1} - \mu^{d0}) - (U^{d1} - U^{d0})$ and costs as $C^d(Z_M) = \mu^d_C(Z_M) + v^d$. Both can be separated into an observed part (indicated by the $\mu$s) and an unobserved part ($U$s and $v$). By definition (under valid assumptions), the instrument $Z_M$ only affects the choice of $M$ by its impact on the costs and has no direct effect on the potential outcomes. More formally, individuals choose $M = 1$ if their net benefit is positive:

$$
\begin{aligned}
M^d &= \mathbb{1}[Y^{d1} - Y^{d0} \geq C^d(Z_M)] \\
&= \mathbb{1}[\mu^{d1} - \mu^{d0} - \mu^d_C(Z_M) \geq U^{d1} - U^{d0} - v^d] \\
&= \mathbb{1}[\mu^d_m(Z_M) \geq V^d] \\
&= \mathbb{1}[F_V(\mu^d_M(Z_M)) \geq F_V(V^d)] \\
&= \mathbb{1}[Pr(M = 1|Z_M] \geq U^d_M] \qquad \forall j \in \{0, 1\}
\end{aligned} \tag{3}
$$

All observed components are brought on the left side, while all unobserved components are on the right side. For notational reasons, each side was simplified to one short expression, i.e., $\mu^d_m(Z_M) = \mu^{d1} - \mu^{d0} - \mu^d_C(Z_M)$ and $V^d = U^{d1} - U^{d0} - v^d$. Finally, we apply the cumulative distribution function $F_V(\cdot)$ (a rank-preserving monotonic transformation) to both sides. This yields the probability for a college degree based on observables (propensity score $p$) on the left side and the unobserved value $U^d_M$ (quantiles of $V$) on the right side of the inequality. The latter can be interpreted as the unobserved resistance to college education. For every higher value of $F_V(\mu^d_M(Z_M))$ induced by $Z_M$, more people select into college education. Individuals who change $M$ due to a marginal change in $Z_M$ are indifferent, and have $U^d_M = p$. The marginal individuals reveal their distastes for $M$, and we can evaluate their potential outcomes at any observed value of the propensity score or, equivalently, at any quantile of their resistance to college education. This is captured in the MTR functions, i.e. the derivatives of the conditional expectation functions with respect to $p$, which are defined as

$$
m_{dm}(p)' = E(Y^{dm}|U_M = p). \tag{4}
$$

Note that covariates are excluded here for simplicity. In section B.2, we show how to estimate MTRs for the possible combinations of $Z_D$, $D$, and $M$ conditional on $X$ and how to derive the complier-specific (to $Z_D$) conditional expectation functions from the results.

## B.2 Estimation

We start by showing how to estimate the TTE using our approach. This, of course, can be estimated by two-stage least squares, but it helps to get the intuition behind the following steps. Up to this point, we left the inclusion of covariates to the approach implicit. From now on, we explicitly include the vector of conditioning covariates $X$ in our notation. First, we need to estimate the share of compliers to $Z_D$. Therefore, we estimate the following first-stage regression

$$D = \pi_0 + \pi_1 Z_D + \tilde{X}'\pi_X + u \tag{5}$$

Here, covariates enter the equations as mean-centered values indicated by $\tilde{X}$. The advantage is that the intercept directly yields the respective outcome mean, which eases notation.[15] Thus, $\pi_0$ directly gives us the share of always takers ($AT_D$). The share of compliers who react to the instrument $Z_D$ regarding treatment choice $D$ is given by $\pi_1$. Lastly, the share of never takes is given by $1 - \pi_0 - \pi_1$.

We then obtain outcome means for each possible combination of $D$ and $Z_D$ by estimating

$$\begin{aligned} Y = {} & \delta_1[\mathbb{1}(D=1)(Z_D=1)] + \delta_2[\mathbb{1}(D=1)(Z_D=0)] \\ & + \delta_3[\mathbb{1}(D=0)(Z_D=0)] + \delta_4[\mathbb{1}(D=0)(Z_D=1)] + \tilde{X}'\delta_X + v \end{aligned} \tag{6}$$

Under valid assumptions, we know the groups that potentially contribute to every $\delta$: $\delta_1$ gives the mean for always-takers and treated compliers (with $D = 1$ and $Z_D = 1$ it is ambiguous if the individual is always taker or treated complier), $\delta_2$ the mean for a subgroup of pure always-takers, $\delta_3$ – vice versa – the mean for never takers and untreated compliers, and $\delta_4$ the mean of never-takers. Knowing the shares and outcome means we can apply the Imbens and Rubin (1997) formula to get the complier-specific potential outcomes:

$$E(Y^1|C_D) = \frac{\delta_1(\pi_0 + \pi_1) - \delta_2\pi_0}{\pi_1} \tag{7}$$

$$E(Y^0|C_D) = \frac{\delta_3(1 - \pi_0) - \delta_4(1 - \pi_0 - \pi_1)}{\pi_1} \tag{8}$$

[15] Without demeaning the covariates in a previous step, one can derive the always-taker share by adding each covariate's coefficient multiplied by its mean to the constant.

The difference between them yields the TTE, i.e., the complier-specific effect of $D$ and $Y$ (LATE).

Now, we are turning to the setting-specific estimation of the conditional expectation functions. It is an adjusted version of the MTR definition used in the separate approach to estimate marginal treatment effects (see Carneiro and Lee, 2009; Brinch et al., 2017). We define the MTR for different values of $Z_D$, $D$, and $M$ to derive the complier-specific (to $Z_D$) MTRs afterwards. This typically results in 8 different conditional expectation functions: 4 different subpopulations (equivalent to the $\delta$s in equations (6)) for $M = 1$ and $M = 0$ each.

We can express the MTR functions at a specific value of the propensity score $P(Z_M) = Pr(M = 1|Z_M, X) = p$ as

$$\begin{aligned} m_{111}(p)' =& E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 1, M = 1) \\ &+ p\frac{\partial E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 1, M = 1)}{\partial p} \end{aligned} \tag{9}$$

$$\begin{aligned} m_{101}(p)' =& E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 0, M = 1) \\ &+ p\frac{\partial E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 0, M = 1)}{\partial p} \end{aligned} \tag{10}$$

$$\begin{aligned} m_{110}(p)' =& E(Y|X = x, P(Z_M) = p, D = 1, M = 0, Z_D = 1) \\ &- (1-p)\frac{\partial E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 1, M = 0)}{\partial p} \end{aligned} \tag{11}$$

$$\begin{aligned} m_{100}(p)' =& E(Y|X = x, P(Z_M) = p, D = 1, Z_D = 0, M = 0) \\ &- (1-p)\frac{\partial E(Y|X = x, P(Z_M) = p, D = 1, M = 0, Z_D = 0)}{\partial p} \end{aligned} \tag{12}$$

for the cases with $D = 1$.[16] For $D = 0$, $m_{01m}(p)'$ and $m_{00m}(p)'$ are generally defined analogously. But, in our setting, we cannot identify the conditional expectations for any combination with $D = 0$ along $U_M$ as there is no variation in $M^0$. Still, with the MTRs given by equations 9 - 12 the $ITE(1)$ and the $DTE(0)$ can be derived.

For the expected values (and the derivatives with respect to $p$) entering the MTR functions, we begin by running a regression of the following form

$$\begin{aligned} Y =& \alpha_{111}[\mathbb{1}(D = 1)(Z_D = 1)(M = 1)] + \alpha_{101}[\mathbb{1}(D = 1)(Z_D = 0)(M = 1)] \\ &+ \alpha_{110}[\mathbb{1}(D = 1)(Z_D = 1)(M = 0)] + \alpha_{100}[\mathbb{1}(D = 1)(Z_D = 0)(M = 0)] \\ &+ \alpha_{010}[\mathbb{1}(D = 0)(Z_D = 1)(M = 0)] + \alpha_{000}[\mathbb{1}(D = 0)(Z_D = 0)(M = 0)] \\ &+ \beta_{111}[\mathbb{1}(D = 1)(Z_D = 1)(M = 1)] \times p + \beta_{101}[\mathbb{1}(D = 1)(Z_D = 0)(M = 1)] \times p \\ &+ \beta_{110}[\mathbb{1}(D = 1)(Z_D = 1)(M = 0)] \times p + \beta_{100}[\mathbb{1}(D = 1)(Z_D = 0)(M = 0)] \times p \\ &+ \tilde{X}'\delta_X + v \end{aligned}$$

[16] Note that the MTR equations differ for $m = 1$ and $m = 0$ in how the slopes enter.

by OLS. The $\alpha_{dz_d m}$s indicate levels and $\beta_{dz_d m}$s slopes at specific values of the propensity score for every combination of $D$, $Z_D$, and $M$. For the linear MTRs, they can directly be plugged into equations 9 to 12.[17] For the semiparametric MTR functions, we follow Schmitz and Westphal (2021) by first determining $\tilde{Y} = Y - (X - \bar{X})'\hat{\delta}_X$ after running the linear model above to remove average variation in $Y$ induced by control variables. Then we regress the purged outcome variable $\tilde{Y}$ nonparametrically on the propensity score.

As we ultimately need the group-specific MTRs of $Z_D$-compliers, we apply the principle of Imbens and Rubin (1997) from equations 7 and 8 to the MTRs instead of outcome means. Knowing the shares from the first stage regression (equation 5), we calculate

$$m_{11}^C(p)' = \frac{m_{111}(p)'(\pi_0 + \pi_1) - m_{101}(p)'\pi_0}{\pi_1}$$
$$m_{10}^C(p)' = \frac{m_{110}(p)'(\pi_0 + \pi_1) - m_{100}(p)'\pi_0}{\pi_1}$$

where $C = C_D$ indicates compliers to $Z_D$. The analogue for $D = 1$ cannot be derived as $m_{011}(p)'$, $m_{001}(p)'$, $m_{010}(p)'$ and $m_{000}(p)'$ are not identified in our setting (see above).

Finally, we need to determine $Y^{dM^l}$. Based on the hypothetical observation rule $Y^{dM^l} = Y^{d1}M^d + Y^{d0}(1 - M^d)$, $E(Y^{dM^l}|U_M, C_D)$ is a weighted mean of $m_{d1}^C(p)'$ and $m_{d0}^C(p)'$ with weights of $E(M^l|U_M, C_D)$. For determining the weights, we can calculate $E(\mathbb{1}[Pr(M = 1|Z_M) \geq U_M^d]|U_M^d)$ for each combination of $D$ and $Z_D$, that is, the fraction of estimated propensity scores larger than the evaluation point $U_M^d = u_M^d$. Without individuals with $M = 1$ and $D = 0$, as in our case, $M^0$ and the corresponding weights along $U_M$ are zero. Applying the Imbens and Rubin (1997) formula once more, yields the conditional expectation $M^1$ for the compliers (along $U_M$).

For each indirect and direct effect reported throughout the paper, to integrate out unobserved heterogeneity, we average the conditional expectations across $U_M$ only for values within the common support interval ($U_M \in (0.107, 0.943)$).

[17] As we do not have full support in our setting (see Figure A.1) and we aim to estimate the conditional expectation functions only within the region of common support, propensity scores falling outside this interval are set to their respective minimum or maximum values. In addition, we account for level differences in $Y$ on either side of this threshold by including dummy variables indicating upward and downward truncation of the propensity score, interacted with all possible combinations of $D$, $Z_D$ and $M$, as additional control variables.