UNIVERSITÄT
TRIER

From education to employment: Understanding educational mismatch using the case of Germany

Vom Fachbereich IV der Universität Trier zur Verleihung des akademischen Grades Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.) genehmigte Dissertation

von

Theresa Regina Geißler

Trier, 2025

Betreuer*in: Prof. Dr. Laszlo Goerke
1. Berichterstatter*in: Prof. Dr. Laszlo Goerke
2. Berichterstatter*in: Prof. Dr. Katrin Muehlfeld
Datum der Disputation: 03.12.2025

Abstract

Many developed countries, including Germany, face a steady rise in the share of individuals obtaining higher education. While rising education itself bears a series of advantages as extensively studied in previous literature, it is also conceptually linked to a higher likelihood of working in an occupation that does not match one's formal qualifications. Previous studies have predominantly evaluated how demographic or job-related aspects correlate with the likelihood of being educationally (mis)matched. However, they have largely ignored institutional facets of the educational system or industrial organization. Moreover, little is known about how private wealth affects educational mismatch or whether job satisfaction is homogenously affected among individuals once such a mismatch occurs. The five projects collected in this thesis aim to answer these open questions in the literature for Germany, using data from the Socio-Economic Panel and employing different time intervals between 1984 and 2022.

Beginning with the educational system in early childhood, Chapter 2 evaluates the impact of school-starting age on the likelihood of over- and undereducation. It exploits the exogenous variation in school-entry rules across federal states and years in Germany with regression discontinuity designs. The results report a negative impact of school-starting age on the likelihood of undereducation, but no systematic relationship with overeducation.

Subsequently, Chapter 3 explores the variation in education costs by leveraging the quasi-experimental setting induced by the time-limited introduction of tuition fees in several German federal states between 2006 and 2014. The increase in education costs among treated graduates results in a significantly higher likelihood of overeducation, which endures even several years post-graduation.

Chapter 4 focuses on the industrial relations system and examines the correlation between trade union membership and the likelihood and extent of educational (mis)match. The results reveal that trade union members report significantly less overeducation at both the intensive and extensive margin and also a higher likelihood of being matched compared to non-members.

Furthermore, the heterogeneity analysis provides evidence that this correlation is driven by improved bargaining power instead of informational advantages.

Chapter 5 focuses on private wealth as a determinant of educational mismatch by investigating the impact of a wealth shock through inheritances, lottery winnings or gifts on the likelihood of over- and undereducation. Due to the diminishing marginal returns of wages with increasing windfall gains the likelihood of undereducation is expected to decrease, while that of overeducation is expected to increase. Empirically, these suppositions are supported for overeducation, as its likelihood increases significantly after the windfall gain. Further analyses reveal that this effect is driven by individuals switching occupations while increasing their leisure time, and it materializes only for medium to large windfall gains.

Contrary to the previous chapters, Chapter 6 focuses on educational mismatch, more precisely on overeducation, as the independent variable. In particular, it investigates the correlation between overeducation and job satisfaction. The results align with the previously established negative correlation for private sector employees exclusively. In contrast, interaction and subsample analyses reveal a positive correlation for public sector employees. This link is driven by individuals with a high degree of altruistic motivation and family orientation.

Deutsche Kurzfassung (German abstract)

Der Anteil der Personen, die eine höhere Bildung erwerben, steigt in vielen Industrieländern, unter anderem in Deutschland, stetig an. Ein höherer Bildungsgrad bringt eine Reihe von Vorteilen mit sich, die in der bisherigen Literatur bereits ausführlich untersucht wurden. Allerdings ist eine höhere Bildung konzeptionell auch mit einer höheren Wahrscheinlichkeit verbunden, in einem Beruf zu arbeiten, der nicht den formalen Qualifikationen entspricht. In früheren Studien wurde vor allem untersucht, wie demografische oder berufsbezogene Aspekte mit der Wahrscheinlichkeit korrelieren, in einer solchen Position zu arbeiten. Institutionelle Aspekte des Bildungssystems oder der Arbeitsbeziehungen wurden hingegen weitgehend außer Acht gelassen. Außerdem ist wenig darüber bekannt, wie sich der private Wohlstand auf die Bildungsinkongruenz auswirkt oder ob die Arbeitszufriedenheit von Personen, bei denen eine solche Inkongruenz auftritt, einheitlich beeinflusst wird. Die fünf in dieser Arbeit gesammelten Projekte zielen darauf ab, diese offenen Fragen in der Literatur für Deutschland zu beantworten, indem sie verschiedene Zeitintervalle der Daten des Sozio-Ökonomischen Panels zwischen 1984 und 2022 nutzen.

Beginnend mit dem Bildungssystem in der frühen Kindheit, evaluiert Kapitel 2 die Auswirkungen des Einschulungsalters auf die Wahrscheinlichkeit von Über- und Unterbildung. Dabei wird die exogene Variation der Einschulungsregeln zwischen den Bundesländern und Jahren in Deutschland genutzt. Die Ergebnisse von Regressions-Diskontinuitäts-Analysen zeigen einen negativen Einfluss des Einschulungsalters auf die Wahrscheinlichkeit der Unterbildung, aber keine systematische Beziehung zur Überbildung.

Anschließend untersucht Kapitel 3 die Variation der Bildungskosten. Das Kapitel nutzt ein Quasi-Experiment, das durch die zeitlich begrenzte Einführung von Studiengebühren in mehreren deutschen Bundesländern zwischen 2006 und 2014 entstanden ist. Der Anstieg der Bildungskosten bei den betroffenen Absolventinnen und Absolventen führt zu einer signifikant höheren Wahrscheinlichkeit von Überbildung, die auch noch mehrere Jahre nach dem Studienabschluss anhält.

Kapitel 4 widmet sich dem System der Arbeitsbeziehungen und untersucht den Zusammenhang zwischen der Mitgliedschaft in einer Gewerkschaft und der Wahrscheinlichkeit und dem Ausmaß der Bildungsinkongruenz. Die Ergebnisse zeigen, dass Gewerkschaftsmitglieder sowohl ein geringeres Ausmaß als auch eine geringere Wahrscheinlichkeit der Überbildung aufweisen. Zudem arbeiten sie mit einer höheren Wahrscheinlichkeit in adäquaten Berufen. Eine Heterogenitätsanalyse weist darauf hin, dass diese Korrelation auf eine verbesserte Verhandlungsmacht und nicht auf Informationsvorteile zurückzuführen ist.

Kapitel 5 konzentriert sich auf das Privatvermögen als Determinante von Bildungsinkongruenz. Es untersucht die Auswirkungen eines Vermögensanstiegs durch Erbschaften, Lotteriegewinne oder Schenkungen auf die Wahrscheinlichkeit von Über- und Unterbildung. Aus konzeptioneller Sicht wird erwartet, dass die Wahrscheinlichkeit einer Unterbildung abnimmt, während die Wahrscheinlichkeit einer Überbildung aufgrund der abnehmenden Grenzerträge der Löhne mit steigenden Zufallsgewinnen zunehmen dürfte. Empirisch werden diese Annahmen für die Überbildung bestätigt, da ihre Wahrscheinlichkeit nach dem Zufallsgewinn deutlich zunimmt. Weitere Analysen zeigen, dass dieser Effekt teilweise darauf zurückzuführen ist, dass Personen den Beruf wechseln und gleichzeitig ihre Freizeit erhöhen. Zudem kommt der Effekt nur bei mittleren und großen Zufallsgewinnen zum Tragen.

Im Gegensatz zu den vorangegangenen Kapiteln konzentriert sich Kapitel 6 auf die Überbildung als unabhängige Variable. Die zuvor etablierte negative Korrelation zwischen Überbildung und Arbeitszufriedenheit kann dabei nur für Beschäftigte des privaten Sektors festgestellt werden. Dahingegen zeigen Interaktions- und Teilstichprobenanalysen eine positive Korrelation für Beschäftigte des öffentlichen Sektors. Dieser positive Zusammenhang wird durch Beschäftigte mit einem hohen Maß an altruistischer Motivation und Familienorientierung getrieben.

Acknowledgements

Reaching an important milestone does not always go smoothly, but can also be associated with difficulties and doubts. Overcoming these is often not the achievement of a single person, but of supervisors, colleagues, family members, and friends. For this reason, I would like to express my deepest gratitude to those who have accompanied me on my path to my doctorate and have helped me not only to “get through” but to grow with every step on the way to completing this thesis.

First and foremost, I would like to thank my doctoral supervisor, Professor Dr. Laszlo Goerke. He gave me the opportunity to write my dissertation at his chair and the institute and supported my research goals from the very beginning. Moreover, I highly value all the opportunities he offered me, may it be in terms of conference visits, attending courses or teaching courses myself. In particular, his detailed guidance on writing scientific papers and the provision of opportunities for internal discussion about ongoing projects have contributed massively to the qualitative development of the dissertation projects.

Secondly, I would like to thank my second supervisor, Professor Dr. Katrin Muehlfeld. By attending her courses as a Master’s student, I was encouraged to work as a student assistant and learned how to conduct research from the very beginning. During this time, I gained my first experience of academic work and research and she supported me in continuing along this path until I completed my doctorate. I am very sure that without her support I would never have ended up in the academic world.

Even though the academic path is, of course, highly influenced by the professorship, colleagues are a source of support that should not be underestimated. I extend my heartfelt thanks to my friends Björn and Marco. Both started their PhD journeys just a short time before me, so that we could share our experiences. In addition, Marco was one of my co-authors and taught me basically everything I know about dynamic DiD estimations. I would also like to thank Sven for being my co-author and for always having an open ear for me as my office neighbour. Apart from these three, I am grateful for Yue, Jonas, Fenet, Alberto, Adam, Nora, and Sumit who shared their expertise with me. I am also very thankful for my legal colleagues. I would especially like to thank Jana and Georg for their mental support, lunches together, and for being good colleagues and friends. I was also supported by our secretaries and our library team, who knew the answer to every possible question. I would particularly like to express my gratitude to Ms. Mielke, as the first point of contact at the institute, and Ms. Kuhn, who always supported me, especially when our secretariat was not occupied. Furthermore, I would like to thank the student assistants for their valuable support in my research and for their help in organizing conferences.

In addition to the people mentioned above, I am deeply indebted to my family and friends who have supported me over the years. There are no words to describe how grateful I am to have them in my life. Especially, I owe a great debt of gratitude to my parents and sisters. Even as a young child, they saw potential in me that I had not even dared to dream of and always encouraged me to go my own way. Without their love, encouragement and comforting words in times of doubt, I would not be the person I am today. Last but not least, I would like to thank Tim: Without you, this work would never have been finished and I would have given up halfway through. Thank you for always reminding me of my abilities, catching my doubts and always pointing out the important areas of life when I lose track.

Trier, July 2025

Theresa R. Geißler

Contents

List of figures

List of figures in appendices

List of tables

List of tables in appendices

Chapter 1

Introduction

On average, the level of education is increasing worldwide. Numbers provided by the OECD reveal that 23.56% of those aged 25 to 34 in 1998 held a tertiary degree, while the share approximately doubled to 47.42% in 2022 (OECD, nd). Similar patterns can be observed in Germany (Statistisches Bundesamt, 2020b). The share of individuals aged above 15 with tertiary degrees, including Diplom, Bachelor's, Master's and PhD degrees, amounted to 14.5% in 2011 and increased by more than a quarter to 18.5% in 2019. In contrast, the proportion of individuals with an apprenticeship or vocational training[1] declined by about 8% in the same period (Statistisches Bundesamt, 2020b). Moreover, relative to the total population, the share of enrolled students increased by approximately 55% between 1998 and 2023 (Statistisches Bundesamt, 2024a,c).[2] This development is at odds with the relatively stable demand for university degrees in job advertisements. Figures from the Institut für Arbeitsmarkt- und Berufsforschung (2024) indicate that the share of vacancies requiring a university degree showed only minor fluctuations and stayed close to 20% throughout 2010 to 2024, while around 60% consistently required vocational training or apprenticeships.

Previous conceptualisations predict higher rates of mismatches between individuals' qualifications and the job requirements as the number of tertiary graduates rises (Charlot and Decreuse, 2005; Dolado et al., 2000; Lazear et al., 2018). In particular, they anticipate that only the best-qualified individual will be selected for the open position, aligning with Signalling and Job Competition Theory (Spence, 1973; Thurow, 1975). An environment with an oversupply of

[1]Apprenticeships and vocational training are obtained in the German double-tier system. These typically last for two to three years. Individuals are employed by a company and are trained on-the-job while they simultaneously visit vocational schools (*Berufsschulen*, Federal Institute for Vocational Education and Training, 2020).

[2]In 1998, the total population amounted to 82,037,011 and the number of students was 1,800,651 across all federal states, representing 2.19%. In 2023, the total population rose to 84,669,326 while student numbers were 2,868,311, equalling 3.39%.

highly educated individuals and a relatively stable number of vacant positions for tertiary graduates (Institut für Arbeitsmarkt- und Berufsforschung, 2024) implies that all positions for those applicants are eventually filled. Individuals who are not selected for one of these jobs but are still highly educated must then apply for jobs with lower educational requirements, inducing the so-called dropping effect and overeducation (Lazear et al., 2018).

As a form of underemployment, overeducation hence refers to a situation in which individuals' formal qualifications exceed the job requirements and represents one type of educational mismatch (e.g., Feldman, 1996; McKee-Ryan and Harvey, 2011). The opposite case, namely undereducation, describes the state in which individuals' formal qualifications fall short of the job requirements (e.g., Duncan and Hoffman, 1981; McGuinness, 2006). When studying these forms of vertical educational mismatch, which will be the focus of this thesis, it is also crucial to distinguish related but distinct concepts. Diverging from educational mismatch, skills mismatch considers the match between individuals' actual skills and those necessary to perform the job (e.g., McGuinness et al., 2018; Robst, 1995b). Moreover, horizontal educational mismatch differs from vertical educational mismatch as it assesses whether individuals work in jobs outside their field of study or training (Robst, 2007). Focusing on either of those, aside from vertical educational mismatch, is beyond the scope of this thesis, particularly due to difficulties involved in accurately assessing both.

While higher education itself yields diverse benefits for individuals and society, such as increasing civic engagement (e.g., Campbell, 2009; Larreguy and Marshall, 2017), higher wages (e.g., Dickson, 2013) or better health (e.g., Brunello et al., 2016; Ross and Wu, 1995), educational mismatch, particularly overeducation, has often been associated with detrimental outcomes. For example, overeducated individuals face a wage penalty compared to equally qualified individuals working in adequate jobs (e.g., Allen and van der Velden, 2001; Duncan and Hoffman, 1981). Moreover, previous evidence points towards lower job satisfaction rates and deteriorated health outcomes among overeducated individuals (Bracke et al., 2013; De Santis et al., 2021; Fleming and Kler, 2008, 2014; Peiró et al., 2010). Still, the repercussions are not limited to the individual level. Inadequate matches between job requirements and employees' education have, for example, also been linked to (firm) productivity (Büchel, 2002; Jacobs et al., 2023; Plesca and Summerfield, 2023; Tsang, 1987). For this reason, it is of great relevance to understand why people work in positions for which they are not adequately educated and what consequences they face. The aim of this thesis is therefore to expand our knowledge of the factors associated with the likelihood of educational mismatch and its repercussions. The studies in this thesis focus on Germany, which presents an interesting case as education, with few exceptions, is generally publicly funded. This implies that an efficient distribution of the labour force is not only important for individuals and firms, but also for the public and for policy makers. In addition, Germany offers a unique educational landscape through its two-tier education system, higher education institutions, and the links between both systems.

All studies gathered in this thesis rely on data from the Socio-Economic Panel (SOEP) provided by the German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung, DIW). The SOEP is a panel dataset that surveys approximately 30,000 individuals living in Germany every year. Information is provided on all areas of individuals' lives encompassing employment relations, education, household, and family life, as well as individual characteristics, including, for example, personality traits, preferences, and habits (DIW, 2022; Goebel et al., 2019). Importantly, the SOEP data allow for three different assessments of educational mismatch that have previously been used in the literature. First, the dataset encompasses information on the educational attainment individuals think is needed for their current job through the question: "What kind of training is usually required for this type of work?" Comparing the answers to this question with a categorical assessment of educational attainment allows for the identification of overeducation, undereducation, and being matched as proposed in Duncan and Hoffman (1981). This measure is referred to as the indirect self-assessment (ISA). Second, the SOEP provides information on individuals' occupations using the International Standard Classification of Occupations (ISCO, International Labour Office, 2012). This classification allows for the differentiation of occupations at the one- to four-digit level. In the broadest, one-digit classification, the ISCO distinguishes between ten occupations. These are assigned to four different skill levels. Matching this information with the corresponding skill levels defined in the International Standard Classification of Education (ISCED, UNESCO, 2006) allows to define educational mismatch as suggested in the job analyst approach (JA, Rumberger, 1981a,b). Finally, the more fine-grained assessment of occupations at the three- and four-digit ISCO level combined with precise information on the average years of education obtained by individuals enables the establishment of a statistical approach. This approach is also referred to as the mean or realised matches measure (MEAN, Clogg and Shockey, 1984; Verdugo and Verdugo, 1989). In this approach, the level of required education is assessed such that it equals the average level of education obtained by one's reference group. The latter is defined as the three-digit ISCO group in each survey year, as supposed by Blásquez and Budría (2012). Educational mismatch is then identified by comparing this measure to individuals' attained years of education. To avoid biases induced by years of education marginally exceeding or falling short of the reference value, it is suggested to define individuals as matched if their years of education fall within a standard deviation range around the average. Individuals are considered overeducated (undereducated) if their years of education exceed (fall short of) the mean by more than one standard deviation only.[3]

While each of the mentioned measures bears advantages (for a detailed discussion, see e.g., Capsada-Munsech, 2019; Chevalier, 2003; Verhaest and Omey, 2006, 2010), this thesis employs the statistical approach for three main reasons: First and foremost, the statistical measure makes it possible to adjust for the general increase in the level of education in society and in occupations over time. This is particularly important when working with the SOEP, as it covers more than four decades in which the average level of education rose sharply (cf., Statistisches Bundesamt, 2020b). Second, it enables a direct comparison of individuals to their peers and is, therefore, not a static definition of educational mismatch but allows for a relative assessment of individuals' positions within their occupations. This entails that, third, this measure depicts educational mismatch as a function of supply and demand mechanisms in the market, implying that its definition is based on real labour market matches instead of theoretical requirement evaluations (cf., Freeman, 1976; Hartog, 2000).[4]

Using these different assessments of educational mismatch, previous research has broadly investigated which individuals are particularly prone to be (mis)matched. These studies have focused particularly on demographic and job-related aspects, including, but not limited to gender, age, education, migration background, part-time employment, or tenure (e.g., Akgüç and Parasnis, 2023; Aleksynska and Tritah, 2013; Battu et al., 1999; Belfield, 2010; Blásquez and Budría, 2012; Caroleo and Pastore, 2018; Di Pietro and Cutillo, 2006; Diem, 2015; Diem and Wolter, 2014; Dolton and Silles, 2008; Fleming and Kler, 2008; Green and McIntosh, 2007; McGoldrick and Robst, 1996; Robst, 1995b; Santiago-Vela and Mergener, 2022; Turmo-Garuz and Bartual-Figueras, 2019; Verhaest and Omey, 2010). Still, certain links have not yet been studied in detail.

[3]Note that the measure for years of education in the SOEP does not correspond to the actual time spent in education, but is adjusted for the qualification that a person has obtained. In this way, factors such as repeating a grade or different school systems do not affect the variable.

[4]Note that a direct self-assessment (DSA) asking individuals whether they are overeducated (undereducated, matched) cannot be applied using the SOEP data as the necessary question does not form part of the survey.

These include aspects related to the educational system, such as regulations on the age at which children start school, and higher education costs, the impact of trade unions, or an individual's wealth. The major part of this thesis, including Chapter 2, Chapter 3, Chapter 4, and Chapter 5, aims at filling these gaps in the literature and explores these antecedents of educational mismatch in detail.

Chapter 2, which is joint work with Sven Hartmann, explores an early childhood intervention. Specifically, it evaluates the impact of school-starting age on the likelihood of over- and undereducation. Starting school older has been previously linked to better performance in school (Attar and Cohen-Zada, 2018; Balestra et al., 2020), which might ameliorate individuals' positions in the application process and reduce (increase) their likelihood of overeducation (undereducation) (Charlot and Decreuse, 2005; Dolado et al., 2000; Lazear et al., 2018; Spence, 1973). At the same time, individuals who start younger have a higher labour market experience compared to their peers, which has been linked to a lower likelihood of overeducation (Peiró et al., 2010; Sicherman and Galor, 1990). Simultaneously, their potentially worse educational performance could disadvantage them in the application process, what could place them at a higher (lower) likelihood of overeducation (undereducation). To uncover the relationship and in order to account for endogeneity in school-starting age, we leverage the exogenous variation in school-entry cutoffs across German federal states and over time. The data reveal non-compliance with the cutoff regulations, for example due to early or late school starters.[5] For this reason, we employ fuzzy regression discontinuity designs (RDD) to estimate the Local Average Treatment Effect (LATE) of school-starting age on the likelihood of educational mismatch. The empirical results reveal a negative effect of school-starting age on the likelihood of undereducation, while no systematic relationship is observed with the likelihood of overeducation. These results are not heterogeneous across sex, birth cohort, or region. A channel analysis provides suggestive evidence that the effect may be partially mediated by occupational choice and suppressed by educational attainment. However, the data structure does not allow for the identification of statistically significant mediations. Overall, we show that the age at which children start school has long-term effects, even on labour market outcomes that manifest more than a decade later.

Moving forward in the educational system, Chapter 3 evaluates how the costs of higher education affect the likelihood of overeducation. If sufficiently high,

[5] In Germany, early starters ("Kann-Kinder") are children who do not turn six until the predefined cutoff date, but perform exceptionally well in kindergarten. In such cases, children may start school earlier.

tuition fees could affect the likelihood of overeducation by pressuring individuals to enter the labour market faster after graduating. This could induce them to accept jobs below their qualifications at higher rates (e.g., Gervais and Ziebarth, 2019; Minicozzi, 2005; Molina and Rivadeneyra, 2021). Besides this, tuition fees could also influence individuals' academic performance through behavioural adjustments such as an acceleration of studies or increased working hours (e.g., Bietenbeck et al., 2023; Garibaldi et al., 2012), as well as through psychological mechanisms such as a deteriorated mental health (e.g., Bruffaerts et al., 2018; McCloud and Bann, 2019; Walsemann et al., 2015). This could in turn impede their transition to the labour market if employers use educational indicators as signals in the selection process (e.g., Spence, 1973). To shed light on this relationship, the analysis exploits the quasi-experimental setting that originated from the time-limited introduction of tuition fees in several German federal states between 2006 and 2014. The results indicate that graduates who were exposed to these tuition fees are more likely to be overeducated post-graduation. This holds robust across a variety of sensitivity checks aiming at the identification strategy as well as selection concerns. Suggestive evidence aligns with the expectations that the effect of tuition fees might be less pronounced among individuals in high-skill jobs and those treated for a shorter period. Moreover, the effect is significantly less pronounced among individuals with one parent possessing upper secondary education. Finally, using the panel structure of the SOEP allows me to show that individuals exposed to tuition fees experience a higher likelihood of overeducation even up to ten years after graduation. This underlines the long-lasting impact of this policy change undertaken in several federal states of Germany on post-graduation labour market outcomes of affected individuals.

Besides the educational system, other institutions may affect individuals' likelihood and extent of educational mismatch in adulthood. In particular, the role of labour market institutions has not yet been examined in detail. Laszlo Goerke and I therefore focus on the industrial relations system in Chapter 4 and examine the role of trade union membership. In contrast to other countries, collective bargaining and trade union membership are not directly intertwined in Germany. Although collective agreements legally apply to union members employed by covered firms, the firms also often pay the agreed wage to employees who are not union members themselves. Besides this, trade union members benefit from two major advantages, namely, a strengthened bargaining power and better information. The informational advantage may prompt individuals to leave situations of overeducation, as they could become aware of the consequences of educational mismatch or of alternative employment opportunities. Moreover, an improved bargaining power can be expected to raise individuals' chances of realising desired job changes. In consequence, we expect a negative (positive) correlation between union membership and the likelihood and extent of overeducation (undereducation) and a positive link with the likelihood of being adequately matched. Applying tobit and logit estimations to the SOEP data, we find that trade union membership is negatively correlated with the likelihood of overeducation, aligning with the expectations. Moreover, we observe a positive correlation with the likelihood of being matched, while no systematic link with undereducation is evident. By considering subgroups of individuals among whom trade union membership is particularly common, the evidence aligns with the bargaining power perspective. In more detail, we observe the linkage between trade union membership and the likelihood of overeducation and being matched, particularly among subgroups of individuals with higher union densities. As informational advantages should be independent of the share of members within a certain subgroup, the heterogeneous results reveal no evidence in line with the informational channel.

While the previous chapters thus evaluate the educational system and labour market institutions as potential antecedents of the likelihood or extent of educational mismatch, Chapter 5 focuses on individuals' wealth. Precisely, Marco Clemens and I analyse the effect of windfall gains on the likelihood of over- and undereducation. In this context, windfalls include inheritances, gifts, and lottery winnings, which, when received, reduce the benefit of wage as a component of earned income (Doorley and Pestel, 2020). As the expectations regarding the impact of windfall gains on the likelihood of over- and undereducation are *ex ante* unclear, we use a simplified utility function framework to derive our hypotheses. Due to the diminishing marginal utility of wages with rising wealth and a trade-off with leisure, we expect the likelihood of overeducation to increase after a windfall gain, while the reverse is expected for undereducation. Using event studies based predominantly on Sun and Abraham (2021), we find evidence aligning with the expectations for overeducation. Specifically, we observe an increase in the likelihood of overeducation after the receipt of a windfall. Moreover, this increase is driven by individuals switching occupations in the aftermath of a windfall, particularly within or out of high-skill occupations. We also find that these job changes result in larger amounts of leisure time for the switchers, as we observe decreasing working and overtime hours after the windfall. In line with the formalisation, these results provide the first empirical evidence that leisure time and an inferior job match quality are traded against each other with rising wealth. Finally, the formalisation suggests that this trade-off should come into effect only if the windfall gain reaches a certain threshold. Supporting this notion, the data reveal significant effects only for medium to large windfall gains. In summary, Chapter 5 presents evidence that overeducation may also present an active choice depending on the financial resources.

Finally, Chapter 6 looks at one outcome of overeducation and investigates the correlation with job satisfaction. Previous research has established a negative correlation between both concepts in most cases (e.g., Burris, 1983; Chuang and Liang, 2022; Fleming and Kler, 2008, 2014; Sánchez-Sánchez and McGuinness, 2015). But, as most of these studies focus on private sector employees or do not distinguish between sectors (e.g., De Santis et al., 2021; Green and Zhu, 2010; Hersch, 1991; McGuinness and Byrne, 2015; McGuinness and Sloane, 2011; Voces and Caínzos, 2021), it is unclear whether this negative relationship also applies to public sector employees. For example, motivational differences between public and private sector employees, as a more altruistic orientation, might affect individuals' evaluations of a job and their educational mismatch. I account for these potential differences by examining the relationship between overeducation and job satisfaction using a moderation framework which interacts overeducation and the employment sector. This approach reveals a negative correlation between overeducation and job satisfaction, which is significantly attenuated by public sector employment. Subsample analyses align with this and report a significant positive link between overeducation and job satisfaction among public sector employees. This positive relationship is most pronounced among individuals who potentially experience the best fit between their preferences and the working conditions offered in the public sector, as the results are primarily driven by individuals with high levels of altruistic motivation and family orientation.

Lastly, Chapter 7 concludes this thesis.

Chapter 2

School-starting age and educational mismatch*

We present the first empirical evidence on the impact of school-starting age on educational mismatch. Using data from the German Socio-Economic Panel, we exploit exogenous variation in school entry laws and apply a fuzzy regression discontinuity design to estimate causal effects. Our findings indicate that one additional month in school-starting age reduces the likelihood of undereducation by 1.1 percentage points, while no consistent effect emerges for overeducation. Initial evidence suggests that these effects might be driven by occupational choice, as younger school starters are more likely to work in high-skill occupations. Conversely, educational attainment might act as a suppressor, amplifying the reduction in undereducation while uncovering a latent positive link between school-starting age and overeducation. Our results underscore the complex role of institutional factors, such as school-entry policies, in shaping educational and labour market outcomes.

Keywords: School-starting age; School entry laws; Educational mismatch; Regression discontinuity design

*This chapter is joint work with Sven Hartmann and currently submitted. We are grateful for helpful comments by Daniela Sonedda and Laszlo Goerke as well as the participants of the III LABORatorio R. Revelli Workshop 2023, the 2024 meeting of the European Society of Population Economics (ESPE), the 15th International SOEP User Conference, the EEA-ESEM congress 2024, the 39th conference of the Italian Association of Labour Economics (AIEL) and the 2025 Annual Conference of the Royal Economic Society (RES). We thank Adrian Hofmeier for excellent research assistance.

2.1 Introduction

Education constitutes one of the largest public expenditures in modern economies. In 2022, the member states of the European Union collectively invested EUR 746 billion in education, representing 4.7% of their combined GDP, while the United States spent over 5% of its GDP on education (Eurostat, 2024; World Bank Group, 2024). This raises concerns among policymakers and the public about whether workers fully utilise the human capital acquired through education in the labour market, particularly in light of simultaneous labour shortages and persistent non-employment in many advanced economies. Given the significant welfare losses caused by the misallocation of education and jobs, researchers are increasingly examining educational mismatch, commonly referred to as over- or undereducation, which occurs when workers' educational attainment does not align with their job requirements (Acemoglu, 1999; Feldman, 1996; Rumberger, 1981a; Tsang, 1987). Currently, almost 40% of workers in OECD member states experience a mismatch between their qualifications and their job requirements (OECD, 2022c). Approximately half of these cases involve individuals who are overeducated for their roles. Nevertheless, numerous advanced economies have implemented policies to further expand the educational attainment of their workforce. This leads to higher government spending and inadvertently exacerbates the prevalence of educational mismatch, which in turn contributes to the simultaneous phenomena of labour market slack and shortages in these countries (United Nations Development Programme, 2024).

A growing body of research has examined the prevalence and labour market consequences of educational mismatch in recent years, focusing primarily on the relationship between educational mismatch and job mobility (e.g., Alba-Ramirez, 1993; Robst, 1995a; Romanov et al., 2017), wages (e.g., Agopsowicz et al., 2020; Artz and Welsch, 2021; Duncan and Hoffman, 1981; Jacobs et al., 2023; Korpi and Tåhlin, 2009), productivity (Büchel, 2002; Kampelmann and Rycx, 2012), or job satisfaction (e.g., Belfield, 2010; Fleming and Kler, 2008, 2014; Verhaest and Verhofstadt, 2016). In contrast to the consequences of educational mismatch, our understanding of its determinants is still quite limited. Research has centred mainly on individuals' socio-demographic characteristics (Aleksynska and Tritah, 2013; McGoldrick and Robst, 1996; Santiago-Vela and Mergener, 2022) but has largely neglected the role of features of the education system itself. Given the high public and private costs associated with educational mismatch, it is crucial to understand how institutional characteristics influence its prevalence.

One such factor is school-starting age, which is determined by enrolment cutoff dates. These regulations create systematic differences when children begin formal education, with potential long-term effects on their educational and occupational trajectories.

In this paper, we examine the causal effect of school-starting age on educational mismatch. Previous research has shown that the age at which children start school has far-reaching implications on their educational and labour market outcomes, with older school starters often achieving better academic outcomes (cf., Attar and Cohen-Zada, 2018; Balestra et al., 2020). This enhances their competitiveness for high-skill jobs but also increases their risk of overeducation if the labour market cannot accommodate their qualifications. Conversely, those starting school relatively young benefit from higher labour market experience compared to their age-mates, which has been shown to reduce overeducation (Peiró et al., 2010) in line with Career Mobility Theory (Sicherman and Galor, 1990). However, the fact that younger school starters tend to perform worse in school may hinder their job application process and place them at a higher risk of educational mismatch. Thus, the relationship between school-starting age and educational mismatch is *ex ante* unclear.

To identify a causal effect of school-starting age, we utilise temporal and regional variation in school enrolment dates across German states and use data from the German Socio-Economic Panel Study (SOEP). While we hardly find any effect regarding the extensive margin of overeducation, our estimates suggest that an additional month of age at school start reduces the likelihood of being undereducated later in professional life by 1.1 percentage points. These findings are robust to various sensitivity checks, including diverse methods to assess educational mismatch. Moreover, we provide suggestive evidence that the impact of school-starting age might be partially mediated (suppressed) by occupational choice (educational attainment).

With the present study, we are contributing to two specific strands of literature. The first examines the determinants of educational mismatch. In this regard, previous research has documented that macroeconomic features such as the sectoral composition of employment and labour demand affect the prevalence of educational mismatch (Davia et al., 2017; Tarvid, 2015). Most studies, however, emphasise individual-level characteristics and show that the likelihood of working in a job that does not align with the own educational attainment varies across socio-demographic factors such as gender (McGoldrick and Robst, 1996; Santiago-Vela and Mergener, 2022), migration background (Akgüç and Parasnis, 2023; Aleksynska and Tritah, 2013), and individuals' socioeconomic background (Ordine and Rose, 2009). Moreover, the risk of educational mismatch decreases once human capital endowments like tenure or on-the-job-training increase (Büchel and Pollmann-Schult, 2004) and furthermore depends on educational aspects such as college quality or the field of study (Büchel and Pollmann-Schult, 2004; Di Pietro and Cutillo, 2006; Liu et al., 2021; Ordine and Rose, 2009, 2011; Robst, 1995b). In the present paper, we provide the first evidence that institutional features of the education system, such as school entry cutoff dates, can affect the prevalence of educational mismatch.

The second strand of literature deals with the effect of school-starting age on labour market outcomes. In the last two decades, extensive research has been conducted evaluating the impact of school-starting age on students' health (Balestra et al., 2020; Dee and Sievertsen, 2018; Mühlenweg et al., 2012), non-cognitive skills (Cornelissen and Dustmann, 2019; Mühlenweg et al., 2012), academic performance (Dhuey et al., 2019; Dicks and Lancee, 2018; Fletcher and Kim, 2016; Ponzo and Scoppa, 2014; Puhani and Weber, 2008) and their educational path (Cook and Kang, 2016; Herdeiro et al., 2025; Li et al., 2022; Mühlenweg and Puhani, 2010; Ponzo and Scoppa, 2014). In contrast, only a few studies examine the long-term effects on labour market outcomes, focusing primarily on wages (Bedard and Dhuey, 2012; Matta et al., 2016; Oosterbeek et al., 2021) or the probability of employment (Dobkin and Ferreira, 2010), with conflicting findings. Several studies suggest that school-starting age leads to initial differences in wages and labour market participation, which gradually diminish over the life-cycle (Black et al., 2011; Fredriksson and Öckert, 2014; Görlitz et al., 2025). Overall, none of the previous studies have investigated the impact of school-starting age on the education-job-match quality.[6]

This study continues as follows: Section 2.2 introduces our empirical strategy, including the assessment of educational mismatch and school-starting age. In Section 2.3, we analyse the effect of school-starting age on educational mismatch and provide several robustness tests. While we evaluate potential channels of our results in Section 2.4, we summarise our findings in Section 2.5.

[6] A study by Fumarco et al. (2022) examines the impact of relative age – defined as the age difference between a student and their oldest classmate – on school-to-labour market transitions and educational match quality. Using Belgian SONAR data for the 1978 and 1980 birth cohorts, they find no effect of relative age on the likelihood of being adequately matched.

2.2 Data and empirical approach

To evaluate the impact of school-starting age on educational mismatch, we use data from the German SOEP (DIW, 2022; Goebel et al., 2019). The SOEP is a longitudinal and representative household survey conducted annually since 1984, offering comprehensive insights into individuals' education, work, family relations, and other aspects of life. This dataset enables the measurement of school-starting age during childhood and the assessment of educational mismatch in adulthood through multiple approaches. Furthermore, the inclusion of regional identifiers allows us to leverage the regional and temporal variation in school-entry regulations across Germany.

2.2.1 Educational mismatch

Educational mismatch is defined by comparing the level of education attained by individuals with the level of education required for their current job. While assessing an individual's educational attainment is straightforward (in terms of years of schooling), prior research established distinct ways to define the level of required education. One commonly used measure is the *realised matches measure* (MEAN), which estimates the required education level for a certain occupation by calculating the average years of schooling within a reference group (Clogg and Shockey, 1984; Verdugo and Verdugo, 1989). By following Blásquez and Budría (2012), we define the reference group based on the three-digit occupational categories of the International Standard Classification of Occupations (ISCO) provided by the International Labour Office (2012) for each survey year. We use the MEAN measure in the main specification as it captures educational mismatch through supply and demand mechanisms (Freeman, 1976; Hartog, 2000), and it easily adapts to changes over time (Capsada-Munsech, 2019; Clogg and Shockey, 1984), such as educational inflation.[7]

At the extensive margin, individuals are considered overeducated if their level of education exceeds the average education in their occupation group by more than one standard deviation σ_{ot} in a given survey year t. Conversely, individuals are undereducated if their education falls short of this reference value, as outlined

[7] Figure A.1 illustrates the average number of years of education over time, which increased from 10.5 years in 1984 to 12 years in 2020.

in Equation 2.1 and Equation 2.2.

$$OE_{ito} = 1 \text{ if } x_{it} > \overline{x_{ot}} + \sigma_{ot} \tag{2.1}$$

$$UE_{ito} = 1 \text{ if } x_{it} < \overline{x_{ot}} - \sigma_{ot} \tag{2.2}$$

Previous studies have shown that the choice of method to measure educational mismatch can strongly impact the results. Therefore, we employ two alternative measures that have commonly been used in related studies as a robustness check in Section 2.3.2. One is the indirect self-assessment (ISA), which is based on respondents' evaluation of the required education for their current position (Duncan and Hoffman, 1981). In addition, the job analyst measure (JA) assesses the educational attainment of an individual and the required education level in a certain job by using the official classification of occupations and their assigned skill levels (Rumberger, 1981a).

2.2.2 School-starting age

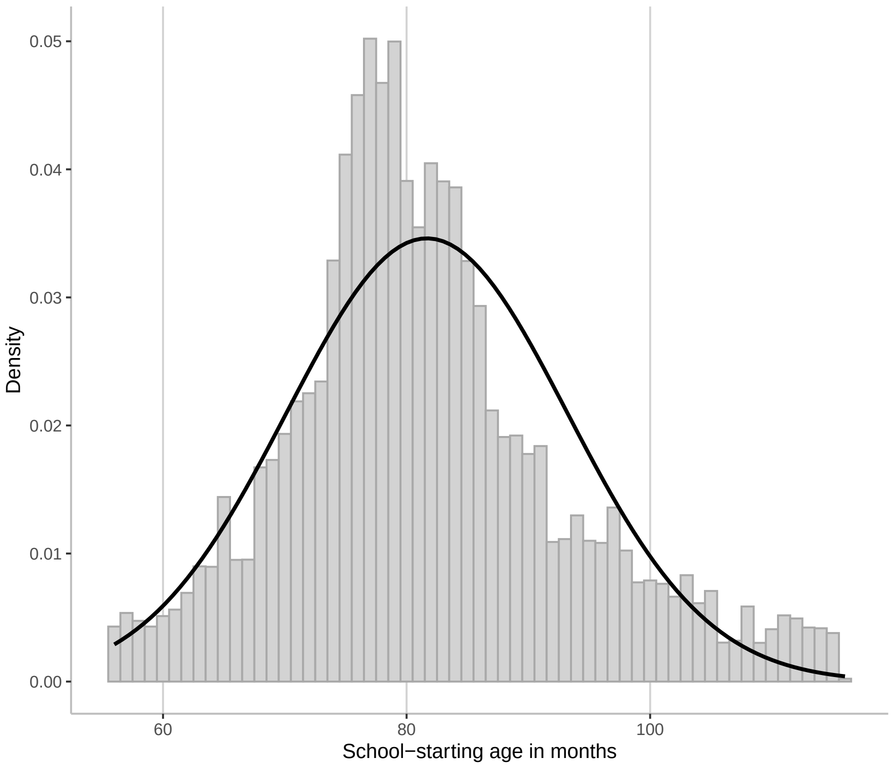
Information on school-starting age in months is provided by the SOEP for all individuals who participated in the survey while still being in school or attending other educational institutions, such as vocational schools. Since the survey contains detailed information on school-leaving degrees, graduation years, dates of birth, and places of residence, we can also determine the school-starting age of the remaining individuals (Bahrs and Schumann, 2020). In addition, we require information on the calendar month of school entry. Since 1990, the school year in West Germany (Federal Republic of Germany) has officially started on August 1st, while in East Germany (German Democratic Republic), it traditionally began on September 1st.[8] The distribution of school-starting age is shown in Figure 2.1. The distribution is mainly concentrated around the age of 80 months or 6.6 years and corresponds to the cutoff rule, which would lead to school-starting ages of roughly 6 to 7 years.

2.2.3 Endogeneity concerns and identification strategy

Previous studies suggest that parents may influence their child's school-starting age based on factors such as the development of their child, family resources, or

[8]Please note that these dates mark the official start of the school year, while the actual start of classes varies from state to state and year to year due to the timing of the summer vacation.

Figure 2.1: Histogram of school-starting age with normal distribution curve



Note: Figure 2.1 displays the distribution of the school-starting age in months. The sample is based on the survey years 1991 to 2020.

personal preferences (see, e.g., Angrist and Krueger, 1992; Bahrs and Schumann, 2020; Black et al., 2011; Fredriksson and Öckert, 2014; Görlitz et al., 2022; Mühlenweg, 2010; Mühlenweg and Puhani, 2010; Oosterbeek et al., 2021). These parental decisions introduce endogeneity, complicating the estimation of a causal relationship between school-starting age in months and educational mismatch.

To address this, we leverage German school-entry regulations, which require children who turn six before a specific cutoff date to begin school in the same year, while those who turn six after the cutoff date start the following year. These cutoff dates are set at the state level and have changed repeatedly over time (see Figure A.2).[9]

[9] In West Germany, we observe cutoffs in March, May, June, August, September, and December, with June being the most frequently used. In the German Democratic Republic, the cutoff date was always in May. These cutoff dates imply different ranges of school-starting ages for compliers, ranging between 5 years and 10 months and 7 years and 7 months, depending on the year of school-start and the federal state.

Using the cutoff dates and individuals' birth months, we define the running variable $dist_{isy} \in \mathbb{Z}$, capturing the distance to the cutoff, as:

$$dist_{isy} = birthmonth_i - cutmonth_{sy} \tag{2.3}$$

where i represents the individual, s is the state of school start[10], and y captures the year of school start.[11] Based on this, we can define the dichotomous variable $older_{isy}$:

$$older_{isy} = \begin{cases} 0 & \text{if } dist_{isy} \leq 0 \\ 1 & \text{if } dist_{isy} > 0 \end{cases} \tag{2.4}$$

This variable facilitates a parametric sharp RDD design, where children born after the cutoff are expected to start school the following year, while those born before comply with entry regulations and start that year (e.g., Angrist and Pischke, 2009; Cook and Kang, 2016; Oosterbeek et al., 2021).

Such a sharp RDD would allow the estimation of the causal effect of being born before versus after the cutoff, if individuals complied with the regulation and if there was no manipulation in the running variable. Although the McCrary-sorting test provided in Figure A.3 may graphically suggest no significant discontinuities in the running variable at the cutoff, the test statistics ($p = 0.000$) reject the null hypothesis of no apparent sorting. This raises concerns about potential manipulation around the cutoff and challenges the causal identification assumptions (for related research see e.g., Huang et al. (2020) and Kim (2021)). Furthermore, using the information on school-starting age reveals cases of non-compliance with the regulations, as illustrated in Figure A.4. To address potential violations of the sharp RDD assumptions and concerns about the endogeneity of school-starting age, we employ fuzzy RDD estimations (see, e.g., Anderson et al., 2011; Bahrs and Schumann, 2020; Dee and Sievertsen, 2018; Fredriksson and Öckert, 2014; Görlitz et al., 2022; Landersø et al., 2017; Mühlenweg and Puhani, 2010; Shin, 2023). We define a range around the cutoff of $\pm$ 4 months[12] and estimate 2SLS using the instrument $older_{isy}$ to account for the endogeneity in school-starting age ($months_i$). This approach yields the

[10]For individuals for whom the precise information on the state of school start is missing, we rely on the current state of residence.

[11]As the cutoff date is always the last day of the month in our sample, we can use full months to define the instrument.

[12]Larger windows around the cutoff improve the precision of our estimations due to a larger sample size. However, this may introduce bias if the selected bandwidth poorly approximates the functional form (for a discussion, see e.g. Cook and Kang, 2016). To ensure the robustness of our results, we implement a more restrictive interval width using a data-driven approach in Section 2.3.2.

local average treatment effect of school-starting age on educational mismatch, identified by those who adhere to the cutoff regulations. The first and second stage equations are shown in Equation 2.5 and Equation 2.6.[13]

$$months_i = \alpha_0 + \alpha_1 older_{isy} + \alpha_2 dist_{isy} + \alpha_3 X'_{it} + \lambda_t + \gamma_b + \delta_s + \epsilon_{it} \tag{2.5}$$

$$EM_{ito} = \beta_0 + \beta_1 \widehat{months}_i + \beta_2 dist_{isy} + \beta_3 X'_{it} + \lambda_t + \gamma_b + \delta_s + \epsilon_{it} \tag{2.6}$$

In the regressions, EM_{ito} is a dummy capturing either overeducation or undereducation of individual i in survey year t and occupation o. The coefficient of interest β_1 captures the effect of school-starting age on the likelihood of overeducation or undereducation. School-starting age is instrumented using a dummy variable $older_{isy}$ which takes the value of one if the individual should have started school in the year after their sixth birthday. The variable $dist_{isy}$ represents the distance to the respective cutoff, while X'_{it} includes a set of control covariates. Finally, λ_t, γ_b, and δ_s account for fixed effects related to the survey year, birth year, and state of school start, respectively, with ϵ_{it} representing the error term.

2.2.4 Covariates

In our main specification, we use covariates which are most likely already determined in childhood consistent with prior studies (e.g. Attar and Cohen-Zada, 2018; Bahrs and Schumann, 2020; Görlitz et al., 2022; Puhani and Weber, 2008).[14] We control for migration background and distinguish between direct or indirect migration. Natives form the reference group (Puhani and Weber, 2008). Sex is accounted for using a dummy variable equal to one for female respondents (Görlitz et al., 2022; McEwan and Shapiro, 2008). Furthermore, we incorporate information about the respondent's parents' educational background by including dummy variables that distinguish between upper secondary, intermediate, and general secondary school degrees. Not having completed school serves as the reference group (see e.g., Attar and Cohen-Zada, 2018; Fredriksson and Öckert, 2014; McEwan and Shapiro, 2008). Moreover, we add the number of siblings and assign a value of zero to those who indicated they have no siblings (Attar and Cohen-Zada, 2018).[15] To adjust for pure age effects,

[13] All estimations are pursued with the software R (version 4.2.3).

[14] If an individual's month of birth, and thus the assigned school-starting age is entirely random, no control variables are required (Mühlenweg and Puhani, 2010). While including exogenous control variables enhances the precision of the estimator, introducing non-exogenous controls could lead to the two-stage least squares estimator being inconsistent.

[15] Note that we observe some variation in the number of siblings within individuals over time. For this reason, X'_{it} in Equation 2.5 and Equation 2.6 varies in i and t.

we include birth-year dummies and thereby control for respondents' age at the time of observation (for a discussion see e.g., Black et al., 2011; Fredriksson and Öckert, 2014). Finally, we include state of school start and survey year fixed effects.

2.2.5 Sample

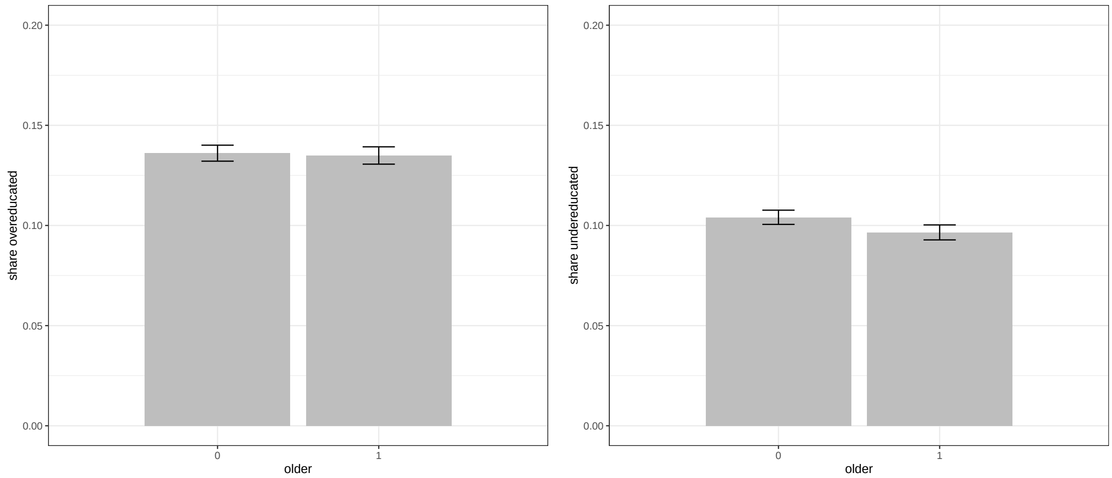
Educational mismatch can only be assessed for the working population. Accordingly, we restrict the sample to observations of individuals in paid employment, excluding those who are unemployed, solo-self-employed, or self-employed in the respective survey year. Additionally, we do not consider observations of individuals after they reached their legal retirement age, while accounting for the step-wise increase from 65 to 67 years as regulated in §235 of the German Social Security Code (*Sozialgesetzbuch)*. Finally, we remove observations with missing values in the defined variables, including the educational mismatch measures,[16] school-starting age, the instrument, the running variable, and all covariates. Our estimation sample includes 53,562 observations from 8,954 individuals, with 36,838 observations from 6,227 individuals within the $\pm$ 4-month window. The latter forms the base for the main estimations. The observation period covers 1991 to 2020.

Table A.1 presents summary statistics applying the $\pm$ 4-month window around the cutoff. The average school-starting age is 81.79 months (6.8 years), while it ranges from 56 months (4.6 years) to 116 months (9.6 years).[17] Approximately 46% were born after the cutoff. The average distance to the cutoff is 0.03 months, translating into roughly one day. Moreover, 14% are classified as overeducated, while 10% are undereducated. Figure 2.2 plots the differences in means of overeducation and undereducation for individuals born before and after the cutoff. Descriptively, we observe little difference in the share of overeducation and a statistically slightly lower share of undereducation among those born after the cutoff. Table A.2 provides the differences in means for the full set of covariates.

[16]We drop observations with missing values in either of the three educational mismatch measures explained above to ensure consistency. This approach allows us to perform robustness checks with alternative measures, enhancing the reliability of the results.

[17]Given the range in school-starting age, in Section 2.3.2 we test whether our results are sensitive to outliers. Excluding individuals with atypically low or high school-starting ages does not alter our findings.

Figure 2.2: Differences in means of main dependent variables by treatment



(a) Overeducation (b) Undereducation

Note: Figure 2.2a and Figure 2.2b report the differences in means of overeducation and undereducation for individuals born before and after the cutoff. The sample is based on those individuals whose birth month is in a $\pm$ 4-month range around the respective cutoff date. The survey years cover 1991 to 2020. 90% confidence intervals plotted.

2.3 Effect of school-starting age on educational mismatch

2.3.1 Main results

Table 2.1 presents the effect of school-starting age on educational mismatch using fuzzy RDD estimations. Columns (1) and (3) present results including the fixed effects exclusively, while columns (2) and (4) add the covariates. The first-stage results confirm that being born after the cutoff significantly increases the school-starting age. Specifically, the school-starting age increases by approximately 2.2 months in both cases. The estimate is statistically significant at the 0.1%-level. Furthermore, the F-statistic is 94 (89.1), exceeding the critical threshold of 10 as recommended by Staiger and Stock (1997). Moreover, the results reveal that school-starting age does not affect the likelihood of overeducation either statistically or quantitatively. In contrast, school-starting age reduces the likelihood of undereducation by up to 1.1 percentage points, with the result being significant at the 0.1%-level. Table A.3 reports the full set of results, including estimates for the covariates. Finally, we observe no heterogeneity in the effect of school-starting age on the likelihood of either over- and undereducation by sex, birth year, or region, as displayed in Figure A.5 in the Appendix.

Table 2.1: Main results

	(1)	(2)	(3)	(4)
	Overeducation		Undereducation	
$\widehat{Months}$	-0.001	-0.001	-0.008**	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Dist	0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Covariates		X		X
First stage:				
Older	2.272***	2.201***	2.272***	2.201***
	(0.235)	(0.234)	(0.235)	(0.234)
F-stat. (1st stage)	94.0	89.1	94.0	89.1
Num. obs.	36838	36838	36838	36838
% correctly predicted	56.19	62.58	55.50	57.30

Note: The estimations are based on data from the SOEP for the years 1991 to 2020. Columns (1) to (4) present fuzzy regression discontinuity results using the variable *older* to instrument school-starting age in months. Columns (1) to (4) use linear probability models and include survey year, birth month, and federal state fixed effects. Columns (2) and (4) add further covariates, including dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by the father and mother. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3.2 Robustness

To ensure the reliability and validity of our findings, we conduct a series of robustness checks.

RDD design, standard errors and sample

Columns (1) to (4) of Table A.4 explore alternative specifications in the fuzzy RDD framework. Columns (1) and (2) incorporate a second-order polynomial to address potential non-linearities in the running variable. The estimates remain consistent with our main findings and show a decrease in the likelihood of undereducation by approximately 1.2 percentage points. To check whether the estimates are sensitive to the selection of the bandwidth around the cutoff, we apply a data-driven approach to determine the bandwidth using the methodology by Calonico et al. (2020) in columns (3) and (4) of Table A.4.[18] This approach suggests a bandwidth of $\pm$ 1.74 months. As the data only allow us to observe full months, we re-estimate the regressions with a $\pm$ 2-month range around the cutoff. These results corroborate our main findings, showing a decrease in the likelihood of undereducation by 1.3 percentage points.

[18] We implement the approach by Calonico et al. (2020) with the R package *rdrobust* (Calonico et al., 2023).

Additionally, we observe a 1.5 percentage point increase in the likelihood of overeducation.

In the base specification, we use heteroskedasticity-robust standard errors (see Wooldridge, 2010). In columns (5) and (6) of Table A.4, we cluster the standard errors at the combined level of the state of school start and the survey year. This allows correcting for state-of-school-start-specific influences and time-varying conditions, such as economic development, during the survey period. The results remain robust and substantiate the findings from the base specification. Finally, in columns (7) and (8) of Table A.4, we account for potential biases in the selection of our sample. In the main sample, school-starting age ranges from 56 to 116 months, spanning five years. To mitigate potential biases introduced by this wide range, we limit the sample to individuals starting school between the ages of 5 and 8 (60 to 96 months). Again, our results hardly change.

Alternative assessment of educational mismatch

The choice of method for measuring educational mismatch can significantly influence the results, as highlighted in previous studies (see, e.g., Capsada-Munsech, 2019; Flisi et al., 2017; Verhaest and Omey, 2006, 2010). For this reason, we take advantage of the SOEP data, which allows us to apply two alternative measures of educational mismatch to account for possible differences between these measures. The first alternative is the indirect self-assessment (ISA), which relies on respondents' evaluation of the educational requirements for their current position (Duncan and Hoffman, 1981). In the SOEP, this is captured through the question: *"What type of education or training is usually required for this type of work?"*. The answer to this question is categorical, including *"No training required"*, *"Professional training completed"*, *"College education completed"* and *"University education completed"*. This assessment is compared to a categorical measure of attained education, distinguishing between a high-school degree, vocational training, college education and university education. The second measure is the job analyst measure (JA), which assesses the level of required education in a certain job by using the official classification of occupations and their assigned skill levels (Rumberger, 1981a). We use the differentiation of skill groups as provided by the ISCO. These are compared to the equivalent skill levels assigned to the educational degree by the International Standard Classification of Education (ISCED, UNESCO, 2006).

The results, reported in Table A.5, are consistent with our main findings. Using the ISA and JA measures, we observe reductions in the likelihood of undereducation by 1.2 and 1.1 percentage points, respectively. Additionally, the effect of school-starting age on the likelihood of overeducation is positive and statistically significant when applying the JA measure.

Set of covariates

In our main analysis, we follow previous studies by only including covariates most likely defined prior to school enrolment, such as an individual's sex (e.g. Bahrs and Schumann, 2020). However, as local labour market conditions may also influence the likelihood of educational mismatch, we conduct an additional robustness test by extending our covariate set to include the regional unemployment rate and the regional share of unemployed individuals by qualification level (Charlot and Decreuse, 2005; Dolado et al., 2000; Lazear et al., 2018). We extract the information on both indicators from the INKAR (*Indikatoren und Karten zur Raum- und Stadtentwicklung*) database.[19] Columns (1) to (4) of Table A.6 show that our main results remain both qualitatively and quantitatively robust. Specifically, we find a consistent decrease in the likelihood of undereducation by 1.2 percentage points. Additionally, conditioning on the regional share of unemployed individuals by qualification level in column (3) reveals a 1 percentage point increase in the likelihood of overeducation.

We also consider the potential role of personality traits, which previous research has linked to the likelihood of educational mismatch (Blásquez and Budría, 2012). Moreover, recent findings by Barabasch et al. (2024) suggest that school-starting age may influence personality traits, raising concerns about potential bad control problems if these traits are included as covariates. For this reason, we do not use them in our main specification. However, to assess the sensitivity of our results, we add dummies for each of the Big Five personality traits in columns (5) and (6) of Table A.6. We classify individuals as open, agreeable, extraverted, conscientious, or neurotic if their reported values exceed the sample mean.[20] While all traits except agreeableness are significantly associated with educational mismatch, the impact of school-starting age remains qualitatively unchanged, aligning with the findings in Section 2.3.

[19]Data on the local unemployment rate has been available since 1998, while information on the regional share of unemployed individuals by qualification level has been available only since 2010.

[20]Since data on Big Five personality traits is available for only six survey waves, we treat these dummies as time-invariant.

2.4 Discussion

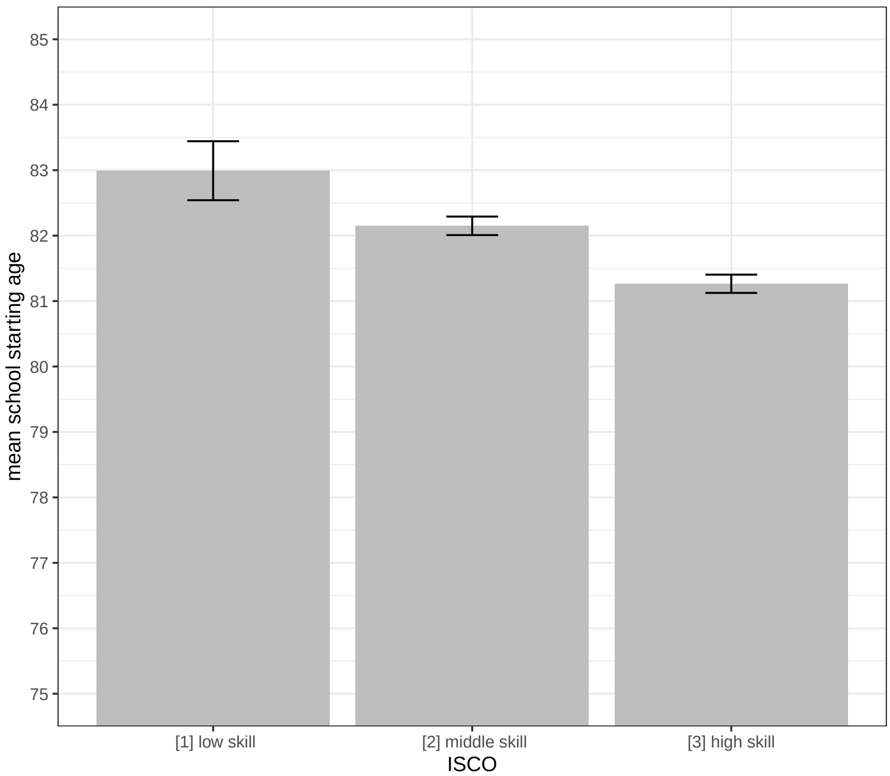
The effects of school-starting age on educational mismatch can arise through two primary channels: Occupational choice and educational attainment. In this section, we explore both mechanisms in order to better understand their contributions to our findings.

2.4.1 Occupational choice

For occupational choice to act as a channel for the effects of school-starting age on educational mismatch, there must be a systematic relationship between individuals' school-starting age and their occupation. Specifically, we would expect the average school-starting age to vary significantly across different types of occupations. Figure 2.3 illustrates the average school-starting age in months for the three skill levels defined by the ISCO classifications. These distinguish between low-skill (elementary positions), middle-skill (clerks, service, agricultural and fishery, craft and trade, plant, and machine operators), and high-skill (legislators, senior officials and managers, professionals, technicians, and associate professionals) occupations. The data reveal a clear pattern: Individuals in low-skill occupations exhibit the highest average school-starting age, followed by those in middle-skill occupations. In contrast, individuals in high-skill occupations tend to have the lowest average school-starting age.

Table 2.2 provides a formal test of the hypothesis that the effect of school-starting age on educational mismatch operates through occupational choice in the spirit of Baron and Kenny (1986). Column (1) estimates the impact of school-starting age on the likelihood of working in a high-skill occupation. The results indicate that each additional month of school-starting age reduces the likelihood of working in a high-skill occupation by 1.9 percentage points, consistent with the descriptive evidence. Columns (2) to (5) assess whether occupational choice mediates the relationship between school-starting age and undereducation. Columns (2) and (4) replicate the baseline specification, while columns (3) and (5) include the potential mediator as an additional covariate. The results show that working in a high-skill occupation increases the likelihood of undereducation by 8.2 percentage points. Additionally, including this mediator reduces the estimated effect of a one-month increase in school-starting age on undereducation from 1.1 to 0.9 percentage points, corresponding to a reduction of around 18%. Despite this comparatively large reduction, the confidence

Figure 2.3: Occupations and school-starting age



Note: Figure 2.3 displays the average school-starting age in months by ISCO skill group. The skill groups cover ISCO-group 9 in "[1] low skill", ISCO-groups 4 to 8 in "[2] middle skill" and ISCO-groups 1 to 3 in "[3] high skill". 90% confidence intervals plotted.

intervals of both estimates overlap. Therefore, we interpret this reduction with caution.

In sum, the results indicate that occupational choice might be a channel through which school-starting age impacts undereducation. Although the reduction in the coefficient is comparably large, the data do not allow for the establishment of a significant mediation. While this may be related to the large time-span lying between school-starting age and job choice, it may also suggest that other mechanisms, such as educational attainment, may play a role in the observed relationship.

2.4.2 Educational attainment

If starting school at an older age affects students' academic performance, this could credibly explain its impact on educational mismatch. Overeducation, for instance, is more likely among individuals with higher levels of educational

Table 2.2: Channel: Occupational choice

Outcome	(1) High skill	(2) Overeducation	(3)	(4) Undereducation	(5)
$\hat{Months}$	-0.019***	-0.001	0.000	-0.011***	-0.009**
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Dist	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High skill			0.023***		0.082***
			(0.005)		(0.005)
First stage:					
Older	2.201***	2.201***	2.203***	2.201***	2.203***
	(0.234)	(0.234)	(0.233)	(0.234)	(0.233)
F-stat.	89.1	89.1	89.4	89.1	89.4
Num. obs.	36838	36838	36838	36838	36838
% correctly predicted	14.35	62.58	62.30	57.31	57.46

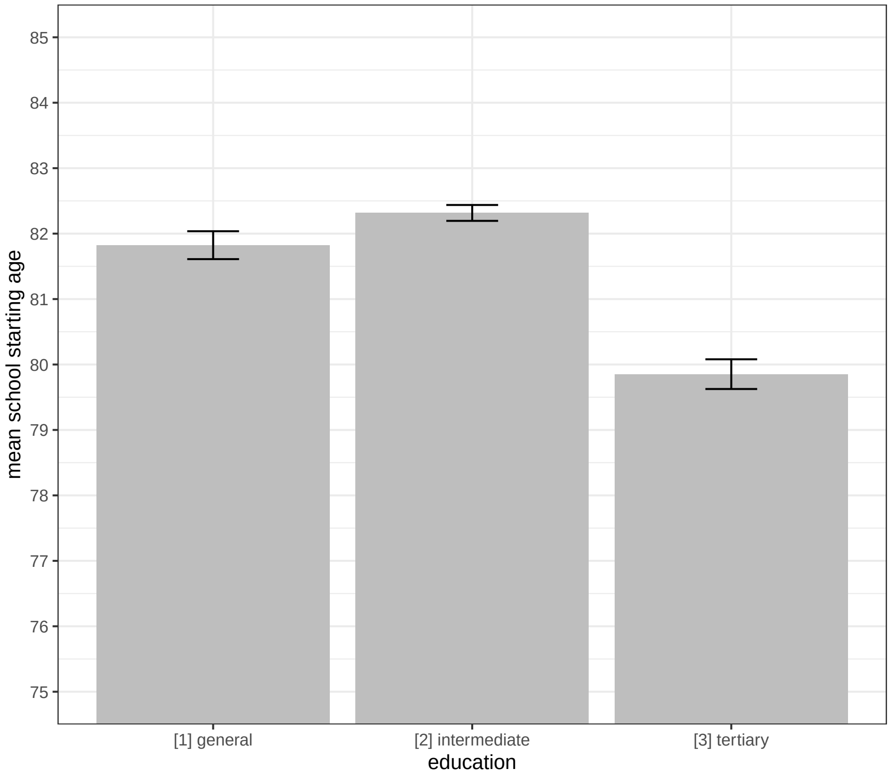
Note: The estimations are based on data from the SOEP for the years 1991 to 2020. Columns (1) to (5) contain results from fuzzy regression discontinuity estimations using the variable *older* to instrument school-starting age in months. Columns (1) to (5) use linear probability models. Column (1) estimates the impact of school-starting age on the likelihood of working in a high-skill occupation (ISCO groups 1 to 3), while columns (2) and (3) ((4) to (5)) estimate the impact of school-starting age on the probability of overeducation (undereducation). Columns (2) and (4) present the baseline specification discussed in Section 2.3. Columns (3) and (5) contain the mediation analysis, adding working in a high-skill occupation as an additional covariate. Controls include dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by father and mother. Survey year, birth month, and federal state fixed effects are included. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

attainment, while undereducation is structurally more common among those with the lowest levels of education.

As evidence on the impact of school-starting age on educational attainment is mixed and partially contradicting (e.g., Chen and Park, 2021; Crawford et al., 2014; Fletcher and Kim, 2016; Martin, 2009; Suziedelyte and Zhu, 2015), we again begin by examining both variables descriptively. Figure 2.4 illustrates the average school-starting age in months by educational degree attained. The findings show that individuals with tertiary degrees tend to have the lowest average school-starting age, while those with intermediate education levels exhibit the highest. This suggests that children who go on to achieve tertiary education tend to start school slightly earlier. Although these differences are statistically significant, their magnitude is relatively small, with average school-starting ages ranging narrowly between 6.6 and 6.8 years.

Table 2.3 presents the formal analysis of educational attainment as a channel. Column (1) demonstrates that an increase in school-starting age reduces the likelihood of attaining a tertiary degree by 1.7 percentage points, aligning with the descriptive findings from Figure 2.4. This negative association suggests that starting school at an older age is linked to a lower probability of obtaining tertiary education. In columns (2) to (5), we test whether owning

Figure 2.4: Education and school-starting age



Note: Figure 2.4 displays the average school-starting age in months by educational degree. The educational degrees cover "[1] general", "[2] intermediate" and "[3] tertiary". 90% confidence intervals plotted.

a tertiary degree is a channel for the effect of school-starting age on the likelihood of educational mismatch. Notably, when we include attainment of a tertiary degree as a covariate, the quantitative effect of school-starting age on undereducation strengthens marginally in quantitative terms. Additionally, its inclusion reveals a significant positive relationship between school-starting age and overeducation. The likelihood of overeducation increases by 0.6 percentage points, or approximately 4%, for each additional month of school-starting age. This result, statistically significant at the 10%-level, might suggest that the effect of school-starting age on the likelihood of overeducation is underestimated in the main analysis due to the suppressing effect of educational attainment. Still, as above, the difference in the point estimates is not statistically different.

2.5 Conclusion

Although empirical evidence on the impact of school-starting age on academic and behavioural outcomes is extensive, evidence on its long-term effects on

Table 2.3: Channel: Educational attainment

Outcome	(1) Tertiary	(2) Overeducation	(3)	(4) Undereducation	(5)
$\hat{Months}$	-0.017***	-0.001	0.006+	-0.011***	-0.012***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Dist	0.002+	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Tertiary			0.378***		-0.114***
			(0.010)		(0.009)
First stage:					
Older	2.201***	2.201***	2.194***	2.201***	2.194***
	(0.234)	(0.234)	(0.234)	(0.234)	(0.234)
F-stat.	89.1	89.1	88.5	89.1	88.5
Num. obs.	36838	36838	36838	36838	36838
% correctly predicted	57.44	62.58	78.07	57.31	55.68

Note: The estimations are based on data from the SOEP for the years 1991 to 2020. Columns (1) to (5) contain results from fuzzy regression discontinuity estimations using the variable *older* to instrument school-starting age in months. Columns (1) to (5) use linear probability models. Column (1) estimates the impact of school-starting age on the likelihood of having tertiary education, while columns (2) and (3) ((4) to (5)) estimate the impact of school-starting age on overeducation (undereducation). Columns (2) and (4) present the baseline specification discussed in Section 2.3. Columns (3) and (5) contain the mediation analysis, adding tertiary education as an additional covariate. Controls include dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by father and mother. Survey year, birth month, and federal state fixed effects are included. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

labour market outcomes, particularly regarding job match quality, remains limited. This study contributes to the literature by examining how school-starting age influences educational mismatch, a key determinant of labour market efficiency. Diverse school-entry cutoffs varying across cohorts and federal states in Germany enable the application of RDD estimations. Using a fuzzy RDD estimation approach, we find compelling evidence that starting school one month older reduces the likelihood of undereducation by 1.1 percentage points in the base specification. This result holds robust to a variety of robustness checks covering the inclusion of a second-order polynomial in the running variable, variations of the bandwidth around the cutoffs as proposed by Calonico et al. (2020), and alternative estimation methods for the standard errors. Moreover, similar effects are observed when following previous studies, suggesting the assessment of educational mismatch in multiple ways (Capsada-Munsech, 2019; Flisi et al., 2017; Verhaest and Omey, 2010).

As educational mismatch is driven by either educational attainment or occupational choice, we finally investigate whether these two mechanisms may mediate our findings. Doing so provides suggestive evidence that the impact of school-starting age on the likelihood of undereducation could be partially mediated by occupational choice. In more detail, starting school one month older is related to a lower likelihood of working in a high-skill occupation, and including the mediator reduces the point estimate for school-starting age by 18.19%. In contrast, when educational attainment is included in the base specification, we observe a 9% increase in the magnitude of the coefficient for school-starting age on the likelihood of undereducation. While the base specification did not reveal a significant impact of school-starting age, starting school older increases the likelihood of overeducation, conditional on educational attainment. Although the inclusion of occupational choice or educational attainment leads to notable changes in the point estimates, the data do not provide sufficient evidence to identify statistically significant differences.

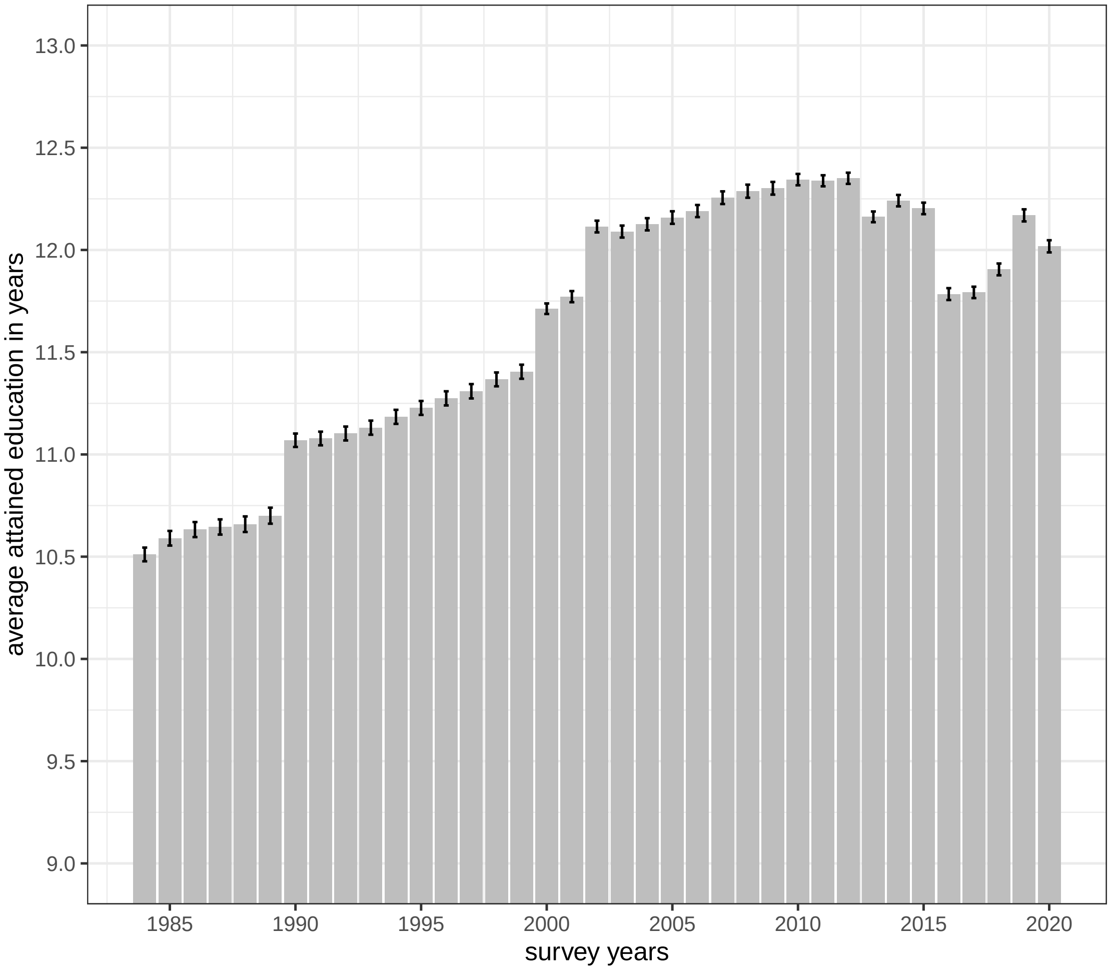
In conclusion, our findings indicate that the age at which students start school has lasting implications for educational mismatch, with potential consequences for labour market outcomes. Specifically, our results show that starting school at an older age reduces the likelihood of undereducation, a state often regarded as preferable due to lower educational investment costs (see, e.g., Duncan and Hoffman, 1981; Sloane et al., 1996). Therefore, our findings also highlight that the advantages of delayed school entry, as emphasised in existing literature, are more nuanced when viewed through the lens of educational mismatch. These insights underscore the trade-offs inherent in rigid school-entry policies, which may inadvertently shape individuals' educational trajectories and labour market prospects. By providing empirical evidence on these complexities, our study contributes to the broader policy debate on optimising school system design balancing educational attainment with labour market efficiency.

Nevertheless, our study has several limitations. First, optimising school entry policies – such as introducing more flexible age cutoffs or tailored early interventions – requires a deeper understanding of the mechanisms and timing through which school-starting age exerts its effects. This would necessitate more detailed data on individual educational pathways. Second, we identify the local average treatment effect rather than the average treatment effect, meaning our results apply specifically to individuals who comply with school-entry rules and cannot be generalised to those who start school earlier or delay entry through red-shirting. Third, our estimations pertain to the German context, where children typically start school at ages six or seven. As school-starting age policies vary internationally, our findings may not be directly transferable to other countries. Additionally, prior research suggests that the effects of school-starting age are more pronounced in systems with early academic tracking (Fredriksson and Öckert, 2014). Given Germany's relatively early tracking system, our estimated effects may represent an upper bound compared to countries with later or less differentiated tracking. To improve the external validity of our findings, further research across diverse educational systems is necessary.

2.6 Appendix A

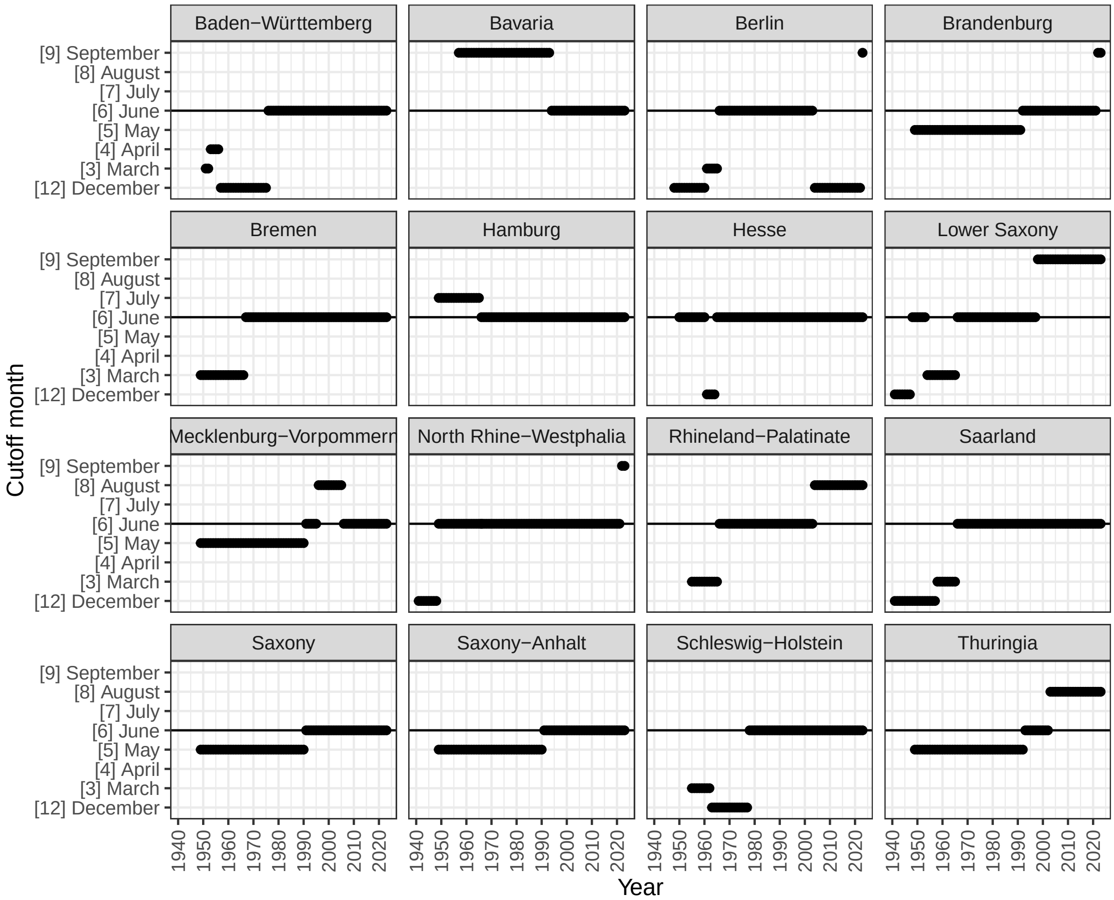
2.6.1 Graphs

Figure A.1: Average years of education by survey year



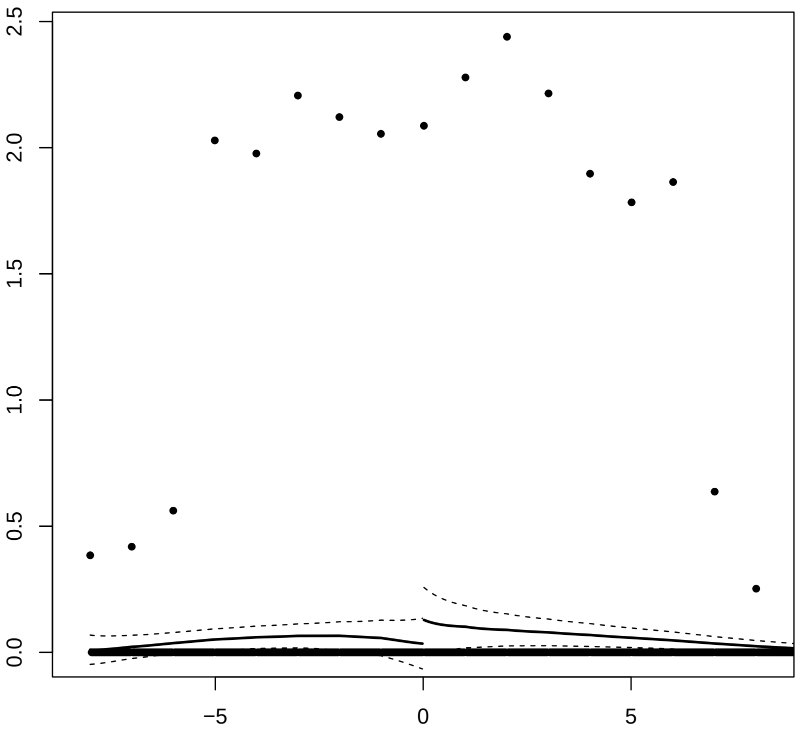
Note: Figure A.1 shows the average years of education attained for all individuals observed in each survey year of the SOEP from 1984 to 2020. 90% confidence intervals plotted.

Figure A.2: Cutoff dates by federal state



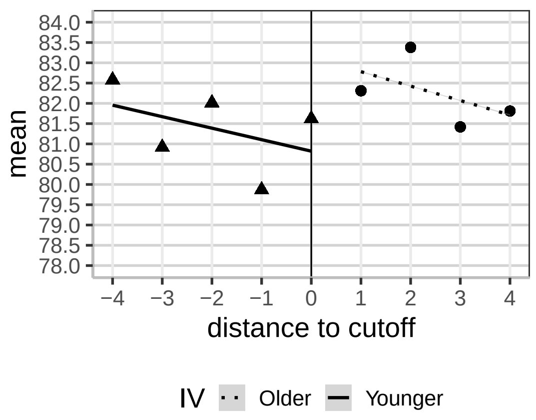
Note: Figure A.2 reports the cutoff dates by year and federal state. Data on the cutoff dates was collected by going through the respective state legislation. The reference line corresponds to June, as this month is the most frequently used cutoff month.

Figure A.3: McCrary manipulation test

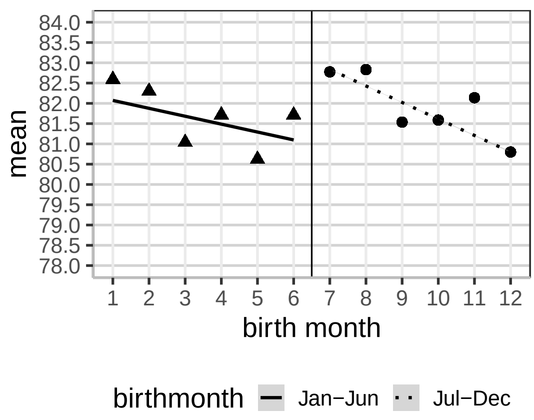


Note: Figure A.3 provides results from the manipulation test for the running variable *dist* as proposed by McCrary (2008). The estimated bandwidth is 5.018, the log difference in heights equals 1.321, and the p-value is 0.000.

Figure A.4: Discontinuity in school-starting age



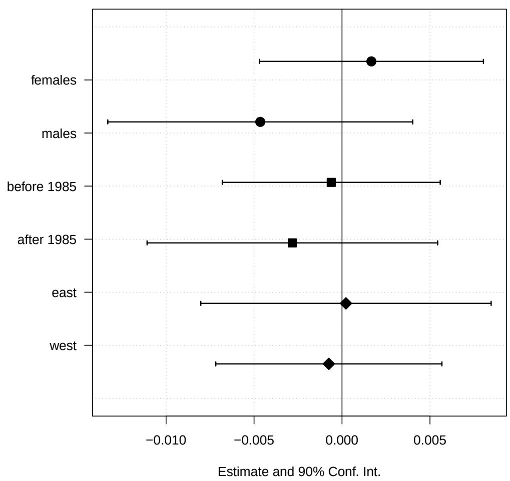
(a) By distance to cutoff



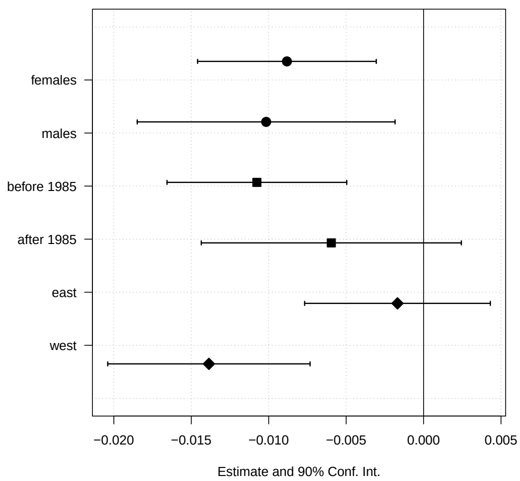
(b) By birth month

Note: Figure A.4a and Figure A.4b display the discontinuity in the school-starting age variable by distance to the cutoff and by birth month.

Figure A.5: Heterogeneity by demographics



(a) Overeducation



(b) Undereducation

Note: Figure A.5a and Figure A.5b display regression results for the likelihood of over- and undereducation from fuzzy RDD estimations for different subsamples, respectively. The subsamples distinguish individuals based on sex, cohort, and region. 90%-confidence intervals displayed.

2.6.2 Tables

Table A.1: Summary statistics

	Mean	SD	Min	Max
School-starting age	81.79	11.37	56.00	116.00
Older	0.46	0.50	0.00	1.00
Dist	0.03	2.54	-4.00	4.00
Overeducation	0.14	0.34	0.00	1.00
Undereducation	0.10	0.30	0.00	1.00
Number siblings	1.51	1.35	0.00	14.00
Female	0.52	0.50	0.00	1.00
Birth year	1970.28	11.01	1939.00	1998.00
Direct migrant	0.03	0.17	0.00	1.00
Indirect migrant	0.05	0.21	0.00	1.00
Native	0.92	0.27	0.00	1.00
Father educ.: upper	0.13	0.34	0.00	1.00
Father educ.: intermediate	0.20	0.40	0.00	1.00
Father educ.: general	0.63	0.48	0.00	1.00
Father educ.: no	0.03	0.18	0.00	1.00
Mother educ.: upper	0.08	0.28	0.00	1.00
Mother educ.: intermediate	0.26	0.44	0.00	1.00
Mother educ.: general	0.62	0.48	0.00	1.00
Mother educ.: no	0.03	0.18	0.00	1.00
Num. obs.		36838		

Note: The sample is based on those individuals whose birth month is in a $\pm$ 4-month range around the respective cutoff date, restricting the survey waves to 1991 to 2020. SOEP weights applied.

Table A.2: Differences in means by instrument *older*

	Older=0	Older=1	Diff.	Std. Error
School-starting age	81.39	82.28	0.889***	0.12
Dist	-1.99	2.42	4.416***	0.01
Overeducation	0.14	0.14	0.00	0.00
Undereducation	0.10	0.10	-0.008*	0.00
Number siblings	1.47	1.56	0.090***	0.01
Female	0.53	0.52	-0.01	0.01
Birth year	1970.97	1969.47	-1.505***	0.12
Direct migrant	0.03	0.03	0.00	0.00
Indirect migrant	0.05	0.05	0.00	0.00
Native	0.92	0.93	0.00	0.00
Father educ.: upper	0.13	0.14	0.00	0.00
Father educ.: intermediate	0.20	0.20	0.00	0.00
Father educ.: general	0.63	0.63	-0.009+	0.01
Father educ.: no	0.03	0.04	0.007***	0.00
Mother educ.: upper	0.08	0.08	0.00	0.00
Mother educ.: intermediate	0.28	0.24	-0.037***	0.01
Mother educ.: general	0.61	0.64	0.028***	0.01
Mother educ.: no	0.03	0.04	0.009***	0.00
Num. obs.	19964	16874		

Note: The sample is based on those individuals whose birth month is in a $\pm$ 4-month range around the respective cutoff date restricting the sample to the survey years 1991 to 2020. SOEP weights applied. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Main results

	(1)	(2)	(3)	(4)
	Overeducation		Undereducation	
$\widehat{Months}$	-0.001	-0.001	-0.008**	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Dist	0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Female		-0.008*		0.001
		(0.004)		(0.004)
Direct migrant		-0.031		0.095***
		(0.021)		(0.021)
Indirect migrant		0.006		0.004
		(0.009)		(0.009)
Number of siblings		-0.009***		0.014***
		(0.001)		(0.001)
Father educ.: general		-0.051***		-0.071***
		(0.013)		(0.017)
Father educ.: intermediate		-0.014		-0.088***
		(0.013)		(0.017)
Father educ.: upper		0.072***		-0.102***
		(0.015)		(0.017)
Mother educ.: general		0.045***		-0.071***
		(0.009)		(0.016)
Mother educ.: intermediated		0.115***		-0.106***
		(0.011)		(0.017)
Mother educ.: upper		0.148***		-0.133***
		(0.015)		(0.019)
First stage:				
Older	2.272***	2.201***	2.272***	2.201***
	(0.235)	(0.234)	(0.235)	(0.234)
F-stat. (1st stage)	94.0	89.1	94.0	89.1
Num. obs.	36838	36838	36838	36838
% correctly predicted	56.19	62.58	55.50	57.30

Note: The estimations are based on data from the SOEP for the years 1991 to 2020. Columns (1) to (4) present fuzzy regression discontinuity results using the variable *older* to instrument school-starting age in months. Columns (1) to (4) use linear probability models and include survey year, birth month, and federal state fixed effects. Columns (2) and (4) add further covariates. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Robustness: Estimation design

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RDD design				Estimation		Sample	
	Quadratic		Cutoff: 2 months		Alternative SE		Outlier	
	OE	UE	OE	UE	OE	UE	OE	UE
$\hat{Months}$	0.000	-0.012***	0.015+	-0.013+	-0.001	-0.011***	-0.003	-0.009**
	(0.003)	(0.003)	(0.008)	(0.007)	(0.004)	(0.003)	(0.003)	(0.003)
Dist	0.000	0.000	-0.010*	0.001	0.000	0.000	0.001	0.004*
	(0.001)	(0.001)	(0.005)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)
Dist^2	0.000	0.000						
	(0.000)	(0.000)						
First stage:								
Older	2.241***		1.313***		2.201***		2.371***	
	(0.238)		(0.316)		(0.240)		(0.177)	
F-stat. (1st stage)	87.4		17.2		89.1		184.1	
Num. obs.	36838	36838	20985	20985	36838	36838	32009	32009
% correctly predicted	62.55	57.26	61.74	57.48	62.58	57.31	62.27	55.71

Note: The estimations are based on data from the SOEP for the years 1991 to 2020 in columns (1) to (8). Columns (1) to (8) contain results from fuzzy regression discontinuity estimations using the variable *older* to instrument school-starting age in months. Uneven columns report results for the likelihood of overeducation (OE), while even columns consider the likelihood of undereducation (UE). Columns (1) and (2) include the second-order polynomial in the running variable. Columns (3) and (4) apply a data-driven cutoff of ± 2 months following Calonico et al. (2020). Columns (5) and (6) alter the estimation of the standard errors. Columns (7) to (8) rely on a sample excluding outliers in the school-starting age variable. Controls include dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by father and mother. Survey year, birth month, and federal state fixed effects are included. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. Standard errors clustered at the state of school start times survey year-level in columns (5) and (6). The percentage correctly reported is displayed following Wooldridge (2010). + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Robustness: Measures

	(1)	(2)	(3)	(4)
	Indirect self-assessment		Job analyst	
	OE	UE	OE	UE
$\hat{Months}$	0.005	-0.012***	0.012***	-0.011**
	(0.004)	(0.003)	(0.004)	(0.004)
Dist	0.004***	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
First stage:				
Older	2.201***			
	(0.234)			
F-stat. (1st stage)	89.1			
Num. obs.	36838	36838	36838	36838
% correctly predicted	55.63	58.02	53.45	56.66

Note: The estimations are based on data from the SOEP for the years 1991 to 2020 in columns (1) to (4). Columns (1) to (4) contain results from fuzzy regression discontinuity estimations using the variable *older* to instrument school-starting age in months. Uneven columns report results for the likelihood of overeducation (OE), while even columns consider the likelihood of undereducation (UE). Columns (1) to (4) rely on different definition methods of educational mismatch, applying the indirect self-assessment (ISA) in columns (1) and (2) and the job analyst measure (JA) in columns (3) and (4). Controls include dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by father and mother. Survey year, birth month, and federal state fixed effects are included. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Robustness: Omitted variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Regional labour market indicators				Personality traits	
	Unempl. rate		Unempl. by qualification			
	OE	UE	OE	UE	OE	UE
$\hat{Months}$	-0.001	-0.012***	0.009*	-0.012***	-0.002	-0.010**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Dist	0.000	0.000	0.000	0.002	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unemployment rate	0.001	-0.002+				
	(0.001)	(0.001)				
Specialist			0.060**	-0.010		
			(0.022)	(0.018)		
Trained			0.000	-0.002		
			(0.004)	(0.003)		
Experts			-0.011	-0.005		
			(0.013)	(0.011)		
Open					0.008*	0.010**
					(0.004)	(0.004)
Extraverted					-0.019***	0.013***
					(0.004)	(0.004)
Neurotic					-0.005	0.012**
					(0.004)	(0.004)
Agreeable					0.005	-0.001
					(0.004)	(0.004)
Conscientious					-0.025***	-0.008*
					(0.004)	(0.004)
First stage:						
Older	2.157***		2.362***		2.267 ***	
	(0.237)		(0.300)		(0.246)	
F-stat. (1st stage)	83.4		62.6		84.9	
Num. obs.	35989	35989	22374	22374	33121	33121
% correctly predicted	57.15	53.05	51.16	53.86	54.61	54.70

Note: The estimations are based on data from the SOEP for the years 1998 to 2020 in columns (1) and (2), 2010 to 2020 in columns (3) and (4) and 1991 to 2020 in columns (5) and (6). Columns (1) to (6) contain results from fuzzy regression discontinuity estimations using the variable *older* to instrument school-starting age in months. Uneven columns report results for the likelihood of overeducation (OE), while even columns consider the likelihood of undereducation (UE). Columns (1) to (6) add new covariates with the local unemployment rate in columns (1) and (2) and the share of unemployed individuals by qualification level in columns (3) and (4). Columns (5) and (6) condition on dummies covering the Big Five personality traits. Controls include dummies for having a direct or indirect migration background, the number of siblings, being female, and the level of education obtained by father and mother. Survey year, birth month, and federal state fixed effects are included. The percentage correctly predicted is displayed following Wooldridge (2010). Heteroskedasticity-robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 3

Who bears the brunt: Tuition fees and educational mismatch*

Exploiting the quasi-experimental introduction of tuition fees in selected federal states of Germany between 2006 and 2014, this study investigates how these fees affect the likelihood of overeducation based on data from the SOEP. Reporting lower bound estimates, the findings show that graduates from fee-charging states are significantly more likely to experience overeducation. There is suggestive evidence that the Intention to Treat Effect (ITT) may be heterogeneous depending on occupational choice after graduation, socioeconomic background and duration of the treatment. Moreover, dynamic analyses reveal that the increase in the likelihood of overeducation due to tuition fee exposure is not short-term but persists for even up to ten years after degree completion.

Keywords: Tuition fees; Overeducation; Quasi-experiment; Germany

*I am grateful for helpful comments by Christina Gathmann, Julio Garbers, Regina Riphan, Joachim Möller, Thorsten Schank, Vivien Voigt, Ugur Aytun, Marco Clemens and Laszlo Goerke, as well as by the participants of the Joint LISER & IAAEU Workshop on Labour and Personnel Economics, the 2025 Annual Conference of the Scottish Economic Society and the TriEcon Workshop on Education Economics.

3.1 Introduction

Access to higher education is often shaped by financial barriers, making it a key concern in both education economics and policy. Human Capital Theory suggests that individuals invest in education to yield higher future earnings (Becker, 1975). However, higher education costs can pose substantial obstacles to human capital accumulation, particularly for individuals with limited financial resources and borrowing constraints. This is particularly relevant in countries such as the US which rely heavily on tuition fees to finance the tertiary education system.[21] This is supported by numerous studies showing that student subsidies, loans, and grants are important measures to buffer the consequences of tuition fees, and might help to equalise opportunities among students from different backgrounds (e.g., Bettinger et al., 2019; Castleman and Long, 2016; Denning, 2019; Dynarski, 2008; Dynarski et al., 2021; Fack and Grenet, 2015; Page et al., 2019; Solis, 2017). In contrast, less is known about the effects of introducing tuition fees in countries where higher education has historically been free. As one of these states, Germany offers a unique example where tuition fees of EUR 500 were temporarily introduced in seven federal states between 2006 and 2014, while all other federal states kept tertiary education free of charge.[22]

Using this quasi-experimental setting, previous research has focused on the impact of tuition fees on enrolment rates and intentions, educational attainment in the population and student mobility (Bahrs and Siedler, 2019; Bietenbeck et al., 2023; Bruckmeier et al., 2013; Dwenger et al., 2012; Hübner, 2012; Minor, 2023). However, long-term consequences are less researched. Some studies suggest that the monetary pressure caused by tuition fees could influence individuals' career choices (e.g., Rothstein and Rouse, 2011). Such an impact on career choices might induce them to accept jobs below their qualifications at higher rates, for example, to accelerate the transition between university and labour market entry (Gervais and Ziebarth, 2019; Minicozzi, 2005). In addition to this mechanism, tuition fees could also deter academic performance or extracurricular engagement, which are important signals in the application process. Weaker signals could, in turn, reduce employment prospects and increase the likelihood of working in jobs that require less education than the attained level (Spence, 1973;

[21] According to the OECD, the annual average costs amounted to USD 9,596 for Bachelor students at public universities and USD 34,041 at private institutions in the US in the academic year of 2022/2023 (for details see Table C5.1 in OECD, 2024).

[22] Importantly, tuition fees are paid directly to the university and differ from term contributions (*Semesterbeitrag*). The latter generally cover administrative costs such as student union contributions or public transport fees and are charged regardless of tuition fees.

Thurow, 1975). If these mechanisms materialised, this could imply substantial long-term consequences for individuals as well as firms and might produce inefficiencies in the labour market. However, empirical studies on such a link are missing. Hence, a potential, important consequence of tuition fees has not yet been investigated.

This study fills this gap by asking: First, do tuition fees affect individuals' likelihood of working in jobs below their qualifications post-graduation? Second, does this link vary across subgroups? And third, does it persist over time? To answer these questions, I operationalise jobs below individuals' qualifications using the concept of overeducation.[23] Overeducation forms part of underemployment (Feldman, 1996) and defines the state in which individuals' formal education exceeds the job requirements (Freeman, 1976).[24] Moreover, I rely on the quasi-experimental setting in Germany to identify the impact of tuition fees and use data on graduates from the Socio-Economic Panel for the years 1985 to 2021 (DIW, 2024; Goebel et al., 2019). The evidence reveals that affected graduates are indeed significantly more likely to be overeducated post-graduation. This result holds robust to a series of sensitivity checks, and concerns regarding sources of selection biases are tackled. Suggestive evidence indicates that the treatment effect is less pronounced among individuals in high-skill occupations, with one parent owning upper secondary education or those treated for a shorter period. Finally, by leveraging the panel structure of the SOEP, this study shows that individuals exposed to tuition fees report a higher likelihood of overeducation not only in the short run but also in the ten years following graduation.

This study adds to the literature in three ways. First, the study extends our knowledge of the consequences of higher education costs. Most related studies evaluate the role of subsidies (e.g., Castleman and Long, 2016; Denning, 2019; Dynarski, 2008; Dynarski et al., 2021; Fack and Grenet, 2015; Marx and Turner, 2019; Monks, 2001; Page et al., 2019; Solis, 2017)[25] or tuition fees (e.g., Andrews and Stange, 2019; Garibaldi et al., 2012; Hassani-Nezhad et al., 2021; Molina and Rivadeneyra, 2021; Neill, 2009) in countries where tuition fees are common. In contrast, the present study evaluates the introduction of tuition fees in Germany,

[23]This study focuses on the likelihood of overeducation. Undereducation is not considered because it is, by definition, highly unlikely among graduates of tertiary education who were affected by the policy change.

[24]Distinct concepts cover overskilling or horizontal educational mismatches, focusing either on whether individuals' skills exceed the job requirements or whether individuals work in positions outside of their field of study (McGuinness et al., 2018; Robst, 1995b, 2007).

[25]Other studies evaluate the impact of support systems to finance education in primary or secondary education (e.g., Angrist et al., 2002; Burlando, 2023).

where higher education was generally free before the intervention. This enables an analysis of tuition fees as a binary policy change. Moreover, by linking this reform to post-graduate overeducation, the study expands on previous research by Bahrs and Siedler (2019), Bietenbeck et al. (2023), Bruckmeier et al. (2013), Dwenger et al. (2012), Fischer and Wigger (2016), Hübner (2012), and Minor (2023) that has largely focused on pre-graduation measures like enrolment decisions and educational intentions.

Second, it complements the literature on the antecedents of overeducation, which has, to a large extent, focused on demographic or job-related factors (e.g., Aleksynska and Tritah, 2013; Battu et al., 1999; Belfield, 2010; Blásquez and Budría, 2012; Caroleo and Pastore, 2018; Carroll and Tani, 2015; Davia et al., 2017; Diem, 2015; Diem and Wolter, 2014; Dolton and Silles, 2008; Fleming and Kler, 2008; Turmo-Garuz and Bartual-Figueras, 2019; Verhaest and Omey, 2010). This study extends this strand of literature by leveraging a quasi-experimental setting that enables the application of more nuanced research designs. While the available data and setting do not allow causal claims, such quasi-experimental approaches are rare in the educational mismatch literature and provide an important extension to predominantly correlational analyses.[26]

Third, the dynamic analysis provided at the end of this study builds on the literature evaluating Career Mobility Theory as an explanation of overeducation. Following this theory, individuals might enter overeducation after graduation to acquire working experience and further training, allowing them to proceed to better-matched jobs in the aftermath (Sicherman, 1991; Sicherman and Galor, 1990). Previous research has mainly estimated whether individuals in overeducation are more likely to participate in training measures, to switch their jobs, or to stay in overeducation for a longer period (e.g., Baert et al., 2013; Büchel and Mertens, 2004; Grunau and Pecoraro, 2017; Kiersztyn, 2013; Roller et al., 2020; Sicherman, 1991; Wen and Maani, 2019). Adding to that, this study evaluates whether the increase in the likelihood of overeducation caused by tuition fees is a short-term phenomenon – helping individuals to transition to matched jobs – or whether it endures even several years after graduation – trapping individuals in overeducation.

The study proceeds as follows: Section 3.2 explains the institutional background and develops the expectations. Section 3.3 describes the dataset, the variables, summarises the sample, and explains the identification strategy. Section

[26] As an exception, Ordine and Rose (2017) investigate the link between the supply of graduates in the labour market and their likelihood of overeducation using a quasi-natural experiment in Italy.

3.4 comprises the main findings, a series of robustness checks, a discussion of potential selection issues, the heterogeneity analyses, and the dynamic analysis. Section 3.5 concludes the study.

3.2 Institutional background and expectations

3.2.1 Quasi-experimental setting

Studying in Germany has generally been free of charge at public universities since the early 1970s for first and consecutive study programmes. This Germany-wide tuition fee exemption was enshrined in the legislation in 2002 when the German parliament (*Bundestag*) passed a law banning tuition fees for all first and consecutive study programmes in Germany.[27] However, following a legal complaint by the federal states of Bavaria, Baden-Württemberg, Hamburg, Saarland, Saxony, and Saxony-Anhalt, the Federal Constitutional Court (*Bundesverfassungsgericht*) ruled that the law violated the German Constitution (*Grundgesetz*) that delegates control over the education system to federal states.[28] As a result of this ruling, seven out of the sixteen federal states in Germany introduced tuition fees of up to EUR 500 per term between 2006 and 2014.

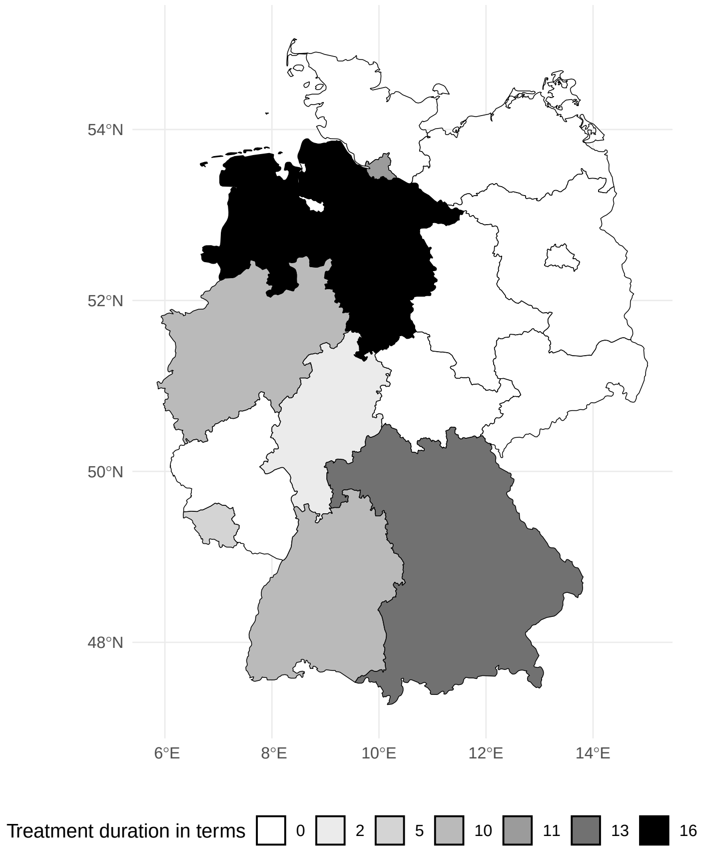
Lower Saxony and North Rhine-Westphalia were the first states to implement fees in the winter term of 2006/07. Subsequently, Bavaria, Hamburg, and Baden-Württemberg followed in the summer term of 2007, and Hesse and Saarland in the winter term of 2007/08. Hesse abolished the fees again after two terms, while Lower Saxony was the last state to remove the fees after the summer term of 2014 and charged tuition fees for over 16 terms.[29] Hence, the

[27] Sechstes Gesetz zur Änderung des Hochschulrahmengesetzes (6. HRGÄndG, Sixth Law Amending the Higher Education Framework Act) Art. 1 Nr. 3.

[28] Bundesverfassungsgericht, Urteil des Zweiten Senats (Federal Constitutional Court, Verdict of the Second Senate), January 26, 2005 - 2 BvF 1/03.

[29] See e.g., Baden-Württemberg Landeshochschulgebührengesetz (Higher Education Fees Act), December 19, 2005, §§3-4, Bayerisches Hochschulgesetz (Bavarian Higher Education Act), May 23, 2006, Art. 71, Hamburgisches Hochschulgesetz (Hamburg Higher Education Act), September 4, 2006, §6b, Hessisches Studienbeitragsgesetz (Hessian Tuition Fees Act), October 16, 2006, §§1-3, Niedersächsisches Hochschulgesetz (Lower Saxony Higher Education Act), December 15, 2005, §11, Hochschulabgabengesetz NRW (Higher Education Fees Act NRW), March 21, 2006, §2, Saarländisches Hochschulgebührengesetz (Saarland Higher Education Fees Act), July 12, 2006, §2.

Figure 3.1: Tuition fees in German federal states



Note: The map displays the federal states of Germany. States that introduced tuition fees between 2006 and 2014 are displayed in distinct shades of grey. The shades are determined by the length of the treatment period counted in terms. The treatment states are: Lower Saxony (winter term 2006/07 - summer term 2014), North Rhine-Westphalia (winter term 2006/07 - summer term 2011), Bavaria (summer term 2007 - summer term 2013), Hamburg (summer term 2007 - summer term 2012), Baden-Württemberg (summer term 2007 - winter term 2011/12), Hesse (winter term 2007/08 - summer term 2008), and Saarland (winter term 2007/08 - winter term 2009/10). Darker shades represent a longer period. Federal states depicted in white never introduced tuition fees and therefore exhibit a treatment duration of zero.

periods for which the fees were active varied largely across the federal states (see Figure 3.1).[30]

This partial and time-wise introduction of tuition fees in Germany provides a quasi-experimental setting enabling to study the influence of education costs on students' labour market outcomes. The following section outlines the mechanisms through which tuition fees may affect these outcomes, particularly with regard to overeducation.

[30]After fees were abolished, many states maintained fees for individuals pursuing a second university degree or for long-term students. Still, such regulations are also implemented in states that never introduced tuition fees for first and consecutive studies between 2006 and 2014. Saxony, for example, regulates that students who exceed the standard period of study by more than four terms due to their own responsibility have to pay EUR 500 for each following term (Sächsisches Hochschulgesetz (Saxon Higher Education Act), May 31, 2023, §13 (2).

3.2.2 Expectations

Assuming that the educational attainment of tertiary graduates is fixed, overeducation can generally arise through two mechanisms: First, individuals could voluntarily apply for jobs below their qualifications; Or second, firms could not hire individuals for jobs requiring their qualifications. This implies that overeducation can result from both the individual's and the firm's choices in the application process.

Regarding the individual's choice, previous research reported that monetary burdens, such as tuition fees or financial support mechanisms during studies, might influence occupational and labour market outcomes not only indirectly but also directly (see, e.g., Bettinger et al., 2019; Black et al., 2023; Field, 2009; Rothstein and Rouse, 2011; Zhang, 2013). Particularly, the financial burden of a loan or grant repayments could motivate individuals to enter the labour market faster. Such acceleration could induce individuals to accept jobs that do not align with their qualifications at higher rates (Gervais and Ziebarth, 2019; Minicozzi, 2005). Molina and Rivadeneyra (2021) provide evidence in line with this supposition and link tuition fees to occupational choice. They examine the abolition of fees in Ecuador in 2008 and demonstrate that individuals transitioned to higher-skilled jobs at higher rates once fees were eliminated.[31] If such a pattern translated to the German case, and individuals affected by tuition fees systematically selected into lower-skilled jobs, this would directly increase overeducation rates among the treated individuals.

Regarding the firm's choice, the conceptual work of Lazear et al. (2018) and Dolado et al. (2000) combined with standard Signalling (Spence, 1973) and Job Competition Theory (Thurow, 1975) suggest that graduates' likelihood of overeducation rises with increasing competition from equally or slightly better-qualified graduates. The reason for this being that firms will select the best-qualified graduate for the vacant position. Thus, for tuition fees to affect the likelihood of overeducation through this mechanism, *ceteris paribus*, the applications of treated individuals would have to convey signals of lower quality compared to those of untreated individuals. This would require tuition fees to affect performance indicators, such as the GPA or time-to-completion, and that the amount of EUR 500 per term was sufficient to provoke such changes.

[31] This effect was observed primarily among individuals from higher socioeconomic backgrounds. Molina and Rivadeneyra (2021) relate this to two explanations: First, individuals from lower socioeconomic backgrounds faced relatively low fees even before their abolition, and second, disparities in pre-university educational quality might deter university enrolment among them.

One plausible behavioural response to the introduction of tuition fees, which may affect graduates' signals, is that students might seek to minimise their total expenses and accelerate their studies (cf., Bietenbeck et al., 2023; Garibaldi et al., 2012). This could affect academic outcomes if students decided to take more exams at a time to complete their studies faster. On the one hand, such an increase in speed could translate into less time for each exam and could induce higher stress levels, possibly harming academic performance. On the other hand, studying for multiple courses at a time may also generate synergies, potentially improving overall performance. Empirical evidence by Garibaldi et al. (2012) on Bocconi University in Italy shows that higher fees reduce the time-to-completion. However, they did not affect the dropouts or academic performance of students. Partly contrary to that, Fricke (2018) investigates a sudden increase in tuition fees in Switzerland and does not find an effect on the likelihood of on-time graduation or academic performance indicators. Overall, the empirical evidence does not suggest that the performance of graduates reduces with rising costs because they aim to complete their studies faster. Hence, this mechanism should not affect graduates signals and thus, their likelihood of overeducation.

Alternatively, affected students could also adjust their behaviour by seeking additional funding for their studies. This may include external sources (e.g., from family or student loans) but also increasing their labour income (cf., Denning, 2019; Kalenkoski and Pabilonia, 2010; Neill, 2015). Transfers by external sources, e.g., parents, could ease students' academic performance as they can spend their time studying and do not have to invest more hours into working. Empirically, anyhow, while students receiving such transfers were shown to work less, study more, and report a higher likelihood of graduating, their academic performance indeed tends to be worse (Bachmann and Boes, 2014; Bodvarsson and Walker, 2004; Hamilton, 2013; Sauer, 2004). As a result, since worse academic performance likely is a negative signal for firms in the application process, one could expect higher likelihoods of overeducation among treated individuals based on this mechanism. In contrast, external subsidies could incentivise students to perform better, particularly when scholarships or grants are tied to academic achievements (e.g., Bernal and Penney, 2019; Montalbán, 2023). While this would be expected to diminish the likelihood of overeducation, purely performance-based subsidies are less prevalent in Germany compared to countries like the United States. If individuals increased their financial resources by working longer hours instead, this would reduce the time available for studying, leisure, or both. This could potentially harm academic performance (see e.g., Baert et al., 2018; Body et al., 2014; Darolia, 2014; Kalenkoski and Pabilonia, 2010; McKee-Ryan et al., 2009; McVicar and McKee, 2002; Montmarquette et al., 2007; Oettinger, 1999; Stinebrickner and Stinebrickner, 2003) and raise individuals' likelihood of overeducation. From another perspective, though, acquiring labour market experience besides studying may also serve as a positive signal in the application process, potentially buffering the risk of overeducation, particularly if gathered in a field related to one's studies (cf., Diem and Wolter, 2014; Verhaest and Omey, 2010).

Beyond behavioural changes, tuition fees might also affect students psychologically, which in turn could be negatively linked to individuals' academic performance. Previous research, for example, found a link between student loans and worse psychological outcomes (see e.g., Walsemann et al., 2015; Zhang, 2013). Taking a more nuanced perspective in their literature review, McCloud and Bann (2019) suggest that not the height of the debt but more subjective measures such as financial concerns are linked to lower mental health. Notably, such psychological deterioration can impede academic performance and success (see e.g., Bruffaerts et al., 2018; Eisenberg et al., 2009). As this would again deter individuals' signals, the likelihood of overeducation could be expected to rise.

Additionally, even if tuition fees neither triggered behavioural nor psychological responses, changes in educational performance and, hence, the likelihood of overeducation could still stem from university-level improvements. If tuition fees rise, they can spend more money to buy better teaching materials, hire more qualified teachers or create better learning spaces. Such improvements could translate into individuals' academic performance and might diminish graduates' risk of overeducation. Anyhow, as with the individual behavioural responses, it remains unclear whether EUR 500 per term and student are sufficient to finance such improvements and if the time interval is sufficiently large to realise them.

To summarise, expectations regarding the impact of tuition fees on the likelihood of overeducation are ambiguous. Consistent with theory, the likelihood of overeducation could be influenced, either by individuals choosing other occupations or by being hired for other jobs. The latter mechanism would require that tuition fees influence individuals' signals in the application process. However, the direction of this relationship is unclear, as tuition fees could deter or improve the academic performance of graduates, depending on the behavioural and psychological responses of individuals and the educational spending of universities.

3.3 Data and methodology

Examining how the introduction of tuition fees affects graduates' overeducation requires data that enables the identification of the type, timing, and place of the obtained degree, as well as the labour situation after graduation. For this reason, I use data from the SOEP (DIW, 2024; Goebel et al., 2019), which allows for the parallel assessment of individuals' labour market outcomes and academic careers. Its panel structure enables the identification of students across different periods, including the treatment period. Finally, it provides information on the socioeconomic background of individuals, which is crucial when evaluating policies that monetarily affect students.

3.3.1 Dependent variable

To define overeducation, I follow Clogg and Shockey (1984) and Verdugo and Verdugo (1989) and use a data-driven approach (MEAN measure) to calculate the required level of education for a job. This measure bears two major advantages in the context of the present study: Firstly, the variable can be implemented by only requiring information on the years of education and occupation of individuals, thereby avoiding any further reduction in sample size. Secondly, the variable allows an adjustment for time trends in education caused, for example, by educational inflation and changing occupational requirements. Since the SOEP contains observations on individuals from four different decades, a time-adjusted measure prevents false conclusions based on contextual inconsistencies (for a discussion, see e.g., Capsada-Munsech, 2019; Verhaest and Omey, 2010).

To apply this measure, the average level of education obtained by workers in the same occupational group is defined as the reference value for the level of required education. As proposed by Blásquez and Budría (2012), the three-digit ISCO classification offers a sufficiently narrow differentiation of occupations, while still allowing for a sufficient size of clusters. Within these clusters, individuals are defined to be overeducated if their years of education x_{it} exceed the average years of education $\overline{x_{ot}}$ in occupation o in survey wave t by more than one standard deviation σ_{ot}. Everyone whose attained education falls within the one standard deviation range around the mean is considered matched (reference category). More formally:

$$OE_{ito} = \begin{cases} 1 & \text{if } x_{it} > \overline{x_{ot}} + \sigma_{ot} \\ 0 & \text{if } \overline{x_{ot}} - \sigma_{ot} < x_{it} < \overline{x_{ot}} + \sigma_{ot} \end{cases} \tag{3.1}$$

3.3.2 Treatment

Individuals are defined to be treated if they obtained their degree in a state that charged fees in the respective period. Therefore, I first extract the year and month in which the degree was obtained from the data.[32] Furthermore, the definition of the treatment variable requires information on the state in which an individual obtained the degree. The dataset contains direct information on this for a subsample of individuals. For the remaining individuals, I look at the federal state of residence in the survey wave in which the degree was obtained.[33] Two critical assumptions are that individuals did not commute to university and that their indicated place of residence equals their true place of residence. If these assumptions were violated, individuals would be considered treated, although they did not pay tuition fees. Such inclusion of non-treated individuals in the treatment group would imply a non-differential misclassification. This would bias the estimated treatment effect toward zero. Hence, the estimates likely represent conservative lower bounds of the true effect (see also e.g., Bahrs and Siedler, 2019; Bietenbeck et al., 2023; Hübner, 2012). However, to properly account for this and as the pattern cannot be tested empirically, the effects are referred to as Intention to Treat Effects (ITT).[34]

Combining information on the timing of the degree and the federal state in which the degree was obtained, I can identify individuals affected by the tuition fee legislation. These individuals form the treatment group, whereas those who obtained a degree outside this period or the treatment states form part of the control group.

3.3.3 Covariates

The estimations include several covariates that are arguably determined before degree completion and related to both the likelihood of overeducation and of being treated. The first set contains demographic covariates, including age, sex, and migration background. Additionally, I include proxies for socioeconomic background. These cover, firstly, the education of one's father and mother as

[32]Direct information on this is available only for a subset of individuals. The remaining respondents are asked each year whether they received a degree since the previous year, and if so, what kind of degree they completed in which month.

[33]Note that this reduces the sample to those for whom direct information is available or to people who participated in the survey in the year in which they graduated, resulting in a small sample size.

[34]For a formalisation on how the "home bias" and spill-over effects threaten the identification of the true Average Treatment Effect on the Treated see Hübner (2012).

they may simultaneously impact the likelihood of studying and post-graduation labour market outcomes. Moreover, parents' educational background arguably proxies the financial resources of young adults (e.g., De Haan, 2011; Dickson et al., 2016; Sikhova, 2023). Secondly, the number of siblings is considered to proxy the distribution of economic resources within families (Downey, 1995). Regional variation is accounted for through fixed effects covering the federal state of residence and the federal state where the individual obtained the degree. Finally, I include survey year fixed effects to rule out time effects.[35]

3.3.4 Sample

Treatment is conditional on visiting a university or college, restricting the sample to graduates. Moreover, I exclusively keep individuals with valid information on the dependent variable. This implies that I drop unemployed or self-employed individuals. Additionally, individuals who pursue studies later in life, for example, in part-time or after having started to work in the labour market, are likely less affected by the monetary burden of tuition fees. Therefore, I restrict the sample to individuals aged 35 and under at the time of degree completion. Also, I exclude the lower 2.5% of the age-at-degree distribution, restricting the sample to individuals who completed their degrees at age 19 at the earliest, as it is uncommon to complete a tertiary degree at even younger ages. As the treatment is time-invariant within individuals, I cannot account for individual time-invariant unobserved heterogeneity. For this reason, the main sample is restricted to the first observation per individual after the degree was obtained. Finally, I make sure that all observed persons have obtained their degree at least one year before the respective survey wave. This avoids overeducation being detected in periods when people are still studying but are already working part-time alongside their studies.

Following this approach, I observe one to at most three treated individuals in Saarland, Hesse, and Hamburg. To avoid biases stemming from this pattern, I only consider the control observations for these federal states. As a result, the treatment effect is identified based on individuals who graduated in Bavaria, Baden-Württemberg, Lower Saxony, and North Rhine-Westphalia. All individuals with missing values are dropped.

[35] The year in which the degree was obtained is not included in the analysis due to its near-perfect correlation with the survey year ($\rho = 0.989^{**}$).

3.3.5 Descriptives

The sample composition is displayed in Table 3.1. Within the sample, 40% of the individuals are overeducated. This number differs from studies relying on German panel data (e.g., Bauer, 2002; Blásquez and Budría, 2012) due to the restriction to graduates, among whom overeducation is, by definition, most likely. Moreover, 11% of the respondents are treated.

The sample is well-balanced regarding sex, as 48% of the respondents are female. The individuals are born between 1953 and 1997, yielding an age range at the time of the survey from 20 to 43 years. Notably, the average age at degree completion is 25.35, aligning with administrative data from the Federal Statistical Office noting the average age at study completion (master's degree) to be 26.5 (Statistisches Bundesamt, 2024d). Individuals have between 0 to 7 siblings, with an average of slightly above 1. The vast majority of them (93%) are natives. 3% (4%) of the individuals have a direct (indirect) migration background. Regarding parental education, I observe 1% of fathers and mothers with no education, 32% (33%) with general education, 28% (41%) with secondary education, and 39% (25%) with upper secondary education degrees.

To identify the effect of tuition fees on the likelihood of overeducation, individuals in treatment and control states should be comparable before the policy change. Otherwise, the estimates might be influenced by factors independent of the actual intervention. For this reason, Table 3.2 compares the differences in means between treated and control states exclusively using the observations of individuals who obtained their degrees before the first tuition fee introduction in 2006.[36] This causes a reduced sample size compared to Table 3.1 and implies that Table 3.2 relies on control group observations exclusively. The results indicate no statistically significant differences in the majority of variables. Still, individuals in the treatment states tend to have a slightly higher number of siblings on average. While this difference is statistically significant, it remains relatively minor, with the average number of siblings being between 1 and 1.5 in both groups. Moreover, the proportion of individuals whose fathers attained secondary education is higher in the control states. This discrepancy may be due to the different school systems in the Federal Republic of Germany and the former

[36]Note that individuals are observed at least one year after graduation, such that 2006 would be expected to be the last year of observation in Table 3.2. However, for few individuals there is a time lack between the year in which they graduated and that in which they are observed for the first time after graduation. Hence, Table 3.2 covers the years 1985 to 2012.

Table 3.1: Summary statistics

	N	Mean	Min	Max
Overeducated	966	0.40	0.00	1.00
Treated	966	0.11	0.00	1.00
Female	966	0.48	0.00	1.00
Number of siblings	966	1.32	0.00	7.00
Birth year	966	1975.59	1953.00	1997.00
Age	966	26.79	20.00	43.00
Age at degree	966	25.35	19.00	35.00
No migrant	966	0.93	0.00	1.00
Direct migrant	966	0.03	0.00	1.00
Indirect migrant	966	0.04	0.00	1.00
Father educ.: no	966	0.01	0.00	1.00
Father educ.: general	966	0.32	0.00	1.00
Father educ.: secondary	966	0.28	0.00	1.00
Father educ.: upper secondary	966	0.39	0.00	1.00
Mother educ.: no	966	0.01	0.00	1.00
Mother educ.: general	966	0.33	0.00	1.00
Mother educ.: secondary	966	0.41	0.00	1.00
Mother educ.: upper secondary	966	0.25	0.00	1.00
Survey year	966	2002.37	1985.00	2021.00
Degree year	966	2000.94	1984.00	2020.00

Note: The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. SOEP weights applied.

German Democratic Republic, which constitutes a large portion of the control group (for a detailed comparison, see, e.g., Riphahn and Trübswetter, 2013).

3.3.6 Identification strategy

I estimate the following linear probability model for the quasi-experimental setting of the introduction of tuition fees in Germany to identify their ITT on the likelihood of overeducation (compare e.g., Bahrs and Siedler, 2019).

$$OE_i = \alpha_0 + \beta_1 treated_i + \beta_2 X_i' + \lambda_t + \gamma_e + \delta_s + \epsilon_i \tag{3.2}$$

OE_i captures overeducation for individual i. α_0 is the intercept. β_1 is the coefficient of interest and identifies the ITT of tuition fee exposure. $treated_i$ equals 1 if the individual obtained a degree in a state and in a year in which tuition fees were charged. X_i' is a vector covering the covariates for each individual i. λ_t, γ_e, and δ_s cover the fixed effects for the survey years, the federal states where the degree was obtained, and the federal state of residence at the time of the interview. ϵ_i is the error term.[37]

[37] All estimations are pursued with the software R (version 4.2.3).

Table 3.2: Differences in means pre-treatment

	Control states		Treatment states		T-test	
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Overeducated	0.450	0.498	0.392	0.489	-0.058	0.050
Female	0.436	0.497	0.358	0.480	-0.079	0.049
Number of siblings	1.166	1.096	1.357	1.141	0.191+	0.112
Birth year	1970.582	7.824	1969.947	7.068	-0.636	0.771
Age	27.399	4.439	27.575	4.130	0.175	0.456
Age at degree	25.896	4.594	25.986	4.019	0.090	0.440
No migrant	0.949	0.221	0.941	0.235	-0.007	0.025
Direct migrant	0.030	0.171	0.039	0.194	0.009	0.020
Indirect migrant	0.021	0.144	0.020	0.139	-0.002	0.015
Father educ.: no	0.005	0.073	0.001	0.032	-0.004	0.003
Father educ.: general	0.357	0.480	0.416	0.494	0.059	0.049
Father educ.: secondary	0.346	0.477	0.239	0.427	-0.107*	0.047
Father educ.: upper secondary	0.291	0.455	0.344	0.476	0.053	0.046
Mother educ.: no	0.004	0.064	0.006	0.076	0.002	0.004
Mother educ.: general	0.423	0.495	0.471	0.500	0.048	0.050
Mother educ.: secondary	0.341	0.475	0.357	0.480	0.016	0.047
Mother educ.: upper secondary	0.232	0.423	0.167	0.373	-0.065	0.042
Survey year	1997.982	5.800	1997.522	5.555	-0.460	0.589
Degree year	1996.478	5.655	1995.933	5.596	-0.545	0.594
Num. obs.	293		396			

Note: The sample is based on SOEP data on graduates from 1985 to 2012. Observations are restricted to the first observation per individual and to individuals who obtained their degrees before the introduction of tuition fees in 2006. The treatment states cover states that introduced tuition fees in 2006 or 2007. SOEP weights applied.

3.4 Results

3.4.1 Main results

Table 3.3 reports the ITT of being affected by tuition fees on the likelihood of overeducation from linear probability models. Column (1) estimates a baseline model including fixed effects for the survey year, federal state of residence, and federal state where the degree was obtained. Column (2) adds demographic covariates, and column (3) conditions on the proxies for socioeconomic background. Using the specification with fixed effects exclusively, the likelihood of overeducation is 11.9 percentage points larger for the treated group. The coefficient increases marginally to 12.8 percentage points once conditioned on the covariates. This corresponds to 28% of the pre-treatment mean of the control group (see Table 3.2). Table B.1 in the Appendix reports the results for the covariates. The likelihood of overeducation increases with the age of the graduate and marginally decreases with the number of siblings. All remaining covariates yield insignificant estimates.[38]

[38]The results are not sensitive to the inclusion of age at graduation as a covariate instead of age (see Table B.2), to the restriction of the sample to individuals who are employed part-time or full-time (see Table B.3), or to the use of the panel structure by including all repeated observations for the individuals covered in the main estimations (see Table B.4).

Table 3.3: ITT of tuition fees on the likelihood of overeducation

	(1)	(2)	(3)
Treated	0.119+	0.133*	0.128*
	(0.060)	(0.050)	(0.049)
λ_t, γ_e, δ_s	X	X	X
Demographics		X	X
Socioeconomic background			X
Num. obs.	966	966	966
R2	0.080	0.096	0.102
R2 Adj.	0.013	0.025	0.024

Note: Columns (1) to (3) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. Column (1) reports estimates including fixed effects for the survey year (λ_t), the federal state where the degree was obtained (γ_e), and the federal state of residence (δ_s). Column (2) adds further demographic covariates and column (3) conditions on proxies for the socio-economic status, respectively. Standard errors are clustered at the level of the state where the degree was obtained and reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.4.2 Robustness

Due to the small sample size and the data limitations, it is critical to ensure the robustness of the findings. Therefore, I employ a series of sensitivity tests in the following section.

Identification strategy

First, I conduct a placebo test (see e.g., Dwenger et al., 2012; Hübner, 2012) altering the treatment period (column (1) of Table B.5). Doing so ensures that the ITT presented in Table 3.3 does not capture differences between treatment and control regions that materialised independently of the policy change. Second, I re-estimate the base model using heteroskedasticity-robust standard errors as proposed in Wooldridge (2010) (column (2) of Table B.5). This addresses concerns that a small number of clusters may lead to overly precise standard errors in the main specification. Third, I estimate the base model using a sample of individuals who obtained their degrees between 2006 and 2014 (column (3) of Table B.5). While this shrinks the size of the control group, it enhances comparability between treated and control individuals by ensuring they attended university during the same period. Notably, this approach also accounts for the changes in the educational system brought about by the Bologna reform. Before the reform, students in Germany pursued *Magister*- or *Diplom*-degrees, which took four to six years to complete. In contrast to the two-tier system implemented with the Bologna reform, students did not face the option to exit early from the degree while still holding an academic qualification. Including earlier cohorts in the base model may therefore lead to a downward bias in the estimated tuition fee effect, as students affected by tuition fees could respond by exiting early with a Bachelor's degree. While this additionally supports the earlier notion of a conservative lower bound estimate, such early exits may be driven not only by the risk of overeducation but also by the financial burden of tuition fees and should therefore be tackled.[39] Fourth, linear probability models are not without criticism because predicted probabilities can exceed the range of 0 to 1. For this reason, column (4) of Table B.5 reports Average Marginal Effects (AME) from logit estimations, while column (5) applies a trimming approach based on Horrace and Oaxaca (2006). Fifth, column (6) of Table B.5 reports results from Difference-in-Differences (DiD) estimations to explicitly test the common trends assumption.[40] In all cases, results align qualitatively with the base results.

Matching and weighting

The identification of the effect of tuition fees on overeducation is threatened if treated individuals differ from their counterfactual in factors other than treatment assignment. While structural differences based on region are partially tackled in the placebo test contained above, I also use matching and weighting methods to explicitly control for differences in observables. Column (7) of Table B.5 shows results from exact matching, where treated and control units are exactly matched on confounders, only allowing for differences in treatment assignment. This procedure reduces the number of observations to 145. Column (8) uses propensity score weighting based on treatment probabilities estimated via a Generalised Linear Model (GLM). Column (9) uses weights from the SuperLearner algorithm, which combines multiple machine-learning models for improved accuracy. While the estimations lack precision in the exact matching procedure, they are qualitatively unaffected in all cases.[41]

[39]Note that there is no significant difference in the years of education between the treatment and control group. The average years of education obtained are 15.76 in the control, and 15.99 in the treatment group. This additionally limits concerns that treated individuals drop out of university earlier at a higher rate.

[40]The estimation equation is $OE_i = \alpha_0 + \beta_1 treat_e + \beta_2 post_y + \beta_3 treat_e \times post_y + \beta_4 X_i' + \lambda_t + \gamma_e + \delta_s + \epsilon_e$ where $treat_e$ equals one if the degree was obtained in a fee-charging state e and $post_y$ captures whether it was obtained after fee introduction in the respective state. δ_3 is the coefficient of interest covering the Difference-in-Difference estimator. To account for the stepwise abolition of fees in the federal states, the sample drops all individuals who obtained their degrees after the fees were again abolished.

[41]Exact matching is implemented using the *MatchIt*-package based on Ho et al. (2011). All weighting techniques are performed with the *WeightIt*-package by Greifer (2023).

Omitted variable bias

As the sample includes the first observation per individual after treatment only, the period between treatment assignment and the transition to the labour market is usually short. Therefore, I cannot rule out that local labour market factors could simultaneously impact the decision to move to a state for studies and the likelihood of overeducation. To account for such potential omitted variable bias, I use additional information from the *Indikatoren und Karten zur Raum- und Stadtentwicklung* (INKAR) database, providing a broad range of statistics on the regional level as well as from the Federal Statistical Office. To capture the impact of the local labour market and economic conditions, I extract information on the local unemployment rate and GDP (columns (1) and (2) of Table B.6). Additionally, I use the annual number of enrolled students per 1,000 inhabitants within each state to capture educational capacity (column (3) of Table B.6). Finally, the introduction of tuition fees implied a monetary burden to students but increased financial resources of the charging states for educational purposes. This might have ameliorated educational quality as formulated in Section 3.2. In turn, this could affect individuals' choice to study in a fee-charging state as well as their job prospects post-graduation. Thus, column (4) of Table B.6 controls for the expenditure on education in each federal state, normalised by the number of students. Including these additional regional indicators does not affect the results qualitatively.

3.4.3 Selection concerns

Despite the provided robustness checks, the results presented in Table 3.3 may still be subject to selection bias. As outlined earlier, the data enable a treatment definition based on direct information about the state in which the degree was obtained or the state of residence when the degree was earned. Still, identification may not be perfectly accurate if, for example, individuals moved between fee-charging and control states without adequately indicating their new place of residence in the survey.[42] Dwenger et al. (2012) provide an in-depth discussion of this issue, noting that internal migration primarily occurs within West or East German federal states, with minimal migration between the two regions. With this in mind, they suggest restricting the control group to East German

[42]Note that in Germany, people who change their place of residence are generally obliged to register their new address within two weeks. However, students who move to study could decide to register their new home as a secondary residence and thus keep their parents' house as their official main residence.

states, none of which introduced tuition fees. Still, applying this restriction reduces the sample drastically. For this reason, I adopt an alternative strategy by incorporating data on internal migration provided by INKAR since 1995. Internal migration captures the net flow of individuals moving in and out of a federal state, while migration for educational purposes focuses specifically on the mobility of individuals aged 18 to 25. Negative values of the indicators indicate a net outflow, and positive values signal a net inflow. Based on this, I create dummy variables that take the value of 1 if the indicator is larger than or equal to 0, and 0 otherwise. While these do not directly approach individual-level selection, they reflect variation in overall internal migration patterns across federal states. Thus, they can be interpreted as a proxy for the average likelihood of mobility associated with living in a particular federal state in a certain year. States that experience a steady outflow of individuals could offer structurally weaker labour markets or fewer higher education opportunities, which could both increase the probability of students leaving the state and raise the likelihood of overeducation after graduation. By including this proxy, I partially account for potential biases arising from differential migration behaviour. Also, these measures make it possible to consider internal migration patterns but prevent a large reduction in the sample size. The results remain qualitatively unaffected (see Table B.7).

A second selection concern is the possibility that individuals who usually would have considered applying to university decide not to study or to delay their application due to the introduced tuition fees. In this case, the results discussed in Table 3.3 would be subject to sample selection bias, as these individuals would no longer appear in the graduate dataset. As this concern cannot be approached using the SOEP data, I analyse administrative data on first-year students provided by the Federal Statistical Office in the following (Statistisches Bundesamt, 2024b). The observation period spans from the winter term of 1998/99 to the winter term of 2023/24 providing 400 federal state-year observations ($25 \times 16 = 400$). If these data revealed substantial deviations in student enrolment trends between treated and control states, it would suggest the presence of an outselection bias. Given that treated and control states differ in the total number of higher education institutions and students, I compute the year-over-year percentage change in first-year student enrolments for each treated state and, collectively, for the control states. Figure B.1 and Figure B.1 illustrate these trends over time.

For Lower Saxony, Hamburg, and Hesse (Figure B.1d, Figure B.1e, Figure B.1f), the trajectories align closely with those of the control states, particularly around the period when tuition fees were first implemented. However, in Bavaria, Baden-Württemberg, Saarland, and North Rhine-Westphalia (Figure B.1a,

Figure B.1b, Figure B.1c, Figure B.1g), a slight divergence emerges during the initial year of tuition fee introduction, with a noticeable gap developing between the treated and the control states. Nonetheless, in all cases, the percentage change in first-year student enrolments eventually surpasses that of the control group, suggesting that any initial decline relative to the control states was offset by a subsequent increase.

To test this more thoroughly, Figure B.2 estimates an event study approach evaluating the effect of fee introduction on the number of students following Sun and Abraham (2021). While the dynamic analysis reveals a small reduction in the number of students in the first period, the number of students in the treatment states subsequently exceeds that in the control states, aligning with the descriptive evidence. Importantly, these estimates are never statistically significant at the 10%-level, indicating no long-term effect on the number of students (for supportive evidence, see e.g., Bruckmeier et al., 2013; Havranek et al., 2018).

The above-presented robustness and selection analyses support the validity of the main results. However, concerns regarding a potential selection on unobservables remain untackled. For this reason, I additionally use the approach proposed by Oster (2019). This approach allows for the estimation of the degree of selection on unobservables relative to observables that would be necessary to reduce the base coefficient (β_1 in Equation 3.2) to zero, $\delta*$.[43] Absolute values of $|\delta*| < 1$ hereby hint at problems with selection on unobservables, while larger values limit such concerns. Applying this approach, $\delta*$ is estimated to be -17.1. This indicates that selection on unobservables would need to be both extremely strong and act in the opposite direction of selection on observables to nullify the effect.[44] This extends the above presented evidence and furthermore limits concerns about selection bias in the base results.

3.4.4 Heterogeneity

After ensuring the robustness of the presented findings and limiting concerns regarding selection bias, it remains open whether one of the mechanisms outlined in Section 3.2 can be observed in the data. While not all of them can be

[43]The estimates are performed using the *"robomit"* package of Schaub (2025). Following Oster (2019) $Rmax$ is estimated based on the *R-squared* of the estimation including all covariates multiplied by 1.3. $\beta*$ is set equal to zero.

[44]For detailed explanations of the interpretation of negative $\delta*$ as well as for further recent studies reporting such, see e.g., Aldén and Hammarstedt (2016), Graham et al. (2017), Lee and Ohtake (2021), Luo (2022).

tested with the data at hand, the following heterogeneity analyses enable me to approach some of them. These heterogeneity analyses use the following form to investigate if the ITT of tuition fees on the likelihood of overeducation is particularly pronounced among certain subgroups:

$$OE_i = \alpha_0 + \beta_1 treated_i + \beta_2 \kappa_i + \beta_3 treated_i \times \kappa_i + \beta_4 X_i' + \lambda_t + \gamma_e + \delta_s + \epsilon_i \quad (3.3)$$

where κ_i denotes the indicators for diverging sources of heterogeneity across subgroups. The coefficient β_3 is the main coefficient of interest and captures this heterogeneity in the ITT of tuition fees on the likelihood of overeducation.

Occupational choice

First, one way in which tuition fees may influence the likelihood of overeducation is through systematic differences in individuals' career choices. As mentioned in Minicozzi (2005) and Gervais and Ziebarth (2019), unsuitable career choices may result from an accelerated transition into the labour market after graduation to cover the financial burden accumulated during studies. Molina and Rivadeneyra (2021) already provided direct evidence on the link between tuition fees and occupational choice and reported that individuals who were no longer affected by tuition fees selected into high-skill occupations at higher rates. If this translated to the German setting, and treated individuals selected into low-skill occupations at higher rates, the ITT of tuition fees on the likelihood of overeducation could be expected to be larger among individuals working in low-skill jobs. To identify this heterogeneity, I use information on individuals' occupations and distinguish the respective ISCO codes into high- and low-skill occupations, where the first three ISCO major groups (managers, professionals, associate professionals/technicians) represent the high-skill occupations (International Labour Office, 2012).[45]

Figure B.3 presents the respective results. As expected and in line with the definition of overeducation, individuals in high-skill occupations report a lower likelihood of overeducation. Also, aligning with the base results (see Table 3.3), the ITT of being exposed to tuition fees is positive and significant. Moreover, as expected, the interaction term between working in a high-skill occupation and being treated is negative. This would indicate that, indeed, the exposure to

[45] Table B.8 shows the ITT of exposure to tuition fees on the likelihood of working in a high-skill occupation. The coefficient is negative, which aligns with the expectations, but is small and insignificant.

tuition fees increases the likelihood of overeducation for individuals working in low-skill occupations, while being less influential among individuals in high-skill occupations. Still, this evidence is only suggestive due to two major reasons: The coefficient of the interaction term is insignificantly estimated, limiting firm conclusions about heterogeneity across occupational choice. In addition, this heterogeneity analysis attempts to explicitly take into account career choice after treatment. However, the differences in the ITT between groups may not only be due to occupational choice, but could also capture certain systematic differences within occupations. Although the base coefficient of the occupation dummy partially accounts for such differences, it might still not encompass all relevant variation.

Socioeconomic background

Second, and as formulated in Section 3.2, individuals could potentially be affected differently by tuition fees depending on their socioeconomic background and, for example, the monetary support received from their parents. Aligning with this supposition, Neill (2009) shows that enrolment rates are most strongly affected among individuals whose parents obtained medium levels of education in Canada. Moreover, using the German case, Bahrs and Siedler (2019) find that the negative effect of tuition fees on educational intentions is stronger among individuals from low-income households. While the SOEP does not provide information on parental income or financial support by parents for a sufficient number of observations, I again employ the education attained by parents as a proxy for socioeconomic background. To maintain interpretability in the interaction setting, I define dummies equaling one if the individuals' father (mother) obtained upper secondary education. Moreover, I create a combined indicator as the sum of these dummies ranging between zero – when neither parent obtained upper secondary education – and two – when both obtained upper secondary education.[46]

Figure B.4a and Figure B.4b report the results separately for the education obtained by mothers and fathers. In both cases, it does not matter for individuals' likelihood of overeducation if the father (mother) obtained upper secondary education, while the ITT of being exposed to tuition fees is positive and significant as before. As indicated by the insignificant interaction terms,

[46]To avoid multicollinearity issues, these heterogeneity checks do not include the categorical assessment of education obtained by mother and father as separate covariates, but solely include the interaction term and the newly created dummies.

this positive ITT does not vary by the educational attainment of either fathers or mothers. Figure B.4c uses the categorical indicator for the joint education obtained by fathers and mothers. As before, whether one or both parental parts obtained upper secondary education is not statistically related to the likelihood of overeducation, while the treatment indicator remains positive and statistically significant. But, in contrast to the previous estimations, I observe a significantly lower likelihood of overeducation for treated individuals with one parental part owning upper secondary education. This aligns with the expectations that a higher socioeconomic background could buffer the influence of the tuition fee introduction. But, in contrast to that, the ITT of tuition fees on the likelihood of overeducation for individuals with two parents owning upper secondary education does not significantly differ from the reference group.

Employment during studies

Third, as outlined in Section 3.2, the likelihood of overeducation might be affected by tuition fees if individuals worked besides their studies to cover the additional expenses. But, theoretical expectations about the direction of the effect were ambiguous. This is, as working besides studies could harm individuals' performance, but could also improve individuals' signals in terms of working experience, potentially reducing the likelihood of overeducation (Diem and Wolter, 2014; Verhaest and Omey, 2010). If the latter of both materialised, one would expect the ITT of tuition fees to be most pronounced among individuals who were not employed. This is because their lack of working experience would disadvantage them in the application process. In contrast, if the performance and signal of individuals were harmed by working besides studying due to reductions in the time available for studying, the opposite, hence, a more pronounced ITT among individuals who worked besides their studies, would be expected. Using information on the employment status of individuals one year before their degree completion, I distinguish between graduates who were not employed during their studies and those who were in either full-time, part-time, or marginal employment.

Figure B.5 presents the results. Notably, having worked besides the studies is significantly linked to a higher likelihood of overeducation. This aligns more with the perspective that working besides studying reduces the time available and might harm academic outcomes and thus, the transition to the labour market. Through adding the interaction of the treatment indicator and the employment variable, the ITT of tuition fees on the likelihood of overeducation becomes insignificant, though remaining positive. Also, the interaction term is positive but imprecisely estimated. Hence, although the evidence aligns more with the perspective that employment besides studies is harmful instead of helpful, the analysis does not provide evidence that the treatment effect is driven by this group.[47]

In addition to the previous sources of heterogeneity, two further aspects should be considered: The duration of the treatment exposure and the federal states. Importantly, these two aspects are inevitably linked to the treatment, as, for example, only treated individuals can differ in the duration of their exposure. For this reason, the interaction approach presented in Equation 3.3 is not applicable here. For the following two heterogeneity analyses, I hence estimate the interaction regression without the baseline effects for the treatment indicator or the source of heterogeneity due to multicollinearity with the interaction term. The regression equation hence reads as follows:[48]

$$OE_i = \alpha_0 + \beta_1 treated_i \times \kappa_i + \beta_2 X_i' + \lambda_t + \gamma_e + \delta_s + \epsilon_i \tag{3.4}$$

Treatment duration

Although the data do not allow for the identification of the exact duration of the degree, they enable me to approximate whether individuals were at least treated for the full period of a standard Bachelor's degree. Notably, this approach allows to account for two questions at once. On the one hand, it proxies differences in the total amount of tuition fees paid. On the other hand, it assesses differences in the timing of treatment, as people who already started their studies with tuition fees may have been affected differently by the tuition fees than those who started without tuition fees and were surprised by the introduction during their studies. As a standard Bachelor's degree takes at least six terms in most cases, every

[47]To validate that the treatment indicator loses significance due to the introduction of the interaction term, I re-estimated the baseline regression on the reduced sample used in Figure B.5. Moreover, I also estimated the baseline regression introducing the employment measure as a covariate without interacting it with the treatment. In both cases, the ITT of exposure to tuition fees is positive and significant. Additionally, as the vanishing significance of the treatment variable could raise the question of whether the treatment effect is mediated by employment besides studying, I regressed the employment indicator on the treatment. The coefficient is close to zero and insignificant, and does, thus, not indicate that the ITT might work through an impact of tuition fees on the employment of students. The estimations are contained in Table B.9.

[48]Importantly, due to the changes in the regression function, the interpretation of the plot changes. In Figure B.3 to Figure B.5, the interaction terms indicate whether the ITT of tuition fees statistically differs between groups. In contrast, in Figure B.6 and Figure B.7, overlapping confidence intervals indicate non-statistically significant differences between the groups, while non-overlapping confidence intervals hint at statistically significant differences.

individual who graduated a minimum of six terms after the introduction of the tuition fees but before their abolition in the respective state can be assumed to be fully treated. Using this approach, 46% of the previously treated individuals were exposed to tuition fees for the full duration of a standard undergraduate programme.

Figure B.6 shows the results from interacting the treatment indicator with the dummy indicating if individuals were treated for at least six terms. Notably, the coefficient for those considered fully treated is slightly larger in magnitude than the base coefficient (see Table 3.3) and is statistically significant. In contrast, the coefficient for those who were treated less than these six terms is quantitatively smaller and imprecisely estimated. These results align conceptually with the anticipation that individuals with longer treatment duration and higher costs report a stronger ITT of tuition fees on the likelihood of overeducation. However, the confidence intervals overlap, preventing the identification of a significant difference in the ITT between both groups.

Federal states

Fifth, differences could arise, for example, if financial resources gathered through the fees were used in different ways by the federal states. They could invest them directly in improving the quality of teaching, or use them for infrastructural investments, while other federal states might only make such investments later. In addition, structural differences between the federal states could induce heterogeneity in the influence of tuition fees on the likelihood of overeducation. While the robustness section already approached whether the main results persist when conditioning for educational expenses on the state-level, Figure B.7 plots interactions of the treatment indicator and the federal states, revealing the ITT for each of them. The abbreviations on the x-axis indicate the federal states such that LS = Lower Saxony, NRW = North Rhine-Westphalia, BW = Baden-Württemberg, and BY = Bavaria.[49] All remaining states are part of the control group for which the interaction term yields zero. The positive ITT reported in Table 3.3 can be observed in Lower Saxony, Baden-Württemberg and Bavaria, but is insignificant in North Rhine-Westphalia. Albeit this, the analysis provides little evidence that structural state differences or differences in the educational system would drive the results.

[49]As discussed in Section 3.3.4, the ITT cannot be identified for Hesse, Hamburg, and Saarland.

In sum, the heterogeneity analyses provide suggestive evidence that the treatment may be less influential among individuals who chose high-skill occupations post-graduation, who have one parent with upper secondary education and for those who were treated for a shorter period. But, in most cases, the data do not allow to identify significant differences between groups. For this reason, although results align with the suppositions in the majority of cases, the evidence is interpreted with caution.

3.4.5 Dynamic analysis

Due to the sample restrictions outlined above, the ITT of tuition fees on the likelihood of overeducation likely applies exclusively to individuals' first jobs post-graduation. Hence, it remains open whether this is a short- or long-term effect. Theoretical considerations related to Career Mobility Theory argue that selecting an overeducated position post-graduation may serve as a temporary state, allowing one to gather work experience to achieve a matched position afterwards (see e.g., Sicherman, 1991; Sicherman and Galor, 1990). The empirical literature, however, reports contradicting results. Several papers report evidence aligning with Career Mobility Theory as overeducated individuals tend to participate more in training, are more likely to switch jobs, or report higher wage growth rates over time (e.g., Grunau and Pecoraro, 2017; Roller et al., 2020; Sicherman, 1991). In contrast, other studies identify overeducation as a "trap," increasing the probability of staying overeducated in the long run (e.g., Baert et al., 2013; Blásquez and Budría, 2012; Büchel and Mertens, 2004; Kiersztyn, 2013; Wen and Maani, 2019). As reported in Table 3.3, tuition fees increase the likelihood of overeducation among treated individuals. If this effect was only short-term, potential negative consequences for individuals and society would be limited. However, if the increased likelihood of overeducation persisted for several years, what would align more with the trap-perspective, this would not only imply an inefficient distribution of formal qualifications in the labour market. Even more, it might also expose individuals to certain disadvantages attributed to overeducation such as the overeducation-pay-penalty or potentially a lower job satisfaction (cf., Allen and van der Velden, 2001; Belfield, 2010; Caroleo and Pastore, 2018; Duncan and Hoffman, 1981; Peiró et al., 2010; Verdugo and Verdugo, 1989).

To investigate whether the ITT of being exposed to tuition fees on the likelihood of overeducation persists, I make use of the panel structure of the SOEP in the following. For each individual, I define the years since the degree completion as the distance between the year in which the degree was obtained and the respective survey year. Using this, I estimate Equation 3.5 where β_1 captures the unique ITT at each distinct time-interval since the degree completion, i.e., the ITT of being treated 1 (2, 3, ...) years post-graduation, while β_2 controls for general trends in the likelihood of overeducation with increasing years since graduation.[50]

$$OE_{it} = \alpha_0 + \beta_1 treated_i \times yearssince_{it} + \beta_2 yearssince_{it} + \beta_3 X_i' + \lambda_t + \gamma_e + \delta_s + \epsilon_i \quad (3.5)$$

To estimate this regression, the following analysis uses all observations of the individuals contained in the data after they completed their degrees.[51] To ensure a sufficient number of observations in both the treatment and control groups across all values of the variable capturing years since degree completion, the analysis is restricted to individuals observed within ten years post-graduation.[52] This results in a sample of 5,889 observations of 965 individuals.[53] Figure 3.2 shows the results for the interaction effect, β_1. While the point estimate is positive and substantial in each of the periods, it is precisely estimated from the fourth interval onward. Moreover, as the time since graduation increases, the point estimate gradually strengthens until around the eighth year and declines slightly afterwards.[54]

In sum, the evidence from the dynamic analyses is more in favour of a trap perspective. The findings indicate that individuals exposed to tuition fees are more likely to experience overeducation even up until ten years after graduation. This suggests enduring consequences of the introduction of tuition fees on labour market outcomes, particularly job match quality, for affected individuals.

[50] Note that years since degree completion can be defined for individuals independent of their treatment assignment, while the interaction term is directly linked to the treatment. For this reason, the treatment cannot be included as separate covariate.

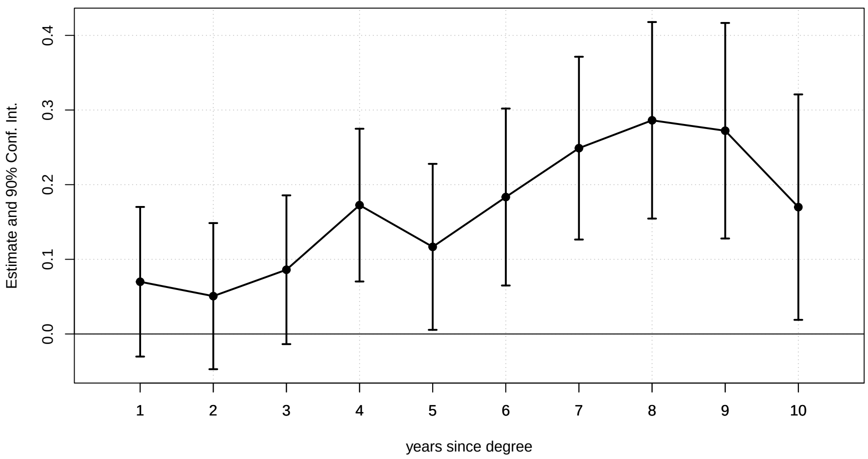
[51] Note that the use of the panel structure and the inclusion of all observations for the years surveyed in Table 3.3 extends the survey years to 2022.

[52] One individual included in the baseline estimations was only observed ten years after graduation and is therefore dropped from the sample.

[53] The results presented in Table B.4 remain qualitatively the same when using this restricted sample.

[54] Figure B.8 in the Appendix additionally plots the baseline coefficients for years since degree completion (β_2) revealing insignificant estimates in all periods.

Figure 3.2: ITT of tuition fees on the likelihood of overeducation: Dynamic analysis



Note: The graph displays the ITT by the years since degree completion based on the interaction of the treatment indicator and the years since degree completion. The regression is based on Equation 3.5 and estimates the ITT conditional on general trends in overeducation with rising years since degree completion. The years since degree completion are not significantly related to the likelihood of overeducation (see Figure B.8). The sample is based on SOEP data using all observations of graduates referred to in Table 3.3 survey between 1985 to 2022. It is restricted to individuals observed within ten years following degree completion. All covariates and fixed effects of Table 3.3 are included.

3.5 Conclusion

Research on the effect of higher education costs is extensive in contexts where tuition fees are common (e.g., Castleman and Long, 2016; Denning, 2019; Dynarski, 2008; Dynarski et al., 2021; Fack and Grenet, 2015; Page et al., 2019; Solis, 2017). In contrast, evidence on rises in tuition fees from countries with historically low or non-existent tuition fees is limited and focuses on educational performance and enrolment (see, e.g., Badillo-Amador and Vila, 2013; Barr, 2019; Bruckmeier et al., 2013; Bruckmeier and Wigger, 2014; Elliott and Soo, 2013; Hübner, 2012; Kane, 2007; McPherson and Schapiro, 1991; Minor, 2023; Neill, 2009; Riphahn, 2012; Savoca, 1990). Consequently, there is a lack of evidence on the impact of tuition fees on post-graduate outcomes, particularly in systems that historically do not rely on such fees. This study provides the first empirical evidence linking higher education costs to post-graduation outcomes in the German setting. Specifically, it exploits the quasi-experimental setting of the introduction of tuition fees in several federal states between 2006 and 2014 on the likelihood of overeducation post-graduation.

Using data on graduates from the SOEP for the period 1985 to 2021 reveals that exposure to tuition fees significantly increases the likelihood of overeducation. This result is not sensitive to the empirical specification. Moreover, concerns regarding potential selection issues can be limited empirically using additional data from INKAR and the Federal Statistical Office, as well as by implementing the approach suggested by Oster (2019). Heterogeneity analyses hint at potential differences based on individuals' occupational choice, socioeconomic background and treatment duration. In dynamic terms, the findings suggest a persistent effect of tuition fees on the likelihood of overeducation. More precisely, the positive link persists in the ten years post-graduation.

Albeit these findings hint at severe and long-lasting impacts of tuition fees on individuals' labour market outcomes, several limitations of the study should be acknowledged. Identifying a sufficiently large sample of treated and control individuals imposes challenges, as the analysis requires participants to be tertiary graduates with identifiable graduation dates, all while being observed after entering the labour market. These requirements, combined with the sample selection criteria, yield a comparably small sample size, potentially limiting generalisability. However, comparisons with similar studies using the same dataset indicate that the sample size is within a typical range (see, e.g., Bahrs and Siedler, 2019). Furthermore, this study focuses on the ITT rather than the average treatment effect (ATE), as it cannot be ruled out that persons in the treated group are actually untreated due to a discrepancy between their assumed and actual place of study. Although undesirable, such non-differential misclassification would introduce a downward bias in the coefficients. This implies that the coefficients are conservative lower bounds and likely even underestimate the true effect (e.g., Bahrs and Siedler, 2019; Bietenbeck et al., 2023; Hübner, 2012). Finally, the ITT is identified by four out of seven states that introduced tuition fees in the respective period due to data limitations. It should be noted that prior studies reported heterogeneity in effect directions, particularly regarding Hamburg and Saarland (Bruckmeier et al., 2013). Therefore, it is emphasised that the evidence obtained in this study cannot be generalised to these states but needs to be supplemented with further results.

Despite these limitations, the insights gained through this study are informative for the debate about the gains and costs of tuition fees. In particular, the study supports the thesis that the introduction of tuition fees cannot only influence the decision for or against higher education but also the individual's outcomes after graduation. While this may harm individuals' careers (Baert et al., 2013), it could also impede labour earnings (Allen and van der Velden, 2001; Dolton and Vignoles, 2000; Duncan and Hoffman, 1981; Rumberger, 1987) and could prevent an efficient allocation of the labour force.

3.6 Appendix B

3.6.1 Main results

Table B.1: ITT: Extended

	(1)	(2)	(3)
Treated	0.119+	0.133*	0.128*
	(0.060)	(0.050)	(0.049)
Age		0.015***	0.017***
		(0.003)	(0.003)
Female		0.041	0.045
		(0.037)	(0.038)
Direct migrant		0.026	0.015
		(0.072)	(0.082)
Indirect migrant		-0.061	-0.066
		(0.066)	(0.063)
Number of siblings			-0.022+
			(0.011)
Father educ.: general			-0.006
			(0.321)
Father educ.: secondary			-0.036
			(0.296)
Father educ.: upper secondary			-0.044
			(0.305)
Mother educ.: general			-0.114
			(0.311)
Mother educ.: secondary			-0.024
			(0.300)
Mother educ.: upper secondary			-0.025
			(0.300)
λ_t, γ_e, δ_s	X	X	X
Num. obs.	966	966	966
R2	0.080	0.096	0.102
R2 Adj.	0.013	0.025	0.024

Note: Columns (1) to (3) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. Column (1) reports estimates, including fixed effects for the survey year (λ_t), the federal state where the degree was obtained (γ_e), and the federal state of residence (δ_s). Column (2) adds further demographic covariates including age, sex and migration background. Column (3) furthermore conditions on proxies for the socio-economic status by including the number of siblings, and dummies, differentiating the education attained by mothers and fathers, respectively. Standard errors are clustered at the level of the state where the degree was obtained. Standard errors are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.2: ITT: Age at degree

	(1)	(2)	(3)
Treated	0.119+	0.127*	0.122*
	(0.060)	(0.053)	(0.052)
Age at degree		0.014***	0.015***
		(0.003)	(0.003)
λ_t, γ_e, δ_s	X	X	X
Remaining demographics		X	X
Socioeconomic background			X
Num.Obs.	966	966	966
R2	0.080	0.094	0.100
R2 Adj.	0.013	0.023	0.022

Note: Columns (1) to (3) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. Column (1) reports estimates, including fixed effects for the survey year (λ_t), the federal state where the degree was obtained (γ_e), and the federal state of residence (δ_s). Column (2) adds further demographic covariates including age at degree (in exchange for age), sex and migration background. Column (3) furthermore conditions on proxies for the socio-economic status by including the number of siblings, and dummies, differentiating the education attained by mothers and fathers, respectively. Standard errors are clustered at the level of the state where the degree was obtained. Standard errors are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.3: ITT: Only full- and part-time employed

	(1)	(2)	(3)
Treated	0.135+	0.148*	0.141*
	(0.075)	(0.067)	(0.064)
λ_t, γ_e, δ_s	X	X	X
Demographics		X	X
Socioeconomic background			X
Num.Obs.	770	770	770
R2	0.085	0.100	0.111
R2 Adj.	0.001	0.011	0.014

Note: Columns (1) to (3) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1985 to 2021 working in full- or part-time. Observations are restricted to the first observation per individual. Column (1) reports estimates, including fixed effects for the survey year (λ_t), the federal state where the degree was obtained (γ_e), and the federal state of residence (δ_s). Column (2) adds further demographic covariates including age, sex and migration background. Column (3) furthermore conditions on proxies for the socio-economic status by including the number of siblings, and dummies, differentiating the education attained by mothers and fathers, respectively. Standard errors are clustered at the level of the state where the degree was obtained. Standard errors are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.4: ITT: Using panel structure

	(1)	(2)	(3)
Treated	0.094*	0.121*	0.121*
	(0.046)	(0.048)	(0.048)
λ_t, γ_e, δ_s	X	X	X
Demographics		X	X
Socioeconomic background			X
Num.Obs.	8654	8654	8654
R2	0.042	0.045	0.054
R2 Adj.	0.035	0.037	0.046

Note: Columns (1) to (3) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1985 to 2021. The sample is based on the same individuals observed in Table 3.3 but includes all repeated observations for these individuals. Column (1) reports estimates, including fixed effects for the survey year (λ_t), the federal state where the degree was obtained (γ_e), and the federal state of residence (δ_s). Column (2) adds further demographic covariates including age, sex and migration background. Column (3) furthermore conditions on proxies for the socio-economic status by including the number of siblings, and dummies, differentiating the education attained by mothers and fathers, respectively. Standard errors are clustered at the level of the state where the degree was obtained interacted with the individual identifier. Standard errors are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p \leq 0.001$.

3.6.2 Robustness

Table B.5: Robustness: Identification strategy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Identification & Estimation						Matching & Weighting		
	Placebo	SE	2006 - 2014	Logit	Trimming	DiD	Exact	PSM	SuperLearner
Placebo	-0.003								
	(0.029)								
Treated/ treated $\times$ post		0.128+	0.237***	0.131*	0.138*	0.109+	0.283	0.116+	0.127+
		(0.076)	(0.055)	(0.060)	(0.050)	(0.063)	(0.223)	(0.062)	(0.062)
Post						-0.096			
						(0.193)			
Treat						-0.540			
						(1.000)			
λ_t, γ_e, δ_s	X	X	X	X	X	X	X	X	X
Demographics	X	X	X	X	X	X	X	X	X
Socioeconomic background	X	X	X	X	X	X	X	X	X
Num. obs.	966	966	244	966	956	928	145	966	966
Effective obs.								927.93	931,71
R2/ McFadden Pseudo R2	0.099	0.102	0.137	0.021	0.103	0.109	0.328	0.116	0.098
R2 Adj.	0.021	0.024	-0.070		0.025	0.027	-0.240	0.039	0.020

Note: Columns (1) to (3) and (5) to (9) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. Column (4) estimates a logit model. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. Column (1) reports estimates using a placebo treatment using different treatment periods. Column (2) uses heteroskedasticity robust standard errors (see Wooldridge, 2010). Column (3) restricts the sample to individuals who obtained their degrees between 2006 and 2014. Column (4) presents Average Marginal Effects (AME) from logit specifications. Column (5) trims the sample following the suggestion of Horrace and Oaxaca (2006) to account for biases in the linear probability specification. Column (6) applies a classical Difference-in-Difference (DiD) approach where the treatment equals 1 if the individual obtained the degree in one of the fee-charging states and the post-variable identifies degrees obtained after introducing fees. Individuals who obtained degrees after the abolition of fees in the fee-charging states are dropped to avoid biases. Columns (7) to (9) report estimates based on the baseline identification strategy, applying weights to correct the covariate balance between the treated and control groups. Column (7) uses exact matching methods. Column (8) applies propensity score matching (PSM) based on a parametric GLM estimating the predicted probability of being in the treatment or the control group. Column (9) uses propensity score weights estimated with the SuperLearner algorithm that combines multiple machine learning models to improve prediction accuracy and flexibility. The covariates used to estimate the weights are the demographic confounders. All covariates and fixed effects are included. Standard errors are clustered at the level of the state where the degree was obtained in columns (1) and (3) to (9) and are reported in parentheses. McFadden Pseudo R2 reported in column (4). + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.6: Robustness: Omitted variables

	(1) Unemployment	(2) GDP	(3) # students/1,000 inhabitants	(4) Expenditure/student
Treated	0.150**	0.117*	0.150**	0.167**
	(0.050)	(0.055)	(0.048)	(0.042)
Unemployment rate	0.007			
	(0.015)			
GDP		0.000		
		(0.000)		
# students (norm.)			0.001	
			(0.007)	
Expenditures (norm.)				-0.013
				(0.017)
λ_t, γ_e, δ_s	X	X	X	X
Demographics	X	X	X	X
Socioeconomic background	X	X	X	X
Num. obs.	667	795	738	618
R2	0.111	0.112	0.114	0.117
R2 Adj.	0.017	0.028	0.026	0.015

Note: Columns (1) to (4) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on university graduates from 1998 to 2021 in column (1), from 1995 to 2021 in column (2), from 1996 to 2021 in column (3), and from 1999 to 2021 in column (4). Observations are restricted to the first observation per individual. Column (1) adds the local unemployment rate. Column (2) contains the GDP, and column (3) additionally conditions on the local number of students per 1,000 inhabitants. Column (4) adds higher education expenditure in EUR 1,000, standardised to the number of students in the respective federal state. The data on the local unemployment rate, GDP, and the number of students per 1,000 inhabitants are extracted from INKAR. Data on the absolute number of students in each federal state, which are used for standardisation, and on expenditure on higher education in each federal state come from the Federal Statistical Office. All covariates and fixed effects are included. Standard errors are clustered at the level of the state where the degree was obtained and are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

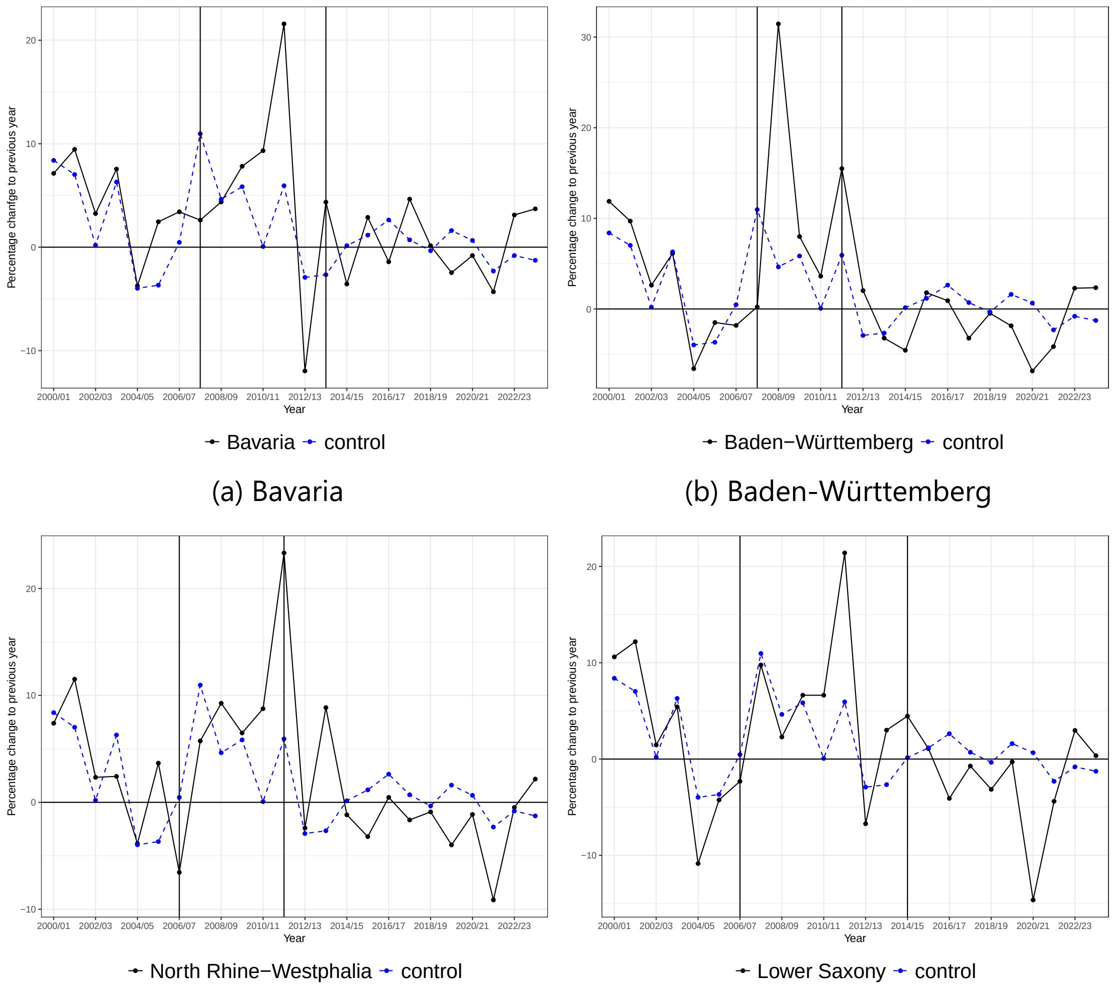
3.6.3 Selection concerns

Table B.7: Selection: Internal migration

	(1) For education	(2) General
Treated	0.147*	0.150**
	(0.050)	(0.051)
Int. migration inflow	-0.021	0.000
	(0.053)	(0.036)
λ_t, γ_e, δ_s	X	X
Demographics	X	X
Socioeconomic background	X	X
Num. obs.	795	795
R2	0.092	0.091
R2 Adj.	0.006	0.006

Note: Columns (1) to (2) contain results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of being overeducated. The sample is based on SOEP data on graduates from 1995 to 2021. Observations are restricted to the first observation per individual. Columns (1) to (2) use information from INKAR on internal migration and add dummies equaling one if the federal state experienced a net inflow of internal migrants in the respective year. This indicator is based on individuals aged 18 to 25 in column (1) capturing migration for educational purposes, while the indicator considers all migration streams in column (2). All covariates and fixed effects are included. Standard errors are clustered at the level of the state where the degree was obtained and are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure B.1: Percentage change in the average number of first-year students



(a) Bavaria

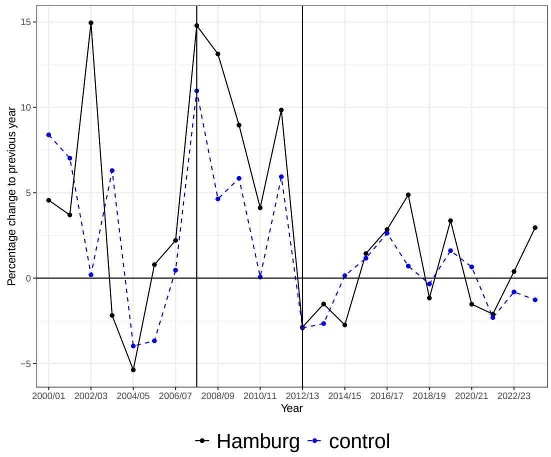
(b) Baden-Württemberg

(c) North Rhine-Westphalia

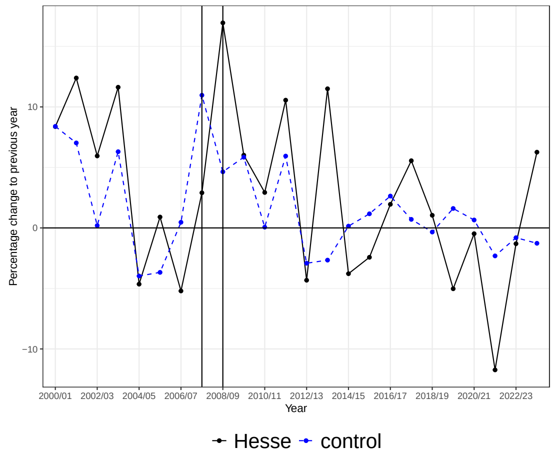
(d) Lower Saxony

Note: This figure is based on administrative data by the Federal Statistical Office (Statistisches Bundesamt, 2024b) on first-year students from winter term 1998/1999 to winter term 2023/24. It displays the percentage change of the average number of first-year students compared to the previous winter term for the treated (Bavaria, Baden-Württemberg, North Rhine-Westphalia, Lower Saxony) in black (solid line) compared to the aggregate of the control states in blue (dotted line). Values above 0 indicate an increase in the number of first-year students compared to the previous winter term, while negative numbers represent a decrease. Due to the estimation of the percentage change compared to the previous winter term, the first observation period (winter term 1998/99) is not included. The reference lines mark the first introduction of tuition fees and the last abolition of fees.

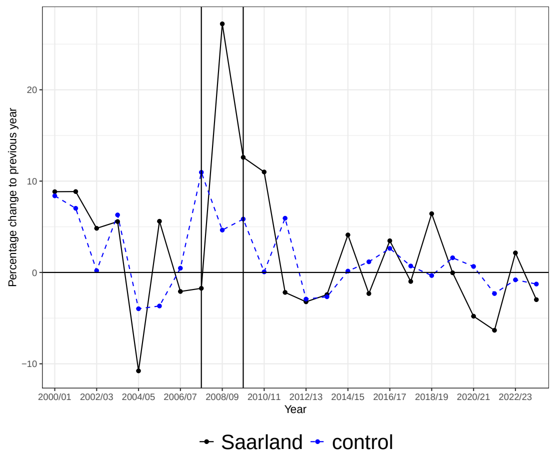
Figure B.1: Percentage change in the average number of first-year students (cont.)



(e) Hamburg



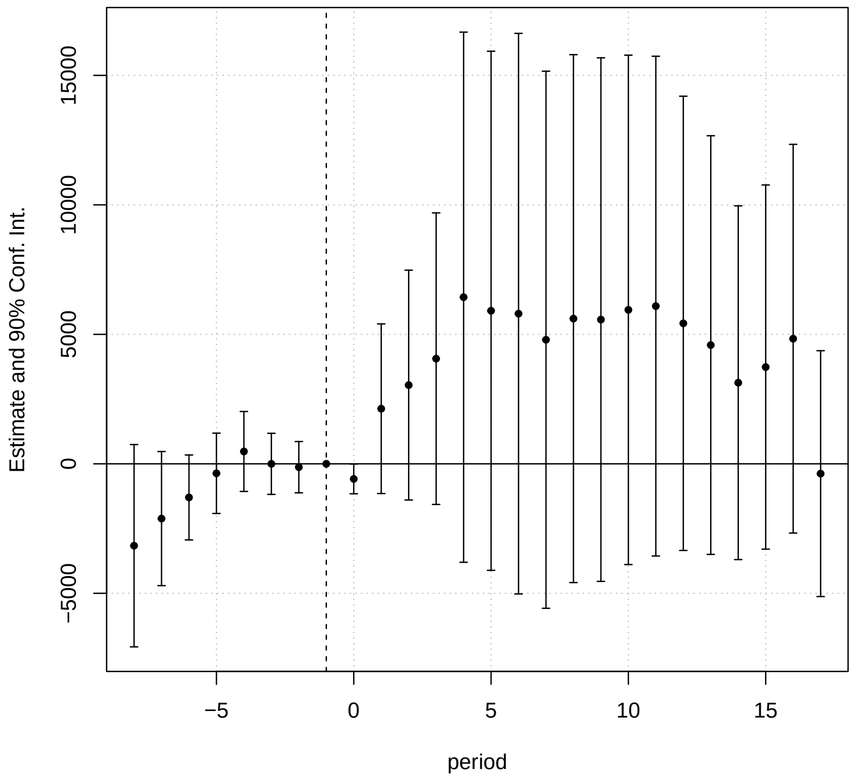
(f) Hesse



(g) Saarland

Note: This figure is based on administrative data by the Federal Statistical Office (Statistisches Bundesamt, 2024b) on first-year students from winter term 1998/1999 to winter term 2023/24. It displays the percentage change of the average number of first-year students compared to the previous winter term for the treated (Hamburg, Hesse, and Saarland) in black (solid line) compared to the aggregate of the control states in blue (dotted line). Values above 0 indicate an increase in the number of first-year students compared to the previous winter term, while negative numbers represent a decrease. Due to the estimation of the percentage change compared to the previous winter term, the first observation period (winter term 1998/99) is not included. The reference lines mark the first introduction of tuition fees and the last abolition of fees.

Figure B.2: Event study: Treatment effect on the number of students



Note: This figure is based on administrative data by the Federal Statistical Office (Statistisches Bundesamt, 2024b) on first-year students from winter term 1998/1999 to winter term 2023/24. It displays results from an event study approach based on Sun and Abraham (2021). To calculate the event study the period is defined based on the year in which the respective winter term started (e.g., winter term 2007/08 translates into cohort= 2007). The cohort equals 2007 for Baden-Württemberg, Saarland, Hesse, Hamburg and Bavaria, and 2006 for Lower Saxony and North-Rhine Westphalia. Federal state and year fixed effects included. Standard errors are clustered at the state level. Average Treatment Effect on the Treated (ATT) = 4452.67.

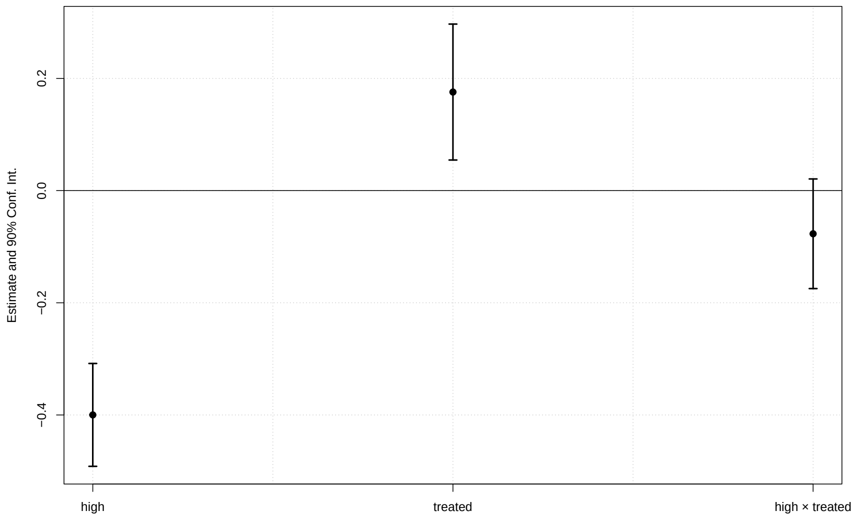
3.6.4 Heterogeneity

Table B.8: ITT: Occupational choice

	(1)
Treated	-0.018
	(0.068)
λ_t, γ_e, δ_s	X
Demographics	X
Socioeconomic background	X
Num.Obs.	966
R2	0.201
R2 Adj.	0.132

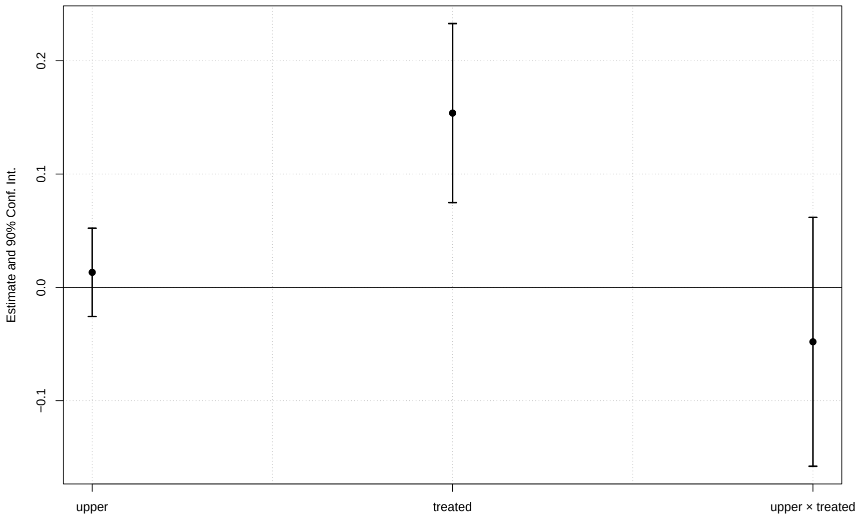
Note: Column (1) contains results from linear probability models estimating the ITT of exposure to tuition fees on the likelihood of working in a high-skill occupation. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. All covariates and fixed effects are included. Standard errors are clustered at the level of the state where the degree was obtained. Standard errors are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure B.3: Heterogeneity by occupation

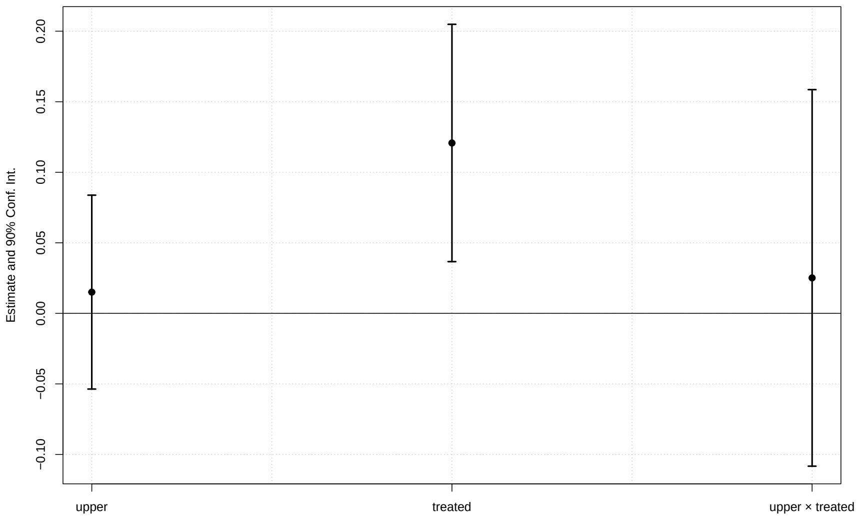


Note: The graph plots the coefficients based on Equation 3.3 focusing on the heterogeneity across individuals working in high- and low-skill occupations. The sample and covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

Figure B.4: Heterogeneity by parental education



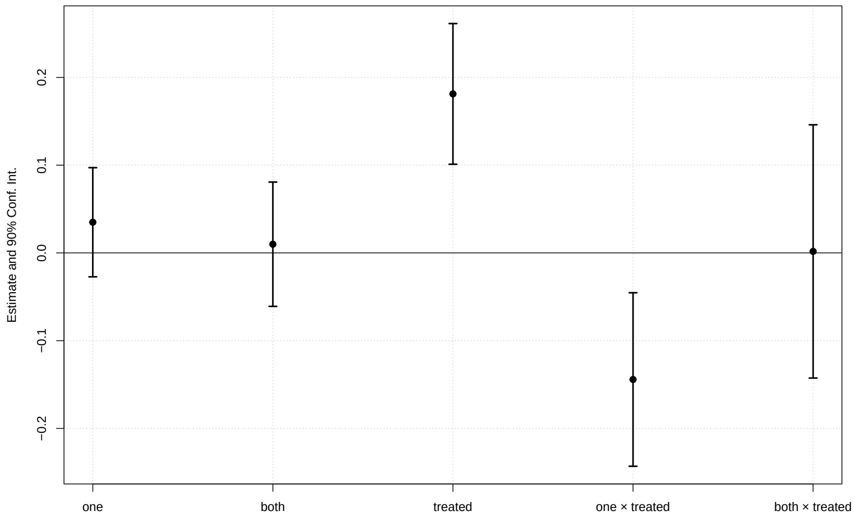
(a) By father education



(b) By mother education

Note: The graph plots the coefficients based on Equation 3.3 focusing on the heterogeneity across parental education. In Figure B.4a (Figure B.4b) "upper" indicates that the individual's father (mother) obtained upper secondary education. Due to multicollinearity issues, the regression does not additionally include the categorical indicators for father and mother education while the remaining covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

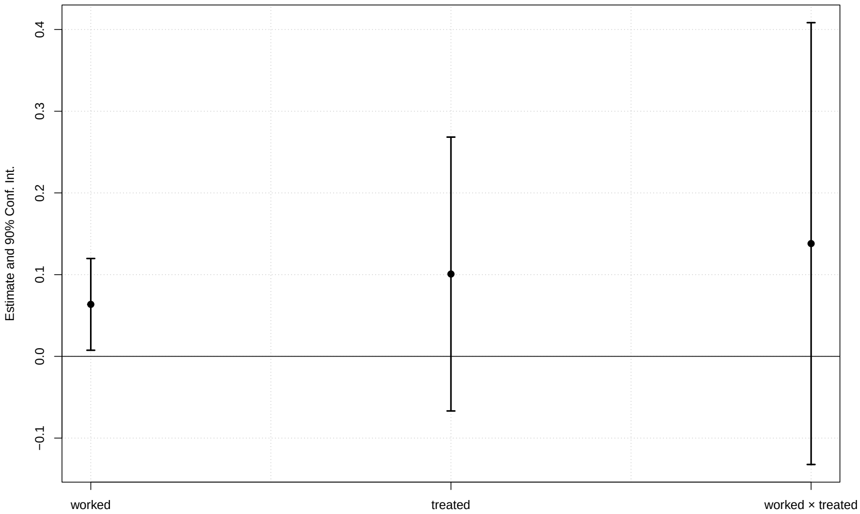
Figure B.4: Heterogeneity by parental education (cont.)



(c) By combination of mother and father education

Note: The graph plots the coefficients based on Equation 3.3 focusing on the heterogeneity across parental education. In Figure B.4c "one" ("both") indicates that one (both) of the parents obtained upper secondary education. Due to multicollinearity issues, the regression does not additionally include the categorical indicators for father and mother education while the remaining covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

Figure B.5: Heterogeneity by employment status during studies



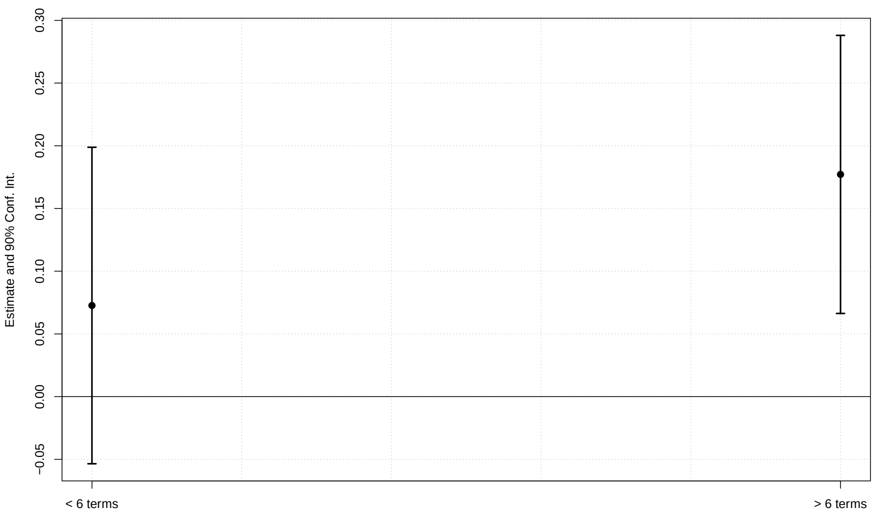
Note: The graph plots the coefficients based on Equation 3.3 focusing on the heterogeneity across employment status during studies. "Working" indicates that the individual was either marginally, part-time or full-time employed one year before university graduation. Due to missing information on the employment status of some individuals one year prior to graduation, the sample size reduces to 737 in this estimation. The covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

Table B.9: Heterogeneity by employment status: Additional analyses

	(1)	(2)	(3)
Dependent	Overeducation		Worked
Treated	0.133+	0.133+	-0.003
	(0.064)	(0.063)	(0.047)
Worked		0.075*	
		(0.028)	
λ_t, γ_e, δ_s	X	X	X
Demographics	X	X	X
Socioeconomic background	X	X	X
Num.Obs.	737	737	737
R2	0.112	0.116	0.212
R2 Adj.	0.011	0.014	0.122

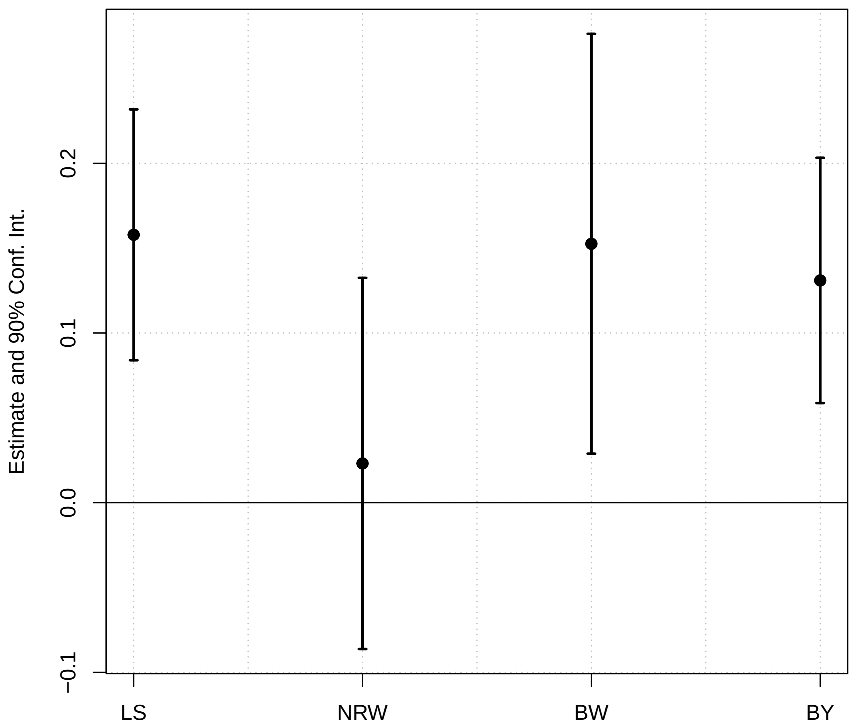
Note: Columns (1) to (3) contain results from linear probability models. Column (1) estimates the same regression as Table 3.3 on the reduced sample used in Figure B.5, while column (2) additionally controls for whether the individuals worked besides studying. Column 3 regresses the treatment indicator on the likelihood of having worked while studying. The sample is based on SOEP data on graduates from 1985 to 2021. Observations are restricted to the first observation per individual. All covariates and fixed effects are included. Standard errors are clustered at the level of the state where the degree was obtained and are reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure B.6: Heterogeneity by treatment duration



Note: The graph displays the ITT for individuals who were treated at least for a period of six terms (Bachelor's degree) compared to those who were treated shorter based on Equation 3.4. The sample and covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

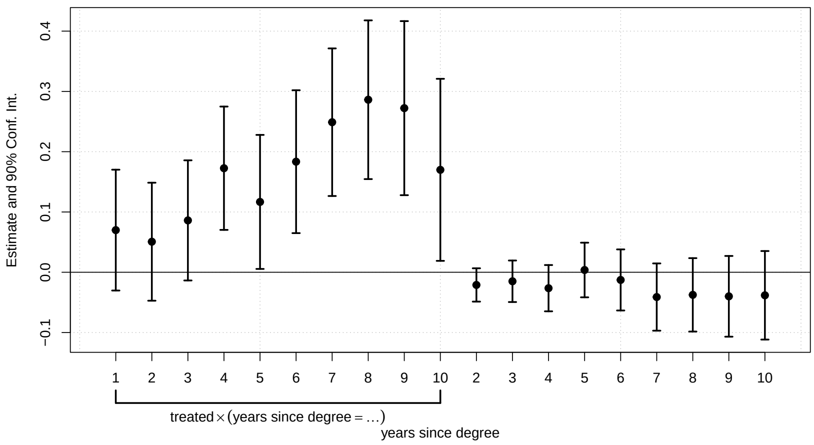
Figure B.7: ITT by federal state



Note: The graph displays the ITT for each federal state based on the interaction of the treatment indicator and the federal state in which the degree was obtained based on Equation 3.4. The sample and the remainder of the covariates are the same as in Table 3.3. 90%-confidence intervals plotted.

3.6.5 Dynamic analysis

Figure B.8: ITT of exposure to tuition fees: Dynamic analysis (extended)



Note: The graph displays the ITT by the years since degree completion based on the interaction of the treatment indicator and the years since degree completion. I.e., the first ten coefficients are the same coefficients are displayed in Figure 3.2. The remaining coefficients represent the baseline coefficients for the link between years since degree completion and the likelihood of overeducation (the first year after graduation forms the reference category). The regression is based on Equation 3.5. The sample is based on SOEP data using all repeated observations of graduates referred to in Table 3.3 surveyed between 1985 to 2022. It is restricted to individuals observed within ten years following degree completion. All covariates and fixed effects of Table 3.3 are included.

Chapter 4

Educational mismatch and trade union membership*

We hypothesise that trade unions assist their members in avoiding situations of educational mismatch. We test this hypothesis using data from the German Socio-Economic Panel and find that trade union membership is negatively associated with overeducation and positively with the likelihood of being educationally matched. These correlations are especially pronounced among core groups of members. Our findings suggest that strong trade union presence within these groups helps avert the adverse consequences of overeducation and educational mismatch. We observe no systematic linkage between union membership and undereducation.

Keywords: Educational mismatch; Trade union membership; German Socio-Economic-Panel

*This chapter is published in Industrial Relations (Geißler and Goerke, 2025) and is joint work with Laszlo Goerke. It is included under the terms of the Creative Commons Attribution License (CC BY 4.0). In order to include the paper in this thesis, the language has been adjusted at several points for example in terms of changes from American to British English. The project was awarded the INFER PhD Paper Award in 2023. An earlier version of this paper, with a different title, is available as INFER Working Paper 2023.14. Special thanks to an anonymous referee for very valuable comments and constructive suggestions. We are also grateful for the helpful comments from participants of the 25th Annual Conference of the International Network for Economic Research (INFER), the 2nd Gender and Economics Workshop of the Université du Luxembourg, and the Young Economist Meeting 2023.

4.1 Introduction

Educational mismatch occurs when job requirements and an employee's qualifications diverge. These situations of mismatch are a widespread phenomenon and can impose substantial costs on workers, firms, and society (e.g., Tsang and Levin, 1985). Accordingly, the determinants of overeducation and, to a lesser extent, undereducation, have been investigated frequently. Similarly, the groups that are particularly affected as well as characteristics associated with a divergence between the required and available qualifications have been studied intensively.[55] In contrast, the effects of labour market institutions, in general, and of trade unions, in particular, have scarcely been looked at.

This is surprising because trade unions can affect the attractiveness of hiring individuals with either excessive or insufficient educational qualifications. Union membership can also influence the employee's decisions to enter or remain in situations characterised by educational mismatch. Such an impact of trade unions may arise because members possess superior information regarding job features, required qualifications, compensation, career prospects, and the costs and benefits associated with changing jobs. In negotiations, trade unions are involved in establishing formal requirements for jobs with employers to define pay scales, they bargain about promotion rules and influence the conditions allowing firms to dismiss employees or prevent such dismissals. Therefore, trade unions can enhance the knowledge of members about the consequences of overeducation, opportunities for leaving such a state of educational mismatch and the likely outcomes of a resulting job change. Trade unions can also improve a member's bargaining power in negotiations with the employer when attempting to avoid overeducation. Such a bargaining power effect can arise because union membership strengthens protection against a job loss or dismissal. Moreover, if fellow members support a request for a better educational match, a firm's costs of declining it can increase. In sum, trade unions may help members to avoid overeducation and reduce its extent by providing information and strengthening the ability to leave a situation of educational mismatch. Comparable information and bargaining power effects may also alter a union member's likelihood and extent of undereducation.

[55] The pertinent analyses have looked at socio-demographic factors, such as gender (e.g., McGoldrick and Robst, 1996; Santiago-Vela and Mergener, 2022), migration background (e.g., Akgüç and Parasnis, 2023; Aleksynska and Tritah, 2013), and educational aspects, including grades, college quality or field of study (e.g., Robst, 1995b; Turmo-Garuz and Bartual-Figueras, 2019). Other studies have investigated job characteristics (Büchel and Pollmann-Schult, 2004; Di Pietro and Cutillo, 2006; Green and McIntosh, 2007) and the composition of the labour market (e.g., Davia et al., 2017; Tarvid, 2015).

Based on these considerations, we empirically analyse the relationship between union membership and various indicators of educational mismatch for Germany. The country represents an interesting research objective because labour relations are highly formalised. Many jobs require the completion of a specific apprenticeship scheme or a university degree in a particular subject or field. Consequently, employees are well aware of the possibility of educational mismatch. Moreover, educational mismatch and especially overeducation have increased during the last decades. Therefore, the scope for a union membership effect has gone up. Additionally, collective bargaining coverage does not imply union membership of covered employees, or vice versa. Accordingly, the German industrial relations system enables individual trade union membership to play a role. Finally, the German trade union federation (Deutscher Gewerkschaftsbund, DGB) has identified educational mismatch as a topic of concern and repeatedly highlighted its societal and individual impact. It has also commissioned a scientific study on the extent and determinants of overeducation (Deutscher Gewerkschaftsbund, 2014).

Using data from the German Socio-Economic Panel (SOEP) for a period of over 30 years, we show that trade union membership is negatively related to the duration of overeducation as well as its likelihood. Moreover, membership is positively linked to the likelihood of being educationally matched, meaning that individuals are neither over- nor undereducated. We observe no systematic correlation with undereducation. Looking at subgroups, the effects can be observed particularly for those individuals who exhibit a comparatively high likelihood of union membership. This suggests that trade unions specifically assist their core members when it comes to addressing educational mismatch, supporting the notion of a bargaining effect. Because informational advantages associated with membership are unlikely to vary across subgroups in accordance with the share of employees belonging to a trade union, our results bring forth no evidence for an informational channel.

The paper provides background information on the institutional set-up in Germany, focusing particularly on industrial relations and educational mismatch in Section 4.2. In Section 4.3, we derive hypotheses about the correlation between an employee's trade union membership and various indicators of educational mismatch. We review related analyses in Section 4.4, while we describe the data and explain the empirical approach in Section 4.5. We present our central findings, several robustness checks, and a heterogeneity analysis in Section 4.6 and finish with some concluding remarks in Section 4.7.

4.2 Institutional set-up

4.2.1 Industrial relations in Germany

As of now, approximately 16% of employees in Germany are members of a trade union. This fraction has declined substantially, from a peak of about 36% just after the re-unification in 1990 (OECD and AIAS, 2023). Union members in Germany generally pay a (tax-deductible) membership fee of 1% of their gross wage. In exchange, they receive financial support in case of a strike, legal advice, and support in employment-related conflicts. In addition, trade unions provide members with information about working conditions and job-related aspects.

In contrast to other countries (see e.g., Bryson, 2014; Fang and Hartley, 2022), there is no robust evidence of wage differences between union members and non-members in Germany (Fitzenberger et al., 1999; Goerke and Pannenberg, 2004; Schmidt and Zimmermann, 1991).[56] A major reason is that union membership is not directly linked to collective bargaining. This can already be observed when comparing the bargaining coverage of close to 40% in the private sector in 2023, respectively almost 90% in the public sector (Hohendanner and Kohaut, 2024), to union density. From a legal perspective, collective bargaining agreements apply to all firms covered by the contract and to union members who work in a covered firm. However, in the vast majority of cases, firms pay all employees the negotiated wage, irrespective of their union membership status. Collective bargaining agreements usually contain provisions concerning remuneration, and often with regard to working time, vacation entitlements, and overtime pay. Moreover, they define pay scales and qualification and tenure requirements that have to be fulfilled to move from one to the next level. Collective bargaining agreements may also include job security and training provisions. While pay is often negotiated annually, bargaining about other components usually takes place less frequently. Firms not subject to collective agreements can negotiate wages and working conditions individually with their staff. However, in about 50% of the relevant contracts, at least pay is aligned to a collective agreement (Hohendanner and Kohaut, 2024).

In addition to collective bargaining, co-determination is still an important element of the German industrial relations system (Jäger et al., 2022).

[56] Bonaccolto-Töpfer and Schnabel (2023) present evidence suggesting a union membership wage premium of about 2.5%, with substantial variations across occupations, using two recent waves of the SOEP.

Co-determination at the plant level in the private sector takes place via works councils. They can be established by a vote among employees in firms with at least five employees. Subsequent to being set up, such votes are repeated at regular intervals. Since these votes are not compulsory, works councils currently exist in less than 10% of all eligible private-sector establishments. Given their prevalence in larger firms, about 40% of private-sector employees are subject to co-determination. This percentage has declined substantially from about 55% in the mid-1990s. In the public sector, an analogue to works councils exists. Such personnel councils are much more widespread (Hohendanner and Kohaut, 2024).

4.2.2 Trade unions and educational mismatch in Germany

The incidence of educational mismatch has risen substantially over the last decades. According to the OECD, the share of individuals possessing more qualifications than required for their job was on average 15.3% in the European Union in 2022. Moreover, 16.8% were underqualified, implying that around every third individual's work exhibited educational mismatch. In Germany, the respective shares were slightly larger with 18.8% being over- and 21% being underqualified (OECD, 2022b). While the OECD provides no data on the development of educational mismatch for the entire German labour force over time, there is such information for graduates aged 25 - 34 between 2011 and 2022. These statistics show an increase of overqualification from 11.7% in 2011 to 18.8% in 2022 (European Centre for the Development of Vocational Training, nd). Considering all employed individuals, data from the SOEP for the period 1984 to 2020 reveals a similar pattern (see Clogg and Shockey (1984); Verdugo and Verdugo (1989), Figure C.1).

Due to the rising prevalence of educational mismatch and its potential negative consequences, trade unions in Germany try to establish conditions that counteract the development of educational mismatch and other forms of underemployment. The DGB, for example, requests a better compatibility of family and work in order to reduce gender frictions in the risk of underemployment (Deutscher Gewerkschaftsbund, 2014). Moreover, trade unions advocate employment agencies to rely on qualification-oriented counselling to reduce the risk of educational mismatch among unemployed individuals who re-enter the labour market (Deutscher Gewerkschaftsbund, 2014). Another example of the German trade unions' stance on underemployment and educational mismatch is the statement on the draft law regulating the employment of civil servants who worked for the post office prior to its privatisation. This draft permitted a broader deployment of civil servants, also to jobs for which they were overqualified. The trade union representing these employees strongly objected to the possibility of assigning such jobs (Vereinte Dienstleistungsgewerkschaft, 2014). Trade unions also engage in strengthening the role of on-the-job training, as this is not only relevant for individuals lacking any kind of apprenticeship or qualification, but particularly for individuals who are overeducated (IG Metall, 2018). This is also visible in the DGB's engagement in enhancing and extending the Continuing Education Act (Deutscher Gewerkschaftsbund, 2023).

In sum, trade unions in Germany have identified the increasing extent of educational mismatch as a development having potentially adverse consequences for its members. In consequence, they promote policies and regulations that help to avoid its expansion.

4.3 Union membership and educational mismatch

Overeducation has been associated with a number of detrimental outcomes, such as lower job satisfaction (e.g., Battu et al., 1999; Iseke, 2014; Verhaest and Verhofstadt, 2016), wage losses (e.g., Cohn and Khan, 1995; Duncan and Hoffman, 1981; Heijke et al., 2003; Korpi and Tåhlin, 2009; Sattinger and Hartog, 2013; Sicherman, 1991), and possibly a deterioration in health (e.g., Bracke et al., 2013; Korpi and Tåhlin, 2009). In consequence, we assume that employees desire to leave a job for which they are overeducated.

Overeducation is arguably more likely to occur or persist when employees are unaware of how to secure a more suitable job, lack information about alternative employment opportunities, or do not fully comprehend the negative consequences associated with overeducation. Trade unions in Germany provide members with job-related information, such as required levels of qualifications or associated wages and benefits. To illustrate the informational role of membership, first, a look at the websites of German trade unions is instructive. The major German unions emphasise the informational benefits when listing the main arguments for becoming a member.[57] Second, the research institute of the DGB runs a continuous online survey to allow for pay comparisons at a very detailed level, *inter alia*, based on an individual's formal qualifications and experience.[58]

[57] See, for example, the websites of the Railroad and Transport Union, EVG (https://www.evg-online.org/deine-vorteile/unsere-leistungen/) and the Mining, Chemical and Energy Industrial Union, IGBCE (https://igbce.de/igbce/werde-mitglied).

[58] See the website https://www.lohnspiegel.de.

As a third example, the IG Metall, the largest German trade union with more than 2 million members, provides comprehensive information on collective bargaining agreements it has negotiated to its members.[59] In addition to these direct informational benefits, union members take a more active stance in social exchanges, and union membership can provide social contacts (Flavin et al., 2010; Flavin and Shufeldt, 2016; Keane et al., 2012; Radcliff, 2005). Greater social capital can also result in better information about working conditions and educational requirements.[60] Therefore, we hypothesise that trade union membership reduces overeducation on account of its informational advantages.

Even if employees know about their overeducation and its adverse consequences, obtaining a job with a better educational match may be difficult, especially if firms benefit from employing overeducated individuals (see, e.g., Sloane et al., 1996). First, employees have to voice their desire for a job change. Second, they must realise their aspirations. If trade unions act as the voice of employees, union members, in particular, are better able to communicate their desire to avoid overeducation than non-members. Moreover, if union members are better protected against management reprisals in case of voicing discontent, for example, because they are less likely to face a dismissal (see, e.g., Berglund and Furåker, 2016; Freeman, 1980; Goerke and Pannenberg, 2011; Ivlevs and Veliziotis, 2017; Pierse and McHale, 2015), they will more often express an ambition to obtain a better job match.[61] As mentioned above, trade unions represent their members in job-related legal disputes. In 2023, about 65,000 cases were handled by lawyers of the member unions of the DGB in labour courts. About one-third of them were related to pay issues (DGB Rechtsschutz, 2024). Since union representation in labour courts is free for members, whereas non-members have to pay their lawyers and possible court-fees themselves, union members face lower expected costs of overcoming situations of overeducation and inadequate pay. These lower costs are one reason why union members possess greater individual bargaining power. A further source for a better individual negotiation position is the support for union members either at the plant-level or in local

[59]See the website https://www.igmetall.de/mitglieder.

[60]This line of reasoning is consistent with evidence that the risk of overeducation (overqualification) is lower for individuals with highly educated fathers or individuals from high status families (Capsada-Munsech, 2015; Erdsiek, 2016; Turmo-Garuz and Bartual-Figueras, 2019; Verhaest and Omey, 2010) because they might provide their children with better information regarding their labour market integration as well as with higher social capital in the form of work-related networks.

[61]Sloane et al. (1999) argue that trade unions may entice employers to change job specifications and recruiting behaviour to reduce educational mismatch. Such an effect would have the same consequences as lower costs of requesting a better match.

offices, which especially the large German trade unions emphasise on their websites.

In sum, overcoming a situation of overeducation requires employees to actively initiate the necessary changes. Union members may be more willing to undertake such steps because they know more about their potentially beneficial consequences and possess a greater bargaining power when determining the conditions of a new job.

Our considerations can be summarised as:
Hypothesis 1.1: Trade union members exhibit lower levels of overeducation than comparable non-members.
Hypothesis 1.2: Trade union members are less likely to be overeducated than comparable non-members.

The correlates of undereducation are less well-established than those of overeducation. There is some evidence that undereducated employees obtain higher wages than comparably educated individuals who exhibit no educational mismatch (Kiker et al., 1997; Verdugo and Verdugo, 1989; Verhaest and Omey, 2012). This suggests that undereducated individuals have to invest less in their formal education to receive the same expected returns as adequately educated employees. Therefore, employees may benefit from undereducation (Sloane et al., 1996). If that is the case, the same mechanisms which help trade union members to avoid or leave situations of overeducation – better information and greater bargaining power – could assist them in acquiring a position for which they are inadequately educated.

Based on the above, we postulate:
Hypothesis 2.1: Trade union members exhibit higher levels of undereducation than comparable non-members.
Hypothesis 2.2: Trade union members are more likely to be undereducated than comparable non-members.

We expect the correlation between union membership and undereducation to be weaker than the relationship to overeducation. This is the case for a number of reasons: First, trade unions are strong advocates of the German apprenticeship system. An important element of this support has been demands that all successful graduates of such vocational training should be offered a matching job in the company where the training took place (cf., IG Metall, 2012). Therefore, the scope for undereducation is limited at this stage. Second, by actively promoting undereducation, trade unions would undermine the system of pay scales which makes wage determination less arbitrary. Consequently, support is likely to be more tacit than in the case of overeducation. Finally, trade unions advocate fair and adequate wages. The argument that an individual's remuneration does not correspond to these requirements is obviously more convincing if an employee is underpaid, but not if overpaid, as it may occur in the case of undereducation (Kiker et al., 1997; Verdugo and Verdugo, 1989; Verhaest and Omey, 2012).

Given that educational mismatch is defined as the presence of over- or undereducation, and that we anticipate the union membership impact on overeducation to dominate, we can state:
Hypothesis 3: Trade union members are more likely to be educationally matched than comparable non-members.

4.4 Previous contributions

To the best of our knowledge, there is no thorough investigation of the impact of an individual's trade union membership on educational mismatch. Some analyses, however, include a union membership indicator as one of many right-hand side variables, without discussing the observed correlation in detail.[62]

McGoldrick and Robst (1996) use the 1985 wave of the PSID to estimate determinants of overeducation for the United States. The sign of the estimated coefficient of the union indicator varies with the measure of educational mismatch. Sloane et al. (1996, 1999) present evidence for the British labour market in the mid-1980s. They observe no systematic correlation between trade union membership and overeducation. Similarly, in the study by Wald and Fang (2008), based on data from the 1999 wave of the Canadian Workplace and Employee Survey (WES), overeducation is not significantly correlated with a union indicator.[63]

Fleming and Kler (2008) focus on Australia and utilise the first wave of HILDA. They report a positive correlation between union membership and overeducation. Mavromaras et al. (2009) use the same dataset and indicate that severe overskilling is more frequent among trade union members. Employing the 2004 sweep of the WERS data in the United Kingdom, Belfield (2010) shows

[62]Other studies consider correlations between educational mismatch and union density (Aleksynska and Tritah, 2013; Davia et al., 2017) or collective bargaining coverage (Jacobs et al., 2021).

[63]Wald and Fang (2008) do not distinguish between individual trade union membership and bargaining coverage.

that the probability of overeducation tends to be higher among union members than non-members in the private, but not in the public sector. Aleksynska and Tritah (2013) focus on immigrants in 22 European countries using ESS data between 2002 and 2009. They provide evidence for a positive relationship between trade union membership and overeducation and a negative linkage with undereducation for native-born individuals, but not for immigrants. Finally, and somewhat in contrast to the insights for Australia, Canada, the United States, and various European countries, Sharma and Sharma (2017) report a negative relationship between union membership and both overeducation and undereducation in India.

In sum, studies for various countries include information about union membership when analysing educational mismatch. None of these investigations analyses this correlation in detail, and they also tend to focus on overeducation.

4.5 Data and methodology

To estimate the correlation between trade union membership and educational mismatch, we use data from the SOEP (version 37, DIW, 2022; Goebel et al., 2019), a representative longitudinal study on households and individuals living in Germany. The survey has been conducted annually since 1984 (Goebel et al., 2019). Until 1989, the year prior to German reunification, the data include only inhabitants from the western part of the country.

4.5.1 Educational mismatch

We employ the so-called statistical approach to measure overeducation and undereducation. It is based on the idea that the average level of formal education in a particular occupation indicates its required level. Employees whose years of education are at least one standard deviation above this average are classified as overeducated, while those falling short by at least one standard deviation are classified as undereducated (Clogg and Shockey, 1984; Verdugo and Verdugo, 1989).

The statistical approach is advantageous compared to other indicators of educational mismatch, such as the indirect self-assessment or the job analyst measure, as it allows the evaluation of an individual's position in the respective reference group (Capsada-Munsech, 2019). It assesses educational mismatch based on labour market demand and supply mechanisms (Hartog, 2000), in line with Freeman's (1976) conceptualisation of overeducation. Furthermore, the statistical measure automatically adjusts to changes in the labour market and job requirements (Capsada-Munsech, 2019; Clogg and Shockey, 1984). Finally, the statistical approach implies a restrictive classification of over- and undereducation, as it usually results in a larger proportion of matched individuals than the alternative measures (Blásquez and Budría, 2012).

Due to rising educational requirements for many jobs and increased participation in higher education, the average level of formal education within specific occupations has likely increased over time as well.[64] To account for this effect, we follow Clogg and Shockey (1984) and Verdugo and Verdugo (1989) and use the mean years of education, $\overline{x}_{ot}$,

$$\overline{x}_{ot} = \frac{1}{I_{ot}} \sum_{i}^{I_{ot}} x_{it} \tag{4.1}$$

as the benchmark to determine the existence of over- and undereducation, where I_{ot} denotes the number of individuals working in occupation o in survey year t and x_{it} the years of education of individual i.

To define an occupation o, we utilise the ISCO classification based on the question *"What is your current position/occupation?"*.[65] Following Blásquez and Budría (2012), we transform the 4-digit variable into a 3-digit classification to define the reference group. The years of education, x_{it}, are extracted from the variable *"Number of years of education"*.

We define an individual as overeducated if $x_{it} > \overline{x}_{ot} + s_{ot}$ and as undereducated if $x_{it} < \overline{x}_{ot} - s_{ot}$, where s_{ot} denotes the standard deviation of the duration of education (in years) of individuals in occupation o in survey year t. An individual who is neither over- nor undereducated belongs to the group of educationally matched employees. Therefore, employees for whom $x_{it} < \overline{x}_{ot} + s_{ot}$ holds are either undereducated or educationally matched and, thus, not classified as overeducated, while those characterised by $x_{it} > \overline{x}_{ot} - s_{ot}$ are not undereducated and, consequently, either overeducated or matched. Finally, we employ two quantitative measures of educational mismatch. Years of overeducation are given

[64] The educational expansion is substantial in Germany. Official labour market data shows an increase in higher education graduates from 227,525 in 1998 to 505,650 (Statistisches Bundesamt, 2022b, 2023b).

[65] The SOEP provides ISCO classifications in two versions 88 and 08, which need to be combined. While some occupational groups remained unaffected by the adaptation, others were recoded in the adapted version (see International Labour Office, 2012). To ensure consistency and currency, we use the adapted ISCO-08 classification in these cases.

by $x_{it} - \overline{x}_{ot} - s_{ot}$ if this difference is positive (and zero otherwise). In turn, years of undereducation are given by $|x_{it} - \overline{x}_{ot} + s_{ot}|$ if $x_{it} - \overline{x}_{ot} + s_{ot} < 0$ holds (and are zero otherwise).

4.5.2 Independent variables

Trade union membership

Our main independent variable stems from the question *"Are you a member of one of the following organisations or confederations?"*, allowing respondents to select either a yes-box or a no-option. The first option is *trade union*.[66] From the responses, we construct a dummy variable, which takes the value of one if an employee belongs to a trade union, and zero otherwise. The questionnaire contains the relevant information in the survey years 1985 and 1989 for the western part of the country and, after the reunification, in 1993, 1998, 2001, 2003, 2007, 2011, 2015 and 2019 for respondents throughout Germany.[67]

Since we observe a respondent's union membership broadly every three or four years, we can identify a change in the membership status if an individual participated in the survey for a sufficient number of waves. However, the data do not allow us to determine the exact timing of the change or to distinguish between single and multiple changes between the observation periods. Moreover, the likelihood of observing a variation in membership is greater for respondents who participate in the survey for longer intervals than for individuals who only answer the questionnaire in a few waves. Finally, the union membership question does not provide information on which trade union an individual belongs to and whether the respondent changed unions.

Control variables

We include several independent variables commonly employed in analyses of educational mismatch. The first set focuses on personal characteristics. In particular, we include a dummy variable set equal to one if the respondent is

[66]The question stems from the questionnaire in 2019. It constitutes the translation of the German text by the authors since the one provided by the SOEP is imprecise. The wording of the question and the list of organisations or confederations vary somewhat over time. The SOEP always provides the option *trade union*.

[67]Information on union membership is additionally available in 1990 for respondents from east Germany. We do not include data from this wave in our basic specification because the wave does not contain data on all the confounding factors listed in the next paragraph.

female (e.g., McGoldrick and Robst, 1996), a categorical measure of migration status (e.g., Aleksynska and Tritah, 2013),[68] variables describing the respondent's age (and age squared divided by 100) (e.g., Di Pietro and Cutillo, 2006; Ordine and Rose, 2009) and dummy variables indicating civil status (e.g., Groot, 1996).[69] As educational mismatch may be related to childcare responsibilities (e.g. Green and McIntosh, 2007; Groot, 1996), we control for the number of children. Moreover, the estimations contain dummy variables indicating the federal state of residence.

The second set of independent variables relates to job-related aspects. Specifically, we include dummy variables for being a white-collar worker (e.g., Duncan and Hoffman, 1981) and a civil servant (e.g., Boll et al., 2016), for having experienced unemployment, having a permanent contract (e.g., Turmo-Garuz and Bartual-Figueras, 2019), being a full-time employee (e.g., Sloane et al., 1999), as well as firm-size dummies (e.g., Battu et al., 1999), and years of tenure (and tenure squared divided by 100) (e.g., Kiker et al., 1997). Moreover, we incorporate dummies differentiating ten industries as well as nine occupational groups.

4.5.3 Estimation sample

We consider employed individuals and exclude self-employed and solo self-employed individuals, as well as all respondents exceeding their legal retirement age from our sample. Moreover, we drop all observations with missing values. This leaves us with a sample of 65,217 observations from 33,235 individuals for our main estimations.

About 19% of all observations stem from members of a trade union (see Table 4.1). In our sample, we observe 15,816 individuals at least in two waves with union membership information. 2,886 individuals changed their union membership status at least once. While, therefore, a sizeable fraction of individuals either join or leave a trade union, we only observe 478 cases of variations in union membership and the educational mismatch status from one wave to the next containing the relevant information.

[68]We differentiate between direct migrants who immigrated to Germany themselves, indirect migrants who were born in Germany, while their parents came from abroad, and a third (reference) group of individuals without any migration background.

[69]We differentiate individuals living together in any kind of relationship, such as marriage or registered partnership, individuals living separate, for example, due to divorce, while the reference group is formed of those who indicate that they have never been in a registered relationship, marriage or comparable, which we will refer to as singles.

Table 4.1: Descriptive statistics

	Mean	Sd	Min	Max
Trade union	0.1906	0.3927	0	1
Female	0.4902	0.4999	0	1
Direct migrant	0.1496	0.3567	0	1
Indirect migrant	0.0512	0.2204	0	1
No migrant	0.7992	0.4006	0	1
Age	41.7915	11.4990	17	65
Single	0.2643	0.4410	0	1
Married or partnered	0.6187	0.4857	0	1
Separate or widowed	0.1170	0.3214	0	1
Number of children	1.4891	1.1905	0	12
White-collar	0.5952	0.4909	0	1
Civil servant	0.0566	0.2311	0	1
Prior unemployment	0.3582	0.4795	0	1
Permanent	0.8736	0.3323	0	1
Full-time	0.6985	0.4589	0	1
Public sector	0.2577	0.4374	0	1
Tenure	10.3390	9.9632	0	50.92
Firm size: < 20	0.2251	0.4176	0	1
Firm size: 20 - 199	0.2890	0.4533	0	1
Firm size: 200 - 1999	0.2288	0.4200	0	1
Firm size: ≥ 2000	0.2571	0.4370	0	1
Num. obs.	65217			

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019.

Table 4.1 also reports descriptive statistics for some further variables (see Table C.1 in Appendix 4.8.2 for a complete documentation). About 49% of the observations stem from females. Respondents are on average 41.8 years old. The majority lives in some kind of registered partnership, has no migration background, works full-time, in a white-collar job, and has a permanent contract. Around one quarter of the sample works in the public sector, while 5.6% are civil servants. Approximately one third has experienced unemployment. On average, individuals have worked for the same company for around 10 years.

4.5.4 Estimation approach

As years of over- and undereducation are censored at zero from below, we estimate random effects tobit models (Tobin, 1958).[70]

Our estimation equation reads:

$$Y_{ito} = \beta_1 + \beta_2 TUM_{it} + \beta_3 X'_{it} + \lambda_t + \eta_m + \gamma_o + \xi_s + \nu_i + \epsilon_{it} \tag{4.2}$$

[70]All estimations are pursued with STATA.

Y_{ito} measures years of under- and overeducation of individual i at time t in occupation o. TUM_{it} equals one if a respondent belongs to a trade union and β_2 is the coefficient of interest. X'_{it} is a vector containing personal and job-related controls. We include year and month dummies λ_t and η_m. γ_o captures occupation effects, while ξ_s represents the industry fixed effects. ν_i includes the individual random effects that are assumed to be independent and identically distributed (i.i.d.), $N(0, \sigma^2_\nu)$ and ϵ_{it} is the error term, which is assumed to be i.i.d. $N(0, \sigma^2_\epsilon)$ and independent of ν_i.

When we analyse the existence of educational mismatch, we estimate random effects logit models of the form:

$$Pr(Y_{ito} \neq 0) = P(\beta_1 + \beta_2 TUM_{it} + \beta_3 X'_{it} + \lambda_t + \eta_m + \gamma_o + \xi_s + \nu_i + \epsilon_{it}), \quad (4.3)$$

where Y_{ito} refers to being undereducated, overeducated or educationally matched.

4.6 Results

In Section 4.6.1, we document educational mismatch for trade union members and non-members. This provides first evidence for our hypotheses. Section 4.6.2 presents the main estimation results and Section 4.6.3 reports the findings from a number of robustness checks. In Section 4.6.4, we consider the correlation for various subgroups to derive insights concerning the mechanisms discussed – information and bargaining power – underlying the observed effects.

4.6.1 Descriptive evidence

Table 4.2 indicates that members of a trade union exhibit, on average, about 1.5 months of overeducation (= 0.1247 x 12), while the respective number for non-members equals almost 2 months. Moreover, the probability of being overeducated (educationally matched) is about 2.9 (3.9) percentage points lower (higher) for union members than non-members. Finally, we observe a one percentage point lower probability of being undereducated among trade union members. Therefore, the descriptive evidence is consistent with Hypotheses 1.1, 1.2, and 3 but reveals no evidence of a positive linkage between union membership and undereducation, as suggested by Hypotheses 2.1 and 2.2.

Table 4.2: Mismatch outcomes by union membership status

	All		TUM = 0		TUM = 1		T-test
	Mean	Sd	Mean	Sd	Mean	Sd	
Years of undereducation	0.1056	0.3900	0.1065	0.3939	0.1020	0.3733	-0.0045
Years of overeducation	0.1583	0.5962	0.1662	0.6115	0.1247	0.5251	-0.0415***
Undereducated	0.1271	0.3331	0.1291	0.3353	0.1187	0.3234	-0.0104**
Overeducated	0.1314	0.3379	0.1369	0.3437	0.1081	0.3106	-0.0288***
Matched	0.7415	0.4378	0.7340	0.4419	0.7732	0.4188	0.0392***
Num. obs.	65217		52789		12428		

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.2 Main findings

Table 4.3 depicts the estimated coefficients for the union membership dummy for panel random effects tobit specifications (see Equation (4.2)). The dependent variables are years of undereducation in Panel A and years of overeducation in Panel B.[71]

The specifications controlling for month and year fixed effects in column (1) reveal a negative correlation between trade union membership and years of under- and overeducation. The correlation with undereducation is no longer statistically significant once we control for demographic characteristics. In contrast, the negative correlation between trade union membership and years of overeducation remains qualitatively unaffected by the successive inclusion of additional control variables. Column (5) indicates that union membership is associated with a reduction in the extent of overeducation by 0.15 years (almost 2 months).

Table 4.4 depicts average marginal effects for panel random effects logit specifications (see Equation (4.3)) relating union membership to the likelihood of undereducation (columns (1) and (2)), overeducation (columns (3) and (4)), and for being educationally matched (columns (5) and (6)). Focusing on the specifications including the full set of control variables in the even-numbered columns, we observe that trade union membership is associated with a 1.3 percentage point lower likelihood of overeducation and an almost 2 percentage points higher likelihood of being matched, while no correlation with undereducation is discernible.[72]

[71] For alternative measures of educational mismatch, such as indirect self-assessment (ISA), job analyst measure (JA), and an extended statistical measure (Duncan and Hoffman, 1981; Jacobs et al., 2021; Rumberger, 1981a,b), we obtain qualitatively similar results as reported in Table 4.3.

[72] The results in Table 4.3 and Table 4.4 remain unaffected by incorporating labour market experience instead of tenure or dummy variables for education levels, and expanding the sample through the 1990-wave, while excluding details about contract duration.

Table 4.3: Panel RE tobit

	(1)	(2)	(3)	(4)	(5)
Panel A: Years of undereducation					
Trade union	-0.0880***	-0.0923***	-0.0807***	-0.0276	-0.0267
	(0.0281)	(0.0281)	(0.0243)	(0.0249)	(0.0252)
Panel B: Years of overeducation					
Trade union	-0.228***	-0.234***	-0.155***	-0.159***	-0.150***
	(0.0394)	(0.0395)	(0.0319)	(0.0323)	(0.0340)
Month/ year FE	X	X	X	X	X
Federal state FE		X	X	X	X
Occupation/ industry FE			X	X	X
Demographics				X	X
Job-related					X
Num. obs.	65217	65217	65217	65217	65217

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. Demographics include dummies for being female, having a direct or indirect migration background, age (level and squared term), dummies for being married or living separately, and the number of kids. Job-related controls incorporate dummies for being a white-collar worker, a civil servant, having experienced unemployment before, owning a permanent contract, working full-time, being a public sector employee, tenure (level and squared term), and three dummies distinguishing firm size. Standard errors presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.4: Panel RE logit: Average marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Undereducated		Overeducated		Matched	
Trade union	-0.00969***	-0.00422	-0.0177***	-0.0129***	0.0306***	0.0188***
	(0.00298)	(0.00332)	(0.00316)	(0.00318)	(0.00474)	(0.00501)
Month & year FE	X	X	X	X	X	X
Further covariates		X		X		X
Num. obs.	65217	65217	65217	65217	65217	65217

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. The further covariates are the same as those listed in Table 4.3. Standard errors, clustered at individual level, in parentheses. The average marginal effects account for the discrete change from the base level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In summary, there is less overeducation among trade union members than among non-members, while the probability of being educationally matched is higher among members. These findings are consistent with Hypotheses 1.1, 1.2 and 3, whereas we observe no correlation between union membership and undereducation, as claimed in Hypotheses 2.1 and 2.2.

Turning to other covariates, we observe less overeducation and a higher likelihood of being educationally matched among females and public sector workers, whereas white-collar and permanent employees exhibit less undereducation and more overeducation, with the former correlation dominating, resulting in a higher likelihood of being educationally matched (for related evidence, see, *inter alia* Aleksynska and Tritah, 2013; Belfield, 2010; Di Pietro and Cutillo, 2006; Green and McIntosh, 2007; Jacobs et al., 2021; Kiker et al., 1997; Ordine and Rose, 2009; Sloane et al., 1999). For immigrants, we observe the reverse.[73] The relationship between age and the indicators of mismatch is generally nonlinear (for related evidence, see, *inter alia* Green and McIntosh, 2007; Joona et al., 2014). Table C.2 and Table C.3 in Appendix 4.8.2 display the results for all right-hand-side variables.

4.6.3 Some robustness checks

In this section, we report the findings from three robustness checks. The first two cater to concerns that the union membership dummy is correlated with other determinants of educational mismatch that have not adequately been accounted for in Equations 4.2 and 4.3. The third robustness check focuses on the comparability of union members and non-union respondents. The findings are documented in Appendix 4.8.3.

Altering a situation of educational mismatch often involves a job change. If taking up a new job is furthermore correlated with a variation in the union status, the estimated coefficient for the union dummy may partially capture the consequences of a job change. In the specifications presented in Tables 4.3 and 4.4 we indirectly account for this possibility as we include tenure as a covariate. We can also tackle this issue directly and use the tenure information to create a dummy variable indicating a job change in the previous year. We then estimate Equations 4.2 and 4.3 including the job change indicator and its interaction with the union membership dummy. Results are displayed in Table C.4 and Figure C.2 in Appendix 4.8.3. The results for the union membership dummy are qualitatively the same as those reported in Tables 4.3 and 4.4. Moreover, the interaction terms are never significantly different from zero. Therefore, we find no evidence that a recent job change affects the relationship between union membership and educational mismatch.

About 40% of employees in Germany work in establishments with co-determination via works councils at the plant level. Since guidelines for the selection of employees for recruitment, transfer, regrading, and dismissal require the approval of the works council, co-determination may affect the extent of educational mismatch. Moreover, the likelihood of union membership is correlated with co-determination (Behrens, 2009). Therefore, the union dummy may partially measure the impact of co-determination at the plant level. Because data on union membership and on whether a respondent works in an

[73]This result can be explained by the definition of overeducation being based on the years spent in formal education. The data reveal that among natives 33% have 13 or more years of formal education, while the same holds for about 13% of direct migrants.

establishment with a works council is available in the SOEP only in the years 2001, 2011, and 2019, we cannot include the works council information in our main specifications. In order to nevertheless ascertain the works council impact, we first re-estimate Equations 4.2 and 4.3 for these years. Doing so, we restrict the sample to individuals who provide valid information on the works council variable and work in establishments with at least five employees, as the relevant law entitles employees to establish works councils only in such establishments. The findings are qualitatively the same as depicted in Tables 4.3 and 4.4, indicating that the change in the sample is without impact. In a second step, we add a dummy variable indicating whether an individual works in an establishment with a works council (see Table C.5 in Appendix 4.8.3). The results remain qualitatively unaffected, and if at all, gain in quantitative strength. Therefore, we obtain no evidence that co-determination at the plant-level affects the relationship between union membership and educational mismatch.

Tables 4.3 and 4.4 and the robustness analyses reported above contain the findings from a correlation analysis. In Germany, there are no certification elections or comparable events that could be interpreted as exogenous variations in trade union membership (see, for example, DiNardo and Lee, 2004; Frandsen, 2021). There have not been substantial changes in the costs of union membership, for example, due to a variation in tax-related subsidies of union membership fees, which could be employed as instrumental variables (see, e.g., Barth et al., 2020; Dodini et al., 2021, 2023a,b). Therefore, the data do not allow us to supplement the correlation analysis with an investigation of causal effects. However, it seems unlikely that our results are affected by reverse causality, e.g., that a decline in overeducation or a greater likelihood of being educationally matched raises the gains from union membership or reduces its costs. Moreover, our findings with respect to overeducation and being educationally matched continue to hold after including extensive sets of demographic and job-related control variables. To further strengthen our findings, we additionally employ an entropy balancing approach (Hainmueller, 2012). When reestimating Equations 4.2 and 4.3, the results are in line with those presented in Tables 4.3 and 4.4 (see Table C.6 in Appendix 4.8.3). This qualitative equivalence also suggests that our specifications are not subject to an omitted variable bias.[74]

[74] In all three cases, the results are qualitatively similar when estimating the tobit-specification.

4.6.4 Heterogeneity

We have argued above that union membership can reduce educational mismatch because it, first, provides individuals with additional information about its consequences and the means to avoid such situations. Second, membership raises the bargaining power in negotiations with the employer and, thus, the expected gains of leaving a job for which an individual is not adequately qualified. To ascertain the relevance of the information and the bargaining power perspective, we subsequently consider the relationship between trade union membership and educational mismatch for subgroups differing in their degree of unionisation. The basic idea is that information about the consequences of educational mismatch can have an impact, irrespective of how many other employees belong to the trade union. Hence, if the informational component of union membership affects educational mismatch, no substantial group-specific differences should be discernible. In contrast, the expected costs of voicing discontent are likely to be lower, while the expected gains from leaving, especially in a situation of overeducation, are likely to be greater the more powerful trade unions are. Accordingly, the bargaining power perspective suggests that the union membership effect is stronger among groups of employees with higher unionisation rates.

According to our data (see Table C.7, in Appendix 4.8.4 for details), union density in Germany is substantially higher among males than females, among full-time compared to part-time employees, in the public than the private sector, and among blue-collar compared to white-collar workers. Union density in the western part of the country slightly exceeds the share of union members among employees in eastern Germany. Moreover, in two SOEP waves with union membership data there is information about collective bargaining coverage (i.e., in 2015 and 2019). We observe that union density is higher for employees covered by a collective bargaining agreement than for those who are uncovered.

To analyse the impact of union density we, furthermore, define $2^5 = 32$ groups of employees according to the respondent's gender, working time, sector, place of residence, and occupation.[75] For each group, we calculate union density, dividing the number of union members by the number of employees, leveraging the sample described in Section 4.5.3. The levels of union density computed in this manner range from 7.55% for the group of female, part-time, white-collar workers in the West German private sector, to 43.6% for the group of full-time,

[75]We do not incorporate collective bargaining in order to use observations for the entire time period.

male, blue-collar workers in the West German public sector. The value closest to the median accounts for 42.57% of the cumulative distribution and yields a density of 15.27%. Groups of employees featuring a density below the level of 42.57% are considered as low-density and those above as high-density groups of employees. Given these distinctions, we estimate Equation 4.3 separately for each of the $6 \times 2 = 12$ subgroups enumerated in the previous paragraph and for the subgroups of high- and low-density employees as defined above.

Table 4.5 depicts the average marginal effects from panel RE logit specifications. The results summarised in Table 4.4 for the entire sample concerning the relationship between union membership and overeducation, respectively (mis)match, can qualitatively also be observed for the subgroups of males, full-time and private sector employees, individuals residing in western Germany, and high-density employees. Trade union members covered by a collective wage agreement or in a blue-collar job are less likely to be undereducated and more likely to be educationally matched than their non-member counterparts. Finally, for white-collar employees, union membership is associated with an increase in the probability of being educationally matched by about the same magnitude as for blue-collar members.[76]

By comparing the different subgroups, we find a negative correlation between trade union membership and overeducation as well as a positive one with the probability of being educationally matched for some employees but not for others. This suggests that the relationship is not due to informational advantages of union members, as it seems implausible that trade unions could restrict information on how to avoid educational mismatch, for example, according to gender or working time. The evidence from Table 4.5 is more in line with the bargaining power perspective. This is evident from our observation of significant positive correlations between union membership and being educationally matched among males, full-time employees, blue-collar workers, and for individuals covered by a collective wage agreement. In addition, the likelihood of being overeducated is lower among males and full-time workers. In all of these subgroups, union membership is relatively widespread. The findings for the group of high-density employees in Panel G of Table 4.5 support this interpretation.

[76] If we estimate tobit specifications, the results for the subgroups considered in Panels A - C and F, G are qualitatively the same. For white-collar workers (Panel E), we find a significant negative correlation between trade union membership and overeducation in the tobit specification. In addition, we observe a negative correlation between union membership and overeducation for individuals living in eastern Germany (Panel D).

Table 4.5: Heterogeneity - Panel RE logit: Average marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Undereducated		Overeducated		Matched	
Panel A: Gender	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
Trade union	-0.00263	-0.00361	-0.0176***	-0.00452	0.0186***	0.0107
	(0.00433)	(0.00521)	(0.00466)	(0.00447)	(0.00663)	(0.00777)
	33249	31968	33249	31968	33249	31968
Panel B: Working time	**Part-time**	**Full-time**	**Part-time**	**Full-time**	**Part-time**	**Full-time**
Trade union	-0.00677	-0.00336	-0.00760	-0.0142***	0.0143	0.0181***
	(0.00746)	(0.00367)	(0.00629)	(0.00398)	(0.00984)	(0.00560)
	19664	45553	19664	45553	19664	45553
Panel C: Sector	**Private**	**Public**	**Private**	**Public**	**Private**	**Public**
Trade union	-0.00559	-0.00446	-0.0176***	-0.00170	0.0239***	0.00928
	(0.00407)	(0.00582)	(0.00378)	(0.00572)	(0.00603)	(0.00872)
	48408	16809	48408	16809	48408	16809
Panel D: Region	**West**	**East**	**West**	**East**	**West**	**East**
Trade union	-0.00322	-0.00557	-0.0160***	-0.00789	0.0203***	0.0138
	(0.00378)	(0.00598)	(0.00379)	(0.00530)	(0.00565)	(0.0108)
	51510	13707	51510	13707	51510	13707
Panel E: Occupation	**Blue**	**White**	**Blue**	**White**	**Blue**	**White**
Trade union	-0.0126***	-0.000529	-0.00621	-0.00397	0.0171**	0.0153**
	(0.00487)	(0.00414)	(0.00442)	(0.00401)	(0.00678)	(0.00703)
	26401	38816	26401	38816	26401	38816
Panel F: Coll. bargaining	**Uncovered**	**Covered**	**Uncovered**	**Covered**	**Uncovered**	**Covered**
Trade union	-0.00821	-0.0152**	0.000292	-0.00265	0.00933	0.0219**
	(0.0135)	(0.00654)	(0.00986)	(0.00669)	(0.0196)	(0.00956)
	9175	12204	9175	12204	9175	12204
Panel G: Union density	**High density**	**Low density**	**High density**	**Low density**	**High density**	**Low density**
Trade union	-0.00340	-0.00870	-0.0133***	-0.00449	0.0186***	0.0117
	(0.00351)	(0.00601)	(0.00399)	(0.00501)	(0.00584)	(0.00912)
Num. obs.	38123	27094	38123	27094	38123	27094

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019 for Panels A - E and Panel G and from 2015 and 2019 for Panel F. The further covariates are the same as those explained in Table 4.3 except for gender in Panel A, full-time in Panel B, public sector in Panel C, and white-collar in Panel E. Standard errors, clustered at individual level, presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The non-significant correlations for the public sector do not necessarily refute the bargaining power argument, as the consequences of overeducation and the role of union membership may differ in the public compared to the private sector. First, educational mismatch is less prevalent in the public sector (see Tables C.2 and C.3 in Appendix 4.8.2). Second, job descriptions in the public sector are usually relatively detailed, and the minimum level of qualification must be stated precisely. Moreover, the payment scheme is much more transparent in the public sector compared to the private sector (Deutscher Bundestag, 2019a,b). Therefore, the informational advantages associated with union membership are likely to be smaller than in the private sector. Third, promotion in the public sector often involves strong elements of experience or tenure. Moreover, the qualification requirements for promotions are usually non-negotiable and apply to all applicants for a vacant position. Accordingly, the impact of bargaining power is likely to be less relevant. Consequently, both the informational benefits of union membership and the resulting bargaining power may play a lesser role in the public than the private sector, and the scope for union membership effects may be smaller. This, in turn, implies that the irrelevance of individual union membership for educational (mis)match in the highly unionised public sector, as depicted in Panel C of Table 4.5, can be consistent with the role of bargaining power.

In sum, the subgroup analysis is largely compatible with the hypothesis that the bargaining power of union members is enhanced by a high percentage of members among employees, which, in turn, enables members of such high-density employee groups to avoid the adverse consequences of educational mismatch. Our investigation, however, finds no evidence that trade union members are better informed about the causes and consequences of educational mismatch compared to non-members.

4.7 Conclusion

We hypothesise that the probability of overeducation and its extent are lower among union members. Conversely, we expect undereducation to be higher among union members. Moreover, we anticipate the overeducation effect to dominate, implying that union members are more likely to be educationally matched than comparable non-members.

Using data from the German Socio-Economic Panel (SOEP) for the period 1985 to 2019, we obtain evidence for a negative correlation between an employee's trade union membership and the incidence of overeducation. We also find that union members exhibit a significantly lower extent of overeducation and a greater likelihood of being educationally matched than non-members. These relationships can be observed, in particular, for males, full-time and private sector employees, individuals residing in western Germany, and respondents working in high union-density surroundings. They are not discernible for females, part-time or public sector employees, and residents of Eastern Germany. As informational advantages of trade union membership would apply uniformly, irrespective of union coverage, the evidence is not consistent with an informational role of union membership but rather supports the bargaining power effect. We cannot identify a robust linkage between union membership and undereducation.

To illustrate the quantitative importance of the effects, the subsequent back-of-the-envelope calculation may be informative. According to Table 4.2, 13.14% of respondents are overeducated with an average duration of overeducation in the entire sample of 0.1583 years. By definition, undereducated and educationally matched individuals have zero years of overeducation. Therefore, each overeducated respondent has, on average, a (uncompleted) duration of overeducation of about 1.2 (=0.1583/0.1314) years. Union membership is associated with a reduction in the duration of overeducation by 0.150 years or 54.75 days, which represents a decrease of 15.2%. Consider this scenario: If the private costs of overeducation account for 5% of the annual gross wage (e.g., Chu Ng, 2003), a union member would experience a reduction in these costs of more than 0.75% (5% x 0.150) of their gross wage. Since the membership fee amounts to 1% of the gross wage, mitigating the negative impacts of educational mismatch can serve as a compelling incentive to join a trade union.

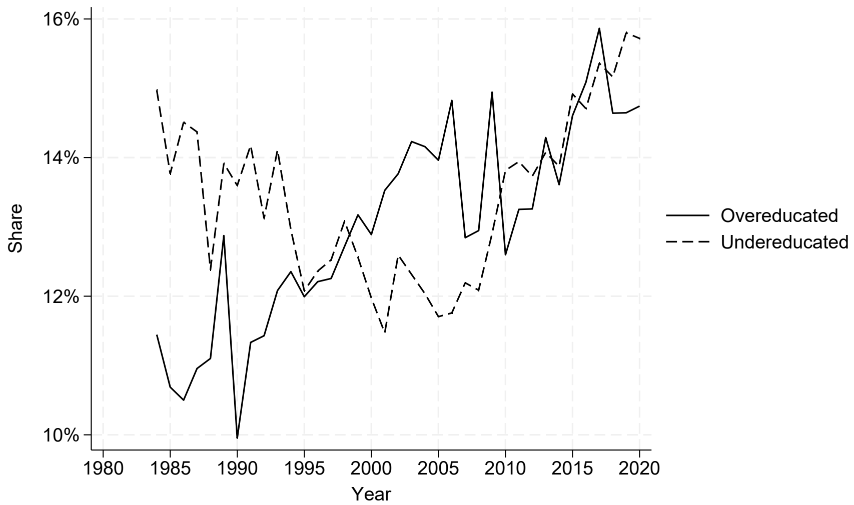
Other analyses that consider trade union membership as one of many determinants of educational mismatch predominantly observe positive correlations. Our findings indicate a negative correlation between union membership and both overeducation and educational mismatch. Whether this difference in findings is due to specific characteristics of the German labour market, trade unions having a different role compared to other countries, or differences in considered employee groups are all issues that we cannot tackle with the data at hand.

While the SOEP data allow for a thorough investigation of the issue under scrutiny, they also have their limitations, as we indicated above. Most importantly, since information on union membership is only available for selected waves of SOEP data, we cannot identify the year an employee joined or left a trade union. Moreover, observed changes in union membership status and situations of educational mismatch are relatively rare. As a result, our data do not enable us to identify fixed effects specification.

4.8 Appendix C

4.8.1 Trends in educational mismatch

Figure C.1: Percentage of overeducated and undereducated individuals in Germany: 1984 to 2020



Note: This graph illustrates the percentage of overeducated (undereducated) individuals as determined by the statistical approach, utilising all available SOEP waves spanning from 1984 to 2020. Individuals who are older than the legal retirement age, those unemployed, and (solo-) self-employed are excluded.

4.8.2 Extended tables

Table C.1: Extended summary statistics differentiated by trade union membership

	All		TUM=0		TUM=1		T-test
	Mean	Sd	Mean	Sd	Mean	Sd	
Trade union	0.1906	0.3927					
Female	0.4902	0.4999	0.5244	0.4994	0.3449	0.4753	-0.180***
Direct migrant	0.1496	0.3567	0.1508	0.3578	0.1447	0.3518	-0.0061
Indirect migrant	0.0512	0.2204	0.0537	0.2255	0.0405	0.1971	-0.0133***
No migrant	0.7992	0.4006	0.7955	0.4034	0.8149	0.3884	0.0194***
Age	41.7915	11.4990	41.2406	11.5680	44.1317	10.8955	2.891***
Single	0.2643	0.4410	0.2792	0.4486	0.2009	0.4007	-0.0783***
Married or partnered	0.6187	0.4857	0.6038	0.4891	0.6821	0.4657	0.0783***
Separate or widowed	0.1170	0.3214	0.1170	0.3214	0.1170	0.3214	0.0000
Number of children	1.4891	1.1905	1.4791	1.2012	1.5313	1.1428	0.0522***
White-collar	0.5952	0.4909	0.6295	0.4829	0.4493	0.4974	-0.180***
Civil servant	0.0566	0.2311	0.0453	0.2079	0.1048	0.3063	0.0595***
Prior unemployment	0.3582	0.4795	0.3852	0.4867	0.2432	0.4291	-0.142***
Permanent	0.8736	0.3323	0.8607	0.3462	0.9283	0.2580	0.0676***
Full-time	0.6985	0.4589	0.6682	0.4709	0.8273	0.3780	0.159***
Public sector	0.2577	0.4374	0.2390	0.4265	0.3371	0.4728	0.0981***
Tenure	10.3390	9.9632	9.2426	9.3176	14.9962	11.1959	5.754***
Firm size: $<$ 20	0.2251	0.4176	0.2605	0.4389	0.0747	0.2629	-0.186***
Firm size: 20 - 199	0.2890	0.4533	0.3045	0.4602	0.2235	0.4166	-0.0809***
Firm size: 200 - 1999	0.2288	0.4200	0.2166	0.4119	0.2804	0.4492	0.0638***
Firm size: $\geq$ 2000	0.2571	0.4370	0.2184	0.4132	0.4214	0.4938	0.203***
Ind: Agriculture	0.0119	0.1083	0.0134	0.1151	0.0052	0.0721	-0.00820***
Ind: Energy	0.0097	0.0982	0.0084	0.0910	0.0156	0.1240	0.00726***
Ind: Manufacturing	0.2641	0.4408	0.2359	0.4246	0.3838	0.4863	0.148***
Ind: Mining	0.0034	0.0584	0.0016	0.0399	0.0112	0.1052	0.00959***
Ind: Construction	0.0608	0.2389	0.0651	0.2467	0.0424	0.2015	-0.0227***
Ind: Trade	0.1440	0.3511	0.1590	0.3657	0.0804	0.2719	-0.0786***
Ind: Transport	0.0512	0.2204	0.0415	0.1994	0.0926	0.2899	0.0511***
Ind: Bank	0.0398	0.1956	0.0442	0.2056	0.0212	0.1439	-0.0231***
Ind: Services	0.3837	0.4863	0.4003	0.4900	0.3129	0.4637	-0.0874***
Ind: Other	0.0314	0.1743	0.0306	0.1722	0.0347	0.1830	0.00409*
ISCO: LegSenOfMan	0.0448	0.2069	0.0476	0.2128	0.0332	0.1792	-0.0143***
ISCO: Prof	0.1399	0.3469	0.1444	0.3515	0.1209	0.3260	-0.0235***
ISCO: TechAssProf	0.2487	0.4323	0.2530	0.4347	0.2305	0.4212	-0.0225***
ISCO: Clerks	0.1293	0.3356	0.1336	0.3403	0.1110	0.3142	-0.0226***
ISCO: Service	0.1224	0.3278	0.1323	0.3388	0.0805	0.2721	-0.0518***
ISCO: AgriculFish	0.0079	0.0886	0.0084	0.0912	0.0059	0.0764	-0.00252**
ISCO: CraftTrade	0.1467	0.3539	0.1309	0.3373	0.2141	0.4102	0.0832***
ISCO: PlantMachine	0.0810	0.2728	0.0675	0.2509	0.1382	0.3451	0.0706***
ISCO: Elementary	0.0791	0.2700	0.0823	0.2748	0.0657	0.2477	-0.0167***
Num. obs.	65217		52789		12428		

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Panel RE tobit - Extended results

Years of...	(1) Undereducation	(2)	(3) Overeducation	(4)
Trade union	-0.0880***	-0.0267	-0.228***	-0.150***
	(0.0281)	(0.0252)	(0.0394)	(0.0340)
Female		0.000482		-0.170***
		(0.0281)		(0.0387)
Direct migrant		1.283***		-1.019***
		(0.0328)		(0.0559)
Indirect migrant		0.564***		-0.374***
		(0.0500)		(0.0751)
Age		-0.0970***		0.0809***
		(0.00615)		(0.00801)
Age (Sq/100)		0.115***		-0.0975***
		(0.00705)		(0.00928)
Married or partnered		-0.000208		-0.143***
		(0.0321)		(0.0365)
Separate or widowed		0.102**		-0.332***
		(0.0416)		(0.0514)
Number of children		0.0621***		-0.0477***
		(0.0111)		(0.0155)
White-collar		-0.565***		0.665***
		(0.0262)		(0.0361)
Civil servant		-1.241***		0.875***
		(0.0610)		(0.0814)
Prior unemployment		-0.0135		0.0609*
		(0.0250)		(0.0326)
Permanent		-0.221***		0.104***
		(0.0271)		(0.0343)
Full-time		-0.163***		0.0759***
		(0.0233)		(0.0291)
Public sector		0.0758***		-0.128***
		(0.0270)		(0.0337)
Tenure		0.00630**		-0.0312***
		(0.00291)		(0.00375)
Tenure (Sq/100)		0.0161**		0.0257**
		(0.00787)		(0.0113)
Firm size: 20-199		-0.0299		0.112***
		(0.0250)		(0.0319)
Firm size: 200-1999		-0.0569**		0.159***
		(0.0278)		(0.0359)
Firm size: $\geq$2000		-0.112***		0.272***
		(0.0288)		(0.0371)
Month & year FE	X	X	X	X
Further covariates		X		X
Num. obs.	65217	65217	65217	65217

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. Further covariates (not displayed): Dummies indicating federal state, industries and occupation. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Panel RE logit: Average marginal effects - Extended results

	(1)	(2)	(3)	(4)	(5)	(6)
	Undereducated		Overeducated		Matched	
Trade union	-0.00969***	-0.00422	-0.0177***	-0.0129***	0.0306***	0.0188***
	(0.00298)	(0.00332)	(0.00316)	(0.00318)	(0.00474)	(0.00501)
Female		-0.00256		-0.0165***		0.0264***
		(0.00344)		(0.00305)		(0.00506)
Direct migrant		0.186***		-0.0669***		-0.124***
		(0.00602)		(0.00368)		(0.00692)
Indirect migrant		0.0658***		-0.0269***		-0.0446***
		(0.00683)		(0.00524)		(0.00955)
Age		-0.0140***		0.00393***		0.0118***
		(0.000870)		(0.000808)		(0.00127)
Age (Sq/100)		0.0166***		-0.00450***		-0.0146***
		(0.000997)		(0.000948)		(0.00148)
Married or partnered		0.000852		-0.0140***		0.0230***
		(0.00420)		(0.00365)		(0.00615)
Separate or widowed		0.0171***		-0.0332***		0.0300***
		(0.00552)		(0.00510)		(0.00815)
Number of children		0.00743***		-0.00358***		-0.00443**
		(0.00135)		(0.00121)		(0.00201)
White-collar		-0.0792***		0.0680***		0.0154***
		(0.00375)		(0.00386)		(0.00538)
Civil servant		-0.152***		0.0831***		0.0735***
		(0.00810)		(0.00702)		(0.0114)
Prior unemployment		-0.00712**		0.00972***		-0.00359
		(0.00323)		(0.00276)		(0.00472)
Permanent		-0.0348***		0.0166***		0.0232***
		(0.00367)		(0.00365)		(0.00566)
Full-time		-0.0230***		0.00382		0.0269***
		(0.00322)		(0.00287)		(0.00480)
Public sector		0.00859**		-0.0148***		0.0102*
		(0.00357)		(0.00343)		(0.00558)
Tenure		0.00116***		-0.00232***		0.00217***
		(0.000411)		(0.000393)		(0.000614)
Tenure (Sq/100)		0.00171		0.000561		-0.00454***
		(0.00109)		(0.00116)		(0.00172)
Firm size: 20-199		-0.00533		0.00854***		-0.00273
		(0.00342)		(0.00316)		(0.00515)
Firm size: 200-1999		-0.00643*		0.0164***		-0.0142**
		(0.00378)		(0.00363)		(0.00576)
Firm size: $\geq$ 2000		-0.0154***		0.0270***		-0.0148**
		(0.00391)		(0.00363)		(0.00584)
Month & year FE	X	X	X	X	X	X
Further covariates		X		X		X
Num. obs.	65217	65217	65217	65217	65217	65217

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. Further covariates (not displayed): Dummies indicating federal state, industries and occupation. Standard errors, clustered at individual level, in parentheses. The average marginal effects account for the discrete change from the base level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

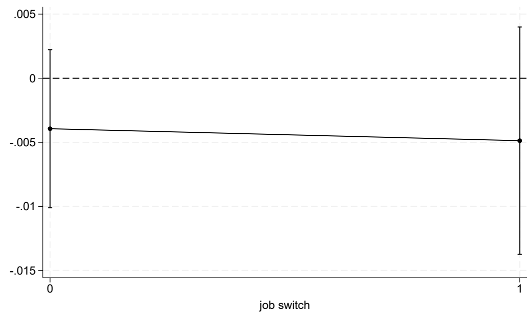
4.8.3 Robustness checks

Table C.4: Panel RE logit: Trade union interaction with job switching

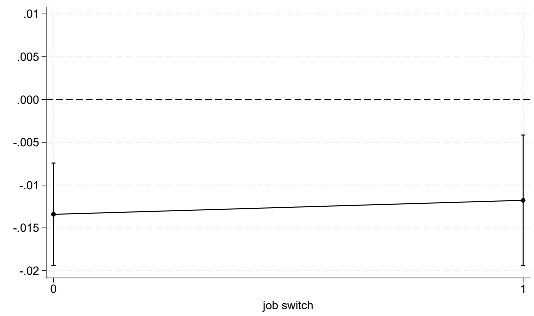
	(1) Undereducated	(2) Overeducated	(3) Matched
Trade union	-0.0991	-0.335***	0.202***
	(0.0951)	(0.0923)	(0.0614)
Job switching	0.0185	-0.0428	0.0321
	(0.0690)	(0.0619)	(0.0437)
Trade union × job switching	-0.0233	0.0397	-0.00701
	(0.152)	(0.132)	(0.0935)
Month & year FE	X	X	X
Further covariates	X	X	X
Num. obs.	65217	65217	65217

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. Job switching is identified based on tenure. Due to the interaction term, Table C.4 does report point estimates. Average marginal effects are displayed in Figure C.2. The further covariates are the same as those listed in Table 4.3. Standard errors, clustered at individual level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

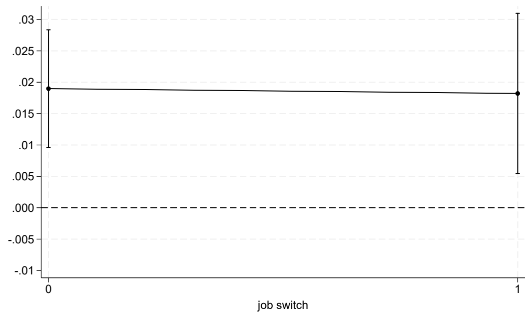
Figure C.2: Panel RE logit: Average marginal effects - Trade union interaction with job switching



(a) Undereducation



(b) Overeducation



(c) Matched

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. This figure displays average marginal effects of the interaction between trade union membership and job switching identified based on tenure. The average marginal effects account for the discrete change from the base level. 90% confidence intervals plotted.

Table C.5: Panel RE logit: Average marginal effects - Including works council information

	(1)	(2)	(3)	(4)	(5)	(6)
	Undereducated		Overeducated		Matched	
Trade union	-0.00932	-0.00880	-0.0227***	-0.0245***	0.0324***	0.0342***
	(0.00600)	(0.00607)	(0.00649)	(0.00647)	(0.00864)	(0.00869)
Council exists		-0.00391		0.0125**		-0.0130
		(0.00573)		(0.00580)		(0.00812)
Month & year FE	X	X	X	X	X	X
Further covariates	X	X	X	X	X	X
Num. obs.	20070	20070	20070	20070	20070	20070

Note: The sample is based on SOEP data from 2001, 2011, and 2019. The further covariates are the same as those listed in Table 4.3. Standard errors, clustered at individual level, in parentheses. The average marginal effects account for the discrete change from the base level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: Logit with entropy balancing weights: Average marginal effects

	(1)	(2)	(3)
	Undereducated	Overeducated	Matched
Trade union	-0.00615	-0.0110**	0.0178**
	(0.00525)	(0.00512)	(0.00690)
Month & year FE	X	X	X
Further covariates	X	X	X
Num. obs.	33325	33325	33325

Note: The sample is based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019. To avoid biases from non-constant weights within individuals, the sample is restricted to the first observation per individual and a cross-sectional logit is estimated. The entropy balancing weights are estimated based on gender, migration background, age, civil status and the number of children. The further covariates are the same as those listed in Table 4.3. Standard errors, clustered at individual level, in parentheses. The average marginal effects account for the discrete change from the base level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.8.4 Heterogeneity analysis

Table C.7: Trade union density among various groups of employees

Panel A: Gender	**Males**	**Females**
	24.49%	13.41%
Num. obs.	33249	31968
Panel B: Working time	**Part-time**	**Full-time**
	10.91%	22.57%
Num. obs.	19664	45553
Panel C: Sector	**Private**	**Public**
	17.02%	24.93%
Num. obs.	48408	16809
Panel D: Region	**West**	**East**
	19.42%	17.69%
Num. obs.	51510	13707
Panel E: Occupation	**Blue-collar**	**White-collar**
	25.92%	14.39%
Num. obs.	26401	38816
Panel F: Coll. bargaining	**Uncovered**	**Covered**
	6.30%	22.62%
Num. obs.	9175	12204

Note: Own calculations based on SOEP data from 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011, 2015, and 2019 for Panels A - E and from 2015 and 2019 for Panel F.

Chapter 5

Unexpected fortunes: Exploring the impact of windfall gains on educational mismatch*

This study is the first to examine the effect of unearned income on educational mismatch, providing a conceptual and empirical analysis. Using data from the German Socio-Economic Panel and dynamic two-way fixed effect models, the results reveal a consistent increase in the likelihood of overeducation following a windfall gain. This effect is driven by individuals switching within high-skill occupations or changing from high-skill to low-skill jobs. For the latter, switches are accompanied by reduced working and overtime hours. The findings are particularly pronounced for medium and large windfalls, individuals with higher household income satisfaction, and younger respondents. Contrary to expectations, the results show no negative correlation between windfall gains and the likelihood of undereducation.

Keywords: Windfall gains; Educational mismatch; Event study; SOEP

*This chapter is joint work with Marco Clemens. We are grateful for helpful comments by Laszlo Goerke, Melanie Arntz, Lisa-Marie Duletzki, Freya Cook, and the participants of the TriECON Workshop on Income, Pay Composition and Labour Market Outcomes, the 16th Interdisciplinary IAB Ph.D. Workshop on Perspectives on Unemployment, the 15th International German Socio-Economic Panel User Conference (SOEP 2024), and the 16th Workshop on Labour Economics.

5.1 Introduction

If private wealth rises due to unearned income, such as inheritances, gifts, or lottery wins, referred to as windfall gains, the relative reward of working to create labour income decreases (see Doorley and Pestel, 2020). As previous studies have reported, such unexpected income shocks induce labour responses such as an adjustment in working hours, job changes, transitions to self-employment, or even a drop out of the labour market (see, e.g., Flèche et al., 2021; Georgellis et al., 2005; Golosov et al., 2024; Holtz-Eakin et al., 1993; Imbens et al., 2001; Joulfaian and Wilhelm, 1994; Lindh and Ohlsson, 1996). These labour responses are partially driven by an increased willingness to trade monetary against non-monetary job conditions and the diminishing marginal utility of labour income as wealth increases (e.g., Doorley and Pestel, 2020; Haywood, 2016). Moreover, job satisfaction tends to increase following substantial windfall gains, particularly among individuals who switch jobs afterward (Haywood, 2016). Still, it remains unclear whether individuals simultaneously adjust their job match quality. Such adjustments may arise if a significant income shock encourages transitions to positions that, while potentially misaligned with their qualifications, better align with personal preferences or interests, for example, in terms of job content.

Educational mismatch provides a framework to explore such adjustments, distinguishing between overeducation, where an individual's education exceeds job requirements, and undereducation, where education falls short (e.g., Acemoglu, 1999; Freeman, 1976; Rumberger, 1981a; Sicherman and Galor, 1990; Tsang and Levin, 1985). Overeducation is generally perceived as an unfavourable employment situation, while undereducation is deemed more favourable (see, e.g., Sloane et al., 1996). This is primarily based on two monetary aspects: First, overeducation presents an over-investment in education, inducing expenditures by the individual and by society (McGuinness, 2006).[77] Second, direct costs are imposed on the individual by overeducation due to the overeducation-pay-penalty. This term describes the phenomenon of overeducated individuals earning less than equally qualified individuals working in an adequate position. In contrast, undereducated individuals experience a wage-premium (see, e.g., Caroleo and Pastore, 2018; Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989). However, when individuals are subject to monetary shocks, such as windfall gains, these perceptions of overeducation and undereducation may shift.

[77] Particularly if education is free as in Germany and financed in large parts through the tax authorities (see, e.g., Brugger et al., 2022).

Due to the lack of research on the relationship between windfall gains and educational mismatch, the expectations are *ex ante* unclear. To address this gap, we employ a simplified utility function approach to define hypotheses on the impact of windfall gains on the likelihood of overeducation and undereducation. This approach predicts that without windfall gains, the utility offered by an undereducated position is largest if the wage premium exceeds the disutility loss from working (e.g., due to potentially longer hours).[78] In contrast, a windfall gain reduces the impact of the wage premium on this mechanism, such that with sufficient size of the windfall, the disutility associated with working in undereducation becomes the decisive part of the inequality (cf., Haywood, 2016). Thus, we expect that individuals' likelihood of working in undereducation reduces after a windfall gain, while the likelihood of working in overeducation is expected to rise.

We empirically test these hypotheses by using German data from the Socio-Economic Panel from 2000 to 2020. We find evidence consistent with our expectations regarding overeducation by applying dynamic two-way fixed effects estimations based on Sun and Abraham (2021). Specifically, the likelihood of overeducation consistently increases following a windfall gain. In line with the formalisation where we model educational mismatch as a function of job choice, this increase is driven by individuals switching jobs from high-skill to low-skill or within high-skill occupations in the aftermath of a windfall. Those job changes are also characterised by reduced working and overtime hours after the windfall gain, indicating that individuals trade an inferior job match quality for leisure time as expected based on the formalisation. Consistent with the expectations and prior studies (e.g., Holtz-Eakin et al., 1993), the results show that the size of the windfall gain is crucial, as small gains do not affect the likelihood of overeducation. Finally, heterogeneity analyses reveal that the reported effects are particularly pronounced for individuals with higher household income satisfaction and younger individuals aged between 18 and 43.

We contribute to two strands of literature. First, we add to the literature on the effects of income shocks induced through windfall gains on labour outcomes (e.g., Belloc et al., 2025; Cabanillas-Jiménez, 2024; Doorley and Pestel, 2020; Elinder et al., 2012; Flèche et al., 2021; Georgellis et al., 2005; Golosov et al., 2024; Haywood, 2016; Holtz-Eakin et al., 1993; Imbens et al., 2001; Joulfaian and Wilhelm, 1994; Kindermann et al., 2020; Lindh and Ohlsson, 1996; Malo and Sciulli, 2021; Schäfer et al., 2011; Sila and Sousa, 2014; Taylor, 2001). Second,

[78]This supposition aligns with previous assumptions by e.g., Sloane et al. (1996) noting that undereducation should be the most preferred state.

our study extends the knowledge on the determinants of educational mismatch (e.g., Akgüç and Parasnis, 2023; Aleksynska and Tritah, 2013; Baran, 2024; Battu et al., 1999; Büchel and Pollmann-Schult, 2004; Carroll and Tani, 2015; Davia et al., 2017; Di Pietro and Cutillo, 2006; Fleming and Kler, 2008, 2014; Liu et al., 2021; McGoldrick and Robst, 1996; Ordine and Rose, 2009, 2011; Robst, 1995b; Santiago-Vela and Mergener, 2022; Tarvid, 2015).

The study proceeds as follows: Section 5.2 derives our hypotheses based on the formal comparison of utilities before and after the windfall gain. Afterward, we provide information on the dataset, the variables of interest, and the empirical strategy in Section 5.3. The results are presented in Section 5.4, which additionally contains robustness checks, and in-depth analyses of the role of occupational choice, the working conditions, and the size of the windfall gain as well as heterogeneity checks. Finally, Section 5.5 provides a summary of the findings and concludes the study.

5.2 Hypotheses

The effects of windfall gains on educational mismatch can be illustrated using a simplified trade-off between leisure and consumption. We assume an individual chooses from four job options, $j \in \{ue, m, oe, 0\}$, where each is associated with distinct educational requirements. For jobs of type j_{ue} the educational requirements exceed the individual's attained education, resulting in undereducation, while jobs of type j_{oe} have requirements below the individual's education, leading to overeducation. Jobs of type j_m represent a match between educational attainment and job requirements. The outside option is denoted by j_0, where the individual is not part of the labour force.

We assume that individuals aim to maximise utility. In this model, utility is a function of two primary factors, consumption and leisure, formalised as:

$$U = f(L, C) \tag{5.1}$$

Leisure L is derived from the time available after accounting for working hours h_j: $L = T - h_j$. The extent of working hours, and consequently the disutility associated with work, likely varies by job type, j. Theoretically, it can be posited that individuals lacking the required education for a given job will need more time to complete tasks compared to those with adequate training. The scarce empirical research on the productivity of mismatched workers in Europe supports this supposition (see, e.g., Büchel, 2002; Grunau, 2016; Jacobs et al., 2023). Building on this, we assume that the working hours required to perform a job h_j are highest for undereducated individuals and lowest for overeducated individuals, such that $h_{ue} > h_m > h_{oe}$.[79] In the case of the outside option, j_0, $h_0 = 0$.

Consumption C in Equation 5.1 is defined as a function of total wealth Y, formulated as $C = Y = w_j + V$. w_j is the wage associated with job j and V represents non-labour income accumulated in the respective year. Consistent with prior findings on wage differences related to educational mismatch (e.g., Caroleo and Pastore, 2018; Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989), we define $w_{ue} > w_m > w_{oe} > w_0$ where the difference between, e.g., w_{ue} and w_m represents a compensating differential (e.g., Alba-Ramirez, 1993; McGuinness and Sloane, 2011; Rosen, 1986; Smith, 1776). w_0 represents the wage of the outside option (not working) and, thus, $w_0 = 0$.

Using a Cobb-Douglas utility function with normalised preferences a for leisure and b for consumption yielding $a + b = 1$, we define:

$$U_j = (T - h_j)^a (w_j + V)^b \tag{5.2}$$

Without a windfall gain (i.e., $V = 0$), the model predicts a choice for undereducation if:

$$(T - h_{ue})^a w_{ue}^b > (T - h_m)^a w_m^b \tag{5.3}$$

$$\underbrace{\left[\frac{T - h_{ue}}{T - h_m}\right]^a}_{<1} > \underbrace{\left[\frac{w_m}{w_{ue}}\right]^b}_{<1} \tag{5.4}$$

This condition shows that undereducation yields a higher utility when the wage premium outweighs the disutility of working h_{ue} hours in undereducation. Similar comparisons apply for overeducation versus matched positions.

Our primary interest is examining how windfall gains, $V > 0$, influence educational mismatch. Using Equation 5.4 to compare job types j_{ue} and j_m with

[79] To the best of our knowledge, no empirical research has yet examined the impact of educational mismatch on actual working hours for individuals in the same occupation. The only related evidence comes from Fleming and Kler (2008) and Fleming and Kler (2014), who examine satisfaction with working hours.

windfall gains, we obtain:

$$\underbrace{\left[\frac{T-h_{ue}}{T-h_m}\right]^a}_{<1} > \left[\frac{w_m+V}{w_{ue}+V}\right]^b \tag{5.5}$$

While the left side remains smaller than 1, the right side monotonically increases with V approaching 1. Thus, the inequality no longer holds if V gets sufficiently large. In sum, a larger V diminishes the marginal utility of the wage premium of undereducation, reversing the decision in favour of the matched position.

The critical value of V at which the individual will decide for the matched position over undereducation can be identified by isolating V, yielding:

$$\frac{\xi w_{ue}-w_m}{1-\xi} < V \text{ with } \xi = \underbrace{\left[\frac{T-h_{ue}}{T-h_m}\right]^{\frac{a}{b}}}_{0<\xi<1} \tag{5.6}$$

In Equation 5.6, ξ approaches 1 if the difference between the disutility of working h_{ue} hours in undereducation and h_m hours in a matched position diminishes. A rise in ξ implies that V must be larger to reverse the individual's decision in favour of the matched position. Similarly, V must be larger for the individual to decide for the matched position if the compensating differential in undereducation is larger. Contrarily, a lower wage premium of undereducation, indicated by a smaller difference between w_{ue} and w_m would necessitate a lower V to reverse the individual's decision in favour of the matched position.[80] Similar mechanisms apply when comparing utilities offered by a matched and an overeducated position. Simulations are displayed in Appendix 5.6.1.

Based on this simplified formalisation, we hypothesise the following effects of windfall gains on educational mismatch:[81]
Hypothesis 1: Windfall gains will reduce the likelihood of undereducation.
Hypothesis 2: Windfall gains will increase the likelihood of overeducation.

[80] This aligns with previous findings that an (expected) increase in wealth increases the demand for leisure (e.g., Basiglio et al., 2023).

[81] In this simple model, the individual chooses the outside option of not working (w_0) if $T^aV^b < (T-h_j)^a(w_j+V)^b$. The critical value of V at which the individual decides not to work increases when the disutility from working in any potential job, proxied by h_j, diminishes. For empirical evidence showing that large windfalls V can induce individuals to drop out of the labour market, see, e.g., Taylor (2001).

5.3 Data and method

We employ data from the Socio-Economic Panel (version 37, DIW, 2022; Goebel et al., 2019) to evaluate the impact of windfall gains on educational mismatch. While the longitudinal household survey has been performed since 1984, information on windfall gains is only contained for the waves 2000 to 2020.

5.3.1 Educational mismatch

In our preferred specification, we define educational mismatch based on the average education achieved in the reference group as proposed by Clogg and Shockey (1984) and Verdugo and Verdugo (1989). Importantly, within this approach, educational mismatch is a function of labour demand and supply (Hartog, 2000). Thus, this assessment not only allows for quick adjustments to changes in the labour market but also defines over- and undereducation relative to one's peers (Capsada-Munsech, 2019).[82] For each survey wave t we estimate the average years spent in education $\overline{x}$ for each occupational group o defined by three-digit ISCO codes (International Labour Office, 2012) following Blásquez and Budría (2012).[83] If the individuals' years of education exceed this reference by more than one standard deviation σ_{ot} they are defined to be overeducated.

$$OE_{ito} \begin{cases} 1 & \text{if } x_{it} > \overline{x}_{ot} + \sigma_{ot} \\ 0 & \text{otherwise} \end{cases} \tag{5.7}$$

In contrast, one is deemed undereducated if the attained education falls short of this mean by more than one standard deviation σ_{ot}.

$$UE_{ito} \begin{cases} 1 & \text{if } x_{it} < \overline{x}_{ot} - \sigma_{ot} \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$

[82] Still, this approach usually results in a more conservative estimate of the share of over- and undereducated individuals compared to alternative measures (e.g., Blásquez and Budría, 2012). To account for this, we apply robustness checks using alternative definitions of educational mismatch.

[83] To give an idea of the level of detail in the three-digit ISCO codes, consider the first major ISCO group "Managers". Managers are divided into four two-digit groups distinguishing between "Chief Executives, Senior Officials and Legislators", "Administrative and Commercial Managers", "Production and Specialized Services Managers", and "Hospitality, Retail and Other Services Managers". For each of these groups, the three-digit ISCO code further differentiates between occupations. Within the last group, for example, the three-digit groups disentangle "Hotel and Restaurant Managers" from "Retail and Wholesale Trade Managers" or "Other Service Managers".

5.3.2 Windfall gains

From 2000 to 2020, the SOEP includes the following survey question: "Did you or another member of the household receive a large sum of money or other assets (car, house, etc.) as an inheritance, gift, or lottery winnings last year?"[84] The respondent can then choose the type of payment they received or respond that they did not receive any payment. In our analysis, we include all three types of windfalls to maintain a large number of observations for the relatively demanding empirical estimation strategy. However, most windfall gains in our sample stem from inheritances or gifts.[85] While providing more power to the estimation, using an aggregate measure also has some drawbacks: For instance, gifts and inheritances may be expected by receivers and occur simultaneously with other events that may influence labour-related outcomes. The data and methods used in this study allow us to investigate these anticipation effects (see Section 5.4). Although the empirical investigation of anticipation effects is essential, they would mean, if present, that individuals adjust their behaviour prior to actual receipt. This would downward bias our estimates. Hence, the presented coefficients can be considered as lower bounds.

Until 2005, individuals were asked about windfalls that exceed EUR 2,500. Thereafter, this threshold fell to EUR 500. To retain a consistent time series, we define the incidence of windfall gains $I_{i,t}$ as equal to 1 if the respondent reports a windfall that exceeds the EUR 2,500 mark, and 0 otherwise. Importantly, we would not expect that a windfall gain below that threshold would lead to a significant labour response.[86] We deflate the amount of the windfall gain $S_{i,t}$ using the consumer price index with base year 2010.[87]

[84] It should be noted that the data do not allow any distinction to be made as to how these profits are used, i.e. whether they are spent, invested or saved.

[85] The number of lottery winners and the quantity of the prizes in the German context is quite small. Over the survey years, we observe only 400 lottery winners in the SOEP, which is low compared to other studies (see, for example, Golosov et al. (2024) for recent US evidence using a sample of around 90,000 winners).

[86] Referring to our model, this is easily shown using the most extreme case, where the disutilities of working in undereducation and a matched position converge, i.e, $\xi \rightarrow 1$. In this case, the individual would be expected to choose the matched position over undereducation only if the compensating differential in undereducation compared to the matched job was smaller than EUR 2,500 or $\sim$ EUR 210 a month.

[87] The raw size of the windfall may be perceived differently depending on the size of the household. For instance, single households likely have a stronger response to a EUR 50,000 gain, which is roughly the sample mean, than a household of a married couple with 2 children. To accommodate for that, in alternative specifications, we normalise $S_{i,t}$ by the number of persons living in the household in year t and obtain a relative measure of windfall size. We use that measure as a robustness measure in Section 5.4.3.

5.3.3 Empirical approach

To estimate the effect of windfall gains on the likelihood of overeducation (undereducation), we begin with the following dynamic two-way fixed effect (TWFE) model:[88]

$$Y^j_{i,t} = \delta_i + \lambda_t + \alpha_1 X'_{it} + \sum_{l=-K}^{-2} \beta_l D^l_{it} + \sum_{l=0}^{L} \beta_l D^l_{it} + \epsilon_{it} \tag{5.9}$$

with $Y^j_{i,t}$ as the outcome of interest. *j* indicates whether overeducation, OE_{ito}, or undereducation, UE_{ito}, is used as an outcome. Furthermore, δ_i is the individual and λ_t the time-fixed component. D^l_{it} represents the relative time period l of individual i with respect to the period of the windfall gain ($l = 0$). Furthermore, we consider a number of control variables X'_{it} that may correlate with both the timing of the windfall and the outcome (see Section 5.3.4). β_l, with $l > 0$, indicates the coefficients of interest and captures the influence of windfall gains on educational mismatch conditional on confounders, common trends, and individual time-invariant components.

However, recent literature on dynamic estimators (for a summary see Roth et al., 2023) suggests that the above estimator may be problematic in our setup. The receipt of windfalls causes an exogenous income shock, which occurs at various points in time. When estimating the effects of such a staggered treatment, dynamic two-way fixed effect estimators, as proposed in Equation 5.9, are likely to produce biased results. Aside from comparisons between treated and non-treated individuals, comparisons between individuals who have both already received treatment lead to negative weighting issues, which can cause such bias (Roth et al., 2023). To adjust for this issue, we make use of the method proposed by Sun and Abraham (2021), which allows for staggered treatment and treatment effect heterogeneity.

Specifically, we estimate the following average treatment effects ($\widehat{ATT}(g,t)$):

$$\widehat{ATT}(g,t) = \frac{1}{N_g} \sum_{i:G_i=g} \left[Y^j_{i,t} - Y^j_{i,g-1}\right] - \frac{1}{N_{\mathcal{G}_{comp}}} \sum_{i:G_i \in \mathcal{G}_{comp}} \left[Y^j_{i,t} - Y^j_{i,g-1}\right] \tag{5.10}$$

[88] All estimations are pursued with the software R.

where t indicates the survey year and g indicates the year in which employees reported a windfall.[89]

G_{comp} indicates the control group. As proposed by Sun and Abraham (2021), in the main specification, we use the last-treated cohort as the control group ($G_{comp} = max_i G_i = 2020$). While the standard estimator can also employ the never-treated individuals as the control group ($G_{comp} = \infty$), we prefer a specification that includes only individuals who, at some point, experience a windfall gain. This approach enhances comparability by controlling for other unobserved characteristics that may differ between those who do and do not receive such gains.[90]

Based on the $\widehat{ATT}(g,t)$ across adoption cohorts, we then estimate the average of the treatment effect l periods after the windfall receipt, weighted by the sample share of each cohort w_g:

$$ATT_l^w = \sum_g w_g \; ATT(g, g+l) \tag{5.11}$$

Dynamic estimates based on Sun and Abraham (2021) rely on parallel trends and no anticipation assumptions. In addition to visually inspecting event-study graphs in that setup, we address those assumptions using different control groups and model specifications that are sensitive under staggered treatment and treatment effect heterogeneity in Section 5.4.

5.3.4 Covariates and final estimation sample

As indicated in the previous section, the SOEP data contains information on confounders that can be linked to both the receipt of windfall gains and educational mismatch. We use information on age (and a squared term), the number of children living in the household, marital status (yes or no), and the logarithm of the total number of persons in the household.[91] *A priori*, one may

[89] For example, $\widehat{ATT}(2001, 2010)$ indicates the average treatment effect in 2010 for agents that reported a windfall in 2001.

[90] In alternative specifications, we also adjust the control group to the last five cohorts and the never-treated individuals and show that the main results are not qualitatively driven by the choice of the counterfactual.

[91] The sample shows a substantial association between household size and the number of children, with a correlation coefficient of about 0.7. However, they may still measure something different, as the logged number of persons in the households captures declining marginal responses with increasing household size. It also adjusts for other life events like divorce and the death of parents who lived in the same household.

predict that people will, for instance, generally receive inheritances later in life when they are middle-aged, married and have more children. This is implicitly also related to a higher number of persons in the household. In our sample, at the time of the windfall gain, 65% of individuals were married, and they were on average 43 years old, had 0.8 children, and lived with 3 persons in the household.

The final sample is restricted to non-missing values of these control variables, windfall information, and the outcomes. We include individuals aged 18 to 64 who are not self-employed or unemployed.[92] Furthermore, by restricting the sample to individuals who at some point receive a windfall gain and after removing missing values for the dependent and independent variables described above, we obtain an estimation sample of 33,094 observations, including 3,888 windfall recipients.

In our estimation strategy, we need comparable treatment and control groups in terms of observable characteristics. To address this point, in Table 5.1, we show the differences in means between the treated cohorts (2000-2019) and the control group (cohort 2020) *before* the treatment, i.e., in l = *-1*. The results indicate no difference between treatment and control groups in the incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}). Furthermore, the findings indicate no differences in all of the potential confounders (see Section 5.3.4). This mitigates concerns that treatment and control groups in our sample are systematically different. We will further evaluate the role of control variables in Section 5.4.1.

5.4 Results

Table 5.2 displays the average treatment effects of windfall gains on educational mismatch. Columns (1)–(4) provide regular TWFE estimations, while columns (5)–(8) show the main estimator based on Sun and Abraham (2021). In addition, we distinguish between the incidence of overeducation, OE_{ito} (columns (1), (2), (5) and (6)), and the incidence of undereducation, UE_{ito} (columns (3), (4), (7) and (8)). For each specification and outcome variable, we display a specification with and without control variables.

The results indicate noteworthy differences between the two estimation strategies, aligning with potential problems of standard TWFE estimations as explained in Section 5.3.3. However, irrespective of the specification, we find

[92]We also use only individuals with non-missing observations for job-related characteristics, including the industry, sector, income and working hours.

Table 5.1: Differences in means between cohorts

	(1)	(2)	(3)	(4)	(5)	(6)
	Early Cohorts (2000-2019)		Last Cohort (2020)			
	Mean	Sd	Mean	Sd	Diff.	P-value
OE_{ito}	0.17	0.38	0.18	0.39	0.01	0.67
UE_{ito}	0.10	0.30	0.09	0.28	-0.01	0.49
Log(# of persons in household)	1.00	0.46	0.99	0.50	-0.01	0.72
Age	42.65	10.94	43.53	11.67	0.88	0.26
Children	0.82	1.06	0.82	1.10	0.00	0.97
Married	0.64	0.48	0.59	0.49	-0.05	0.14
Num. obs.	2508		238			

Note: The table displays the differences in means (Column (5)) between individuals in cohorts 2000-2019 (Columns (1)-(2)) and the last treated cohort (Columns (3)-(4)). The significance levels are indicated by the p-values in column 6. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1.

Table 5.2: Effect of windfalls on educational mismatch

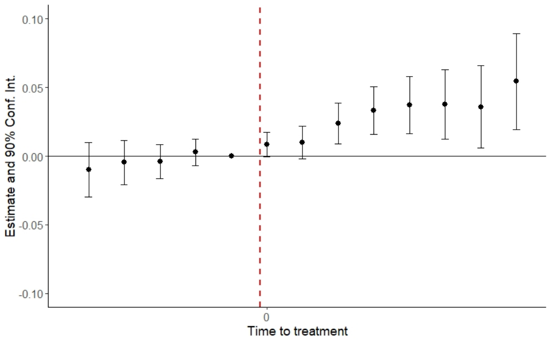
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	TWFE Estimator				Sun and Abraham (2021)			
Dependent variables:	OE_{ito}		UE_{ito}		OE_{ito}		UE_{ito}	
Windfall gain	0.012**	0.012**	0.006	0.007	0.031***	0.031***	0.008	0.009
	(0.005)	(0.005)	(0.004)	(0.004)	(0.011)	(0.011)	(0.008)	(0.008)
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	33094	33094	33094	33094	33093	33093	33093	33093
R^2	0.732	0.732	0.667	0.667	0.736	0.736	0.672	0.673
Within R^2	0.0004	0.0006	0.0001	0.0007	0.015	0.015	0.015	0.020

Note: The results display the effect of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on standard TWFE (Columns (1) to (4)) and the Sun and Abraham (2021) dynamic estimator (Columns (5) to (8)). Using the Sun and Abraham (2021) dynamic estimator, the estimates correspond to ATTs. The model subsequently adds the control variables mentioned in Section 5.3.4. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

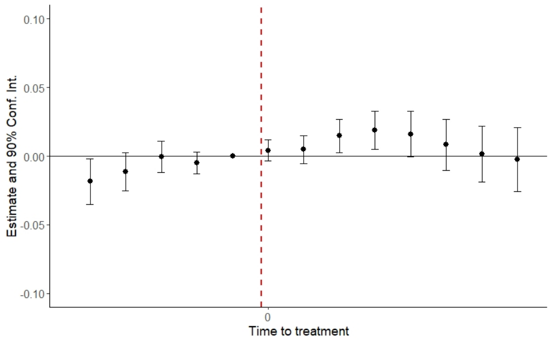
a significant positive effect of windfall gains on overeducation, OE_{ito}. In the main specification (column 6), the estimated treatment effect is 0.031. Hence, with respect to the pre-treatment mean of 0.17, the likelihood of being overeducated increases by approximately 18% after the windfall gain. In contrast to our theoretical expectations, windfall gains appear to have no effect on undereducation, UE_{ito}. Furthermore, including additional control variables (see Section 5.3.4) does not change the results. This limits concerns that our estimation might be driven by observables other than the windfall receipt.

Figure 5.1 displays the dynamic treatment effects based on Sun and Abraham (2021). The results for overeducation, OE_{ito}, (Figure 5.1a) show that the likelihood of overeducation increases in the first five years (*l = 0 to l = 4*) after the windfall gain and stabilises thereafter. Importantly, the graph shows no violation of the parallel trends or the anticipation assumption, as the coefficients in the pre-treatment period do not differ statistically from the reference period (*l = -1*). The results for undereducation, UE_{ito}, (Figure 5.1b) also indicate positive

Figure 5.1: Effect of windfalls on educational mismatch - Main event studies



(a) OE_{ito}



(b) UE_{ito}

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure 5.1a) and undereducation, UE_{ito}, (Figure 5.1b) for each relative treatment period. The sample includes the survey years between 2000 and 2020.

effects in two post-periods. However, it appears that pre-trends are in play, which contradicts the parallel trends assumption and may suggest anticipation effects. Besides that, the estimates are quantitatively smaller compared to those for overeducation and converge back to zero after *l=5*.

5.4.1 Robustness

To provide robustness to our main findings, we conduct a number of tests.[93] First, to infer potential violations of parallel trends, we employ alternative specifications in Figure D.2. To do so, we again use estimations based on Sun and Abraham (2021), but extend the control group to the last five cohorts of the sample in Figure D.2a and Figure D.2b. With that, the size of the control group increases, which mitigates concerns that initial treatment effects are due to specific characteristics of individuals treated in 2020. We also use the never-treated individuals as a control group in Figure D.2c and Figure D.2d. The sample size thereby increases to 212,516 observations. The results align with a significant increase in overeducation, OE_{ito}, and no violation of the parallel trends

[93] As of space reasons, we include figures displaying the dynamic estimations for all robustness checks exclusively if the number of observations stays roughly unchanged. For those, the ATT is presented in the caption of the respective figure. For robustness checks that alter the sample size more severely, we present tables and figures jointly.

assumption, while for undereducation, UE_{ito}, the assumption is consistently violated.[94]

Second, to account for the possibility that people were already receiving the treatment during period *l = -1*, we add *l = -2* as a reference period in our primary estimation. However, the results in Figure D.3 do not contradict our earlier findings. Similarly, when adopting a more balanced sample by confining the observations to individuals that were observed before and after the treatment, we find similar results as in the main specification (see Figure D.4).

Third, we make adjustments in the definition of the outcome variables. So far, the outcome for overeducation (undereducation) includes both matched and undereducated (overeducated) individuals in the reference category (classified as zero). While this is likely a reasonable specification, assuming that treatment effects are linear, moving from a state of undereducation to overeducation is rather extreme and less likely. Hence, in an alternative specification, we only compare the educational mismatch state with being matched (classified as zero). The findings in Figure D.5 are consistent with our main findings. Furthermore, in Figure D.6, we show the main results when using alternative outcome measures, namely a job analyst measure (Rumberger, 1981a, Figure D.6a and Figure D.6b) and indirect self-assessment (Duncan and Hoffman, 1981, Figure D.6c and Figure D.6d). While none of the average treatment effects are significant, the findings on overeducation defined by the job analyst measure (Figure D.6a) are significant at the 11%-level with a positive treatment effect. Still, results using both alternative methods need to be treated with caution, as both analyses reveal considerably larger pre-trends, even in the estimations using overeducation as the outcome.

So far, we have only looked at binary outcomes, hence the effect of the windfall on the likelihood of being over- or undereducated. However, the intensive margin may also be affected. To account for this, we use the years of educational mismatch, years of overeducation, and years of undereducation as outcome variables. The results in Figure D.7 show no intensive margin (or linear) effect, as the coefficients remain insignificant in these specifications. To elaborate this further, we estimate linear probability models on the likelihood of being overeducated by more than 0.1 (0.3, 0.5, 0.7, 0.9) years in Figure D.8. These reveal

[94] In other specifications, we use individuals that had not yet received a windfall gain as a control group by employing an estimator based on Gardner (2022) and Callaway and Sant'Anna (2021) that allows for staggered treatment. Specifically, Gardner (2022)'s approach employs a two-stage difference-in-difference estimator, and Callaway and Sant'Anna (2021) uses inverse probability weighting. The findings can be obtained upon request and are qualitatively in line with our main results.

that the treatment effects on overeducation turn insignificant for more than 0.5 years of overeducation, barely missing significance (p-value = 0.11). Hence, these results suggest that while a windfall receipt increases the likelihood of selecting overeducation at the extensive margin, the intensive margin is less severely affected. For undereducation (see Figure D.9), the results indicate significant and positive effects up to 0.3 years of undereducation as well. However, again, the pre-period coefficients violate the parallel trends assumption.

Fifth, the aggregate windfall measure is built on three components: Lottery winnings, inheritances, and gifts. Arguably, a specification with an exogenous income shock captured by lottery winnings would be the most reliable, as inheritances and gifts may be anticipated by individuals. However, in the SOEP data, lottery winnings constitute only a small fraction of windfall gains, and they are relatively small.[95] In the main estimation sample, the average lottery winner received EUR 19,000, compared to EUR 35,800 for gifts and EUR 57,700 for inheritances. When dividing the sample by the type of windfall, the results in Table D.1 show positive treatment effects on the probability of overeducation for all the windfall types. However, we observe sufficient significance levels only for inheritances. Despite the quantitatively larger coefficient for lottery winnings, the estimate only reaches a p-value of 0.2, likely due to the small sample size. In the dynamic estimations (see Figure D.10), none of the subsample specifications with overeducation, OE_{ito}, as an outcome violates the parallel trends assumption. This mitigates concerns that these types of windfalls are anticipated.

Still, as the statistical power of the estimations is limited when considering lottery winnings exclusively, we employ an alternative approach to control for anticipation effects. Besides information on the windfall gain itself, the SOEP provides information on whether an inheritance or gift is expected. Precisely, respondents are asked: "What do you think, are you going to inherit something or receive a gift of substantial value (again) in the future?". Table D.2 splits the sample according to this question into those who expected no gift or inheritance and those who expected one or did not know. While the estimate is positive and quantitatively comparable to the base specification (see columns (5) and (6) of Table 5.2), the ATT of windfall gains on the likelihood of overeducation reaches statistical significance only in the sample of individuals who did not expect the windfall. While not a perfect measure, this finding further limits concerns regarding potential anticipation effects. Moreover, considering the dynamic analysis in Figure D.11, the parallel trends assumption is not validated in either subsample. For undereducation, we find no significant treatment effects and

[95] We observe only approximately 400 winners and 1,565 observations.

we observe significant pre-trends in at least one of the pre-treatment periods (Figure D.12).

Last, when accounting for additional control variables including industry, occupation, working in the public sector and education level, the main results do not change (see Figure D.13). In sum, the reported evidence aligns with the formalised expectations that windfall gains increase the likelihood of overeducation, while no support for the hypothesis regarding undereducation is found.

5.4.2 Role of occupational choice and working conditions

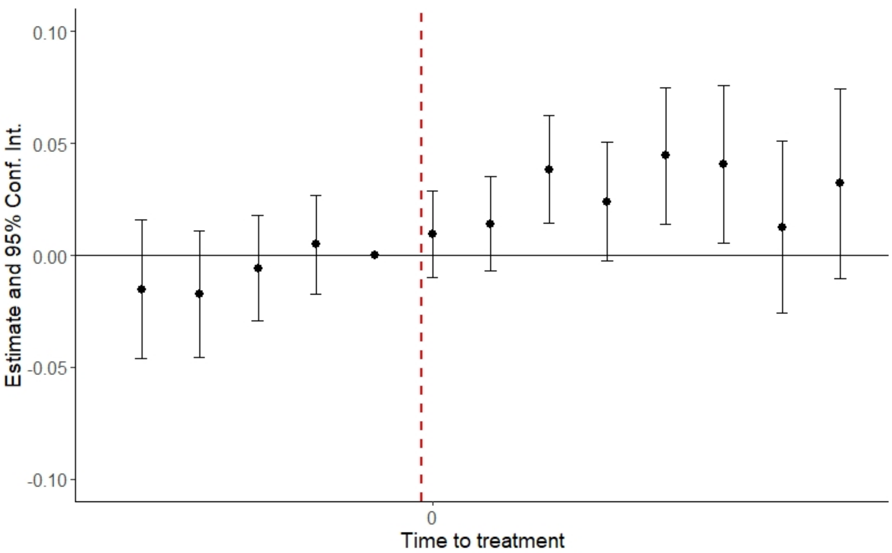
In the formalisation (Section 5.2), we argued that individuals make occupational choices that vary in the degree of educational (mis)match j amplified by the occurrence and size of a windfall gain V. Our main results demonstrate an increased likelihood of overeducation in the aftermath of a windfall, suggesting a link between windfall gains and the probability of either switching occupations or increasing the level of education.[96] In Figure 5.2, we investigate the impact of windfall gains on the likelihood of changing occupations assessed as a change at the three-digit ISCO level compared to the previous interview.

The results show a significant effect on the likelihood of changing occupations, amounting to an increase of 2.5 percentage points after the windfall gain. Given a pre-treatment mean of 0.177, the probability of switching occupations increases by approximately 14%. Particularly, in the dynamic analysis, the likelihood of changing the occupation increases two to four years after the windfall gain. The insignificant and close to zero treatment effects in the first two periods align with the supposition that changing occupations may take a certain amount of time (cf., Golosov et al., 2024).

Still, changing occupations is not sufficient to induce overeducation *per se*. The new occupation needs to differ in the job requirements and thus the required education and skills. Therefore, we further analyse occupational transitions by job type and skill level to evaluate the transition in more detail. Table 5.3 reports the treatment effects on the probability of certain job transitions depending on skill levels. We distinguish high-skill and low-skill occupations, where high-skill occupations cover ISCO-88 major groups 1, 2, 3, 6, and 7, while low-skill occupations capture the remaining ISCO-88 major groups (see Eurofund, 2024).

[96]Results for the effect of windfall gains on the likelihood of increasing the level of education are insignificant and can be obtained upon request.

Figure 5.2: Effect of windfalls on changing occupations



Note: Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on changing occupations (classified on the three-digit ISCO88 level) for each relative treatment period. The sample includes the survey years between 2000 and 2020.

Table 5.3: Windfall gains and job transitions by skill level

	(1)	(2)	(3)	(4)
Dependent variables: Skill transition	High to low	High to high	Low to high	Low to low
Windfall gain	0.009*	0.019**	0.002	0.002
	(0.005)	(0.009)	(0.005)	(0.005)
Num. obs.	27884	30187	27905	27833
R^2	0.202	0.299	0.197	0.269
Within R^2	0.018	0.017	0.017	0.018

Note: The results display the ATT of windfall gains on the probability of specific job transitions. For the outcomes, we distinguish between switches from 1) high to low-skill occupations, 2) high to high-skill occupations, 3) low to high-skill, and low to low-skill occupations. The estimations are based on the Sun and Abraham (2021) dynamic estimator. The sample includes the years between 2000 and 2020. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We define four distinct dummy outcomes capturing whether, compared to the last observed interview, the individual switched from high-skill to low-skill in column (1), within high-skill in column (2), low-skill to high-skill in column (3), and within low-skill occupations in column (4). In line with the base results of an increase in overeducation after a windfall gain, we observe a significant increase in the likelihood of switching from a high-skill to a low-skill occupation. Moreover, the likelihood of switching within the high-skill occupations is positive and significant, too. In contrast and aligning with the non-finding for the likelihood of undereducation, windfall gains do not affect the likelihood of switching from low- to high-skill occupations or within low-skill occupations.

Taken together, the above analysis provides suggestive evidence that the higher likelihood of overeducation induced by windfall gains is determined by individuals switching either from high-skill to low-skill or within high-skill occupations. Still, the propositions formulated in Section 5.2 predict that the transition between jobs of type j also depends on ξ and thus the ratio of the disutility from working in either of the potential jobs. In other words, we expect the working conditions experienced by the individuals to change with the windfall gain. To shed some light on these mechanisms, we look at different proxies for working conditions in Table 5.4.

Specifically, we regress weekly working hours, overtime, logged net earnings, and job satisfaction on the incidence of a windfall gain. In order to investigate whether these working conditions change among individuals switching out of high-skill occupations, we apply TWFE specifications and provide the aggregate treatment effects instead of dynamic estimators.[97] In this specification, the coefficient for windfall gains is the treatment effect for non-switchers and those who switch out of low-skill occupations, among whom we observe higher earnings but worse working conditions (working hours). For those who transition out of high-skill employment (arguably higher likelihood of overeducation), we find overall lower working hours (0.333 - 0.449 = -0.116) and lower overtime hours (0.068 - 0.187=-0.119), but similar earnings (0.027 - 0.024 = 0.003). Importantly, the results for working and overtime hours hold when including occupation fixed effects and align with the hypothesis formulated in our theoretical framework.[98] For job satisfaction, all the coefficients are insignificant but point in the same direction.[99]

In sum, the provided evidence suggests that the effect of windfall gains on the likelihood of overeducation arises due to individuals switching occupations in the aftermath. Particularly, we observe a positive linkage between windfall gains and switching from high- to low-skill occupations directly related to the definition of educational mismatch. Still, the quantitatively stronger results for the likelihood of

[97]Subsample restrictions based on specific job transitions would yield too little power in the framework of Sun and Abraham (2021).

[98]We also run regressions to see if the transitions from low-skill to high-skill employment or moving into undereducation are consistent with our hypothesis. However, contrary to our assumptions, the results show that such employment shifts are not significantly related to longer working hours and even lower wages, which contradicts our assumptions and might be one explanation for the overall null result. These findings can be obtained upon request.

[99]We also test a specification in which we compare transitions out of high-skill employment only to the non-switchers. Additionally, we test specifications, in which we keep occupation or education constant, in line with the assumptions on working hours ($h_{ue} > h_m > h_{oe}$) and wages ($w_{ue} > w_m > w_{oe}$) formulated in Section 5.2. All findings are qualitatively the same and can be obtained upon request.

Table 5.4: Windfall gains and other labour outcomes

	(1)	(2)	(3)	(4)
Dependent variables:	Working hours	Overtime	Earnings (log)	Job satisfaction
Windfall gain	0.333**	0.068	0.027***	-0.049
	(0.157)	(0.057)	(0.007)	(0.035)
Switch out of high	0.197	0.034	0.010	-0.004
	(0.203)	(0.089)	(0.009)	(0.051)
Windfall gain $\times$ Switch out of high	-0.449*	-0.187*	-0.024**	0.091
	(0.258)	(0.108)	(0.012)	(0.063)
Num. obs.	33094	32482	33094	32480
Within R^2	0.034	0.008	0.079	0.0004

Note: The results display the effect of windfall gains on working hours, overtime, logged net earnings and the level of job satisfaction. The dummy "switch out of high" indicates whether, compared to the previous interview, the individual switched from high to low-skill occupations or high to high-skill occupations. The estimations are based on TWFE with the same control variables as in the main specification. The sample includes the years between 2000 and 2020. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

switching within high-skill occupations allow for the supposition that the incentive to switch occupations through the external wealth increase does not suffice to induce fundamental changes in individual occupations.[100] Moreover, we find no evidence of an effect on labour earnings. Thus, while transitions from high-skill to low-skill occupations involve an increase in leisure time ($T - h_j$), as expected, (cf., Basiglio et al., 2023), these changes do not come at the cost of reduced earnings w_j.

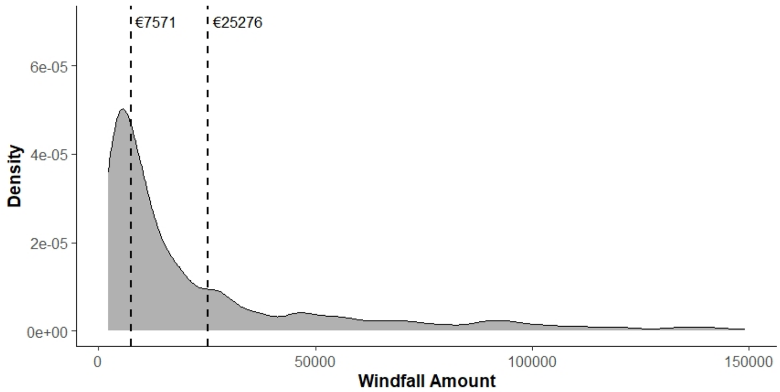
5.4.3 Role of the size of the windfall gain

As formalised in Section 5.2, the expectations regarding the effect of windfall gains on educational mismatch depend crucially on whether the windfall gain exceeds the critical value of V (see Equation 5.6). In line with the supposition that the impact of windfall gains may depend on their size, Sila and Sousa (2014) reported negative effects on employment probability and working hours among males only for windfalls exceeding EUR 50,000. Moreover, some studies investigating the impact of windfalls on outcomes related to labour supply or job choice indicate non-linear relationships (e.g., Malo and Sciulli, 2021). Consistent with the Carnegie effect (see, e.g., Holtz-Eakin et al., 1993), this reported inverse u-shape may be related to the increasing likelihood of leaving the labour market if windfalls get very large (Georgellis et al., 2005; Taylor, 2001).

The SOEP data not only allows us to evaluate the receipt of a windfall gain, but also to differentiate its size. We categorise the measure into three levels, $S_{i,t}^{SMALL}$, $S_{i,t}^{MEDIUM}$, $S_{i,t}^{LARGE}$ and choose a data-driven approach defining the categories

[100] This is also supported by supplementary findings showing that individuals are not more likely to switch industries in response to a windfall gain. The results can be obtained upon request.

Figure 5.3: Distribution of windfall gains size ($S_{i,t}$)



Note: The figure shows the distribution of windfall gains in Germany. Windfall gains include lottery winnings, gifts and inheritances. The sample includes the survey years between 2000 and 2020 and is restricted to observations where a windfall occurred. The dashed lines indicate the cutoff points for the definition of $S_{i,t}^{SMALL}$, $S_{i,t}^{MEDIUM}$ and $S_{i,t}^{LARGE}$. Furthermore, for better visibility, the figure is limited to $S_{i,t}^{r} <$ EUR$150,000$.

based on the tertiles of $S_{i,t}$. Figure 5.3 shows the distribution of the deflated windfall gains $S_{i,t}$ with the dashed lines indicating the cutoff points for $S_{i,t}^{SMALL}$, $S_{i,t}^{MEDIUM}$ and $S_{i,t}^{LARGE}$. The distribution is left-skewed with a long right tail: In our specification $S_{i,t}^{SMALL}$ ranges from EUR 2,204[101] to EUR 7,571, $S_{i,t}^{MEDIUM}$ from EUR 7,583 to EUR 25,276 and $S_{i,t}^{LARGE}$ above that with a maximum of EUR 2,256,299. The sample median of $S_{i,t}$ denotes at EUR 12,036.

In Table 5.5, we show results for subsamples depending on $S_{i,t}^{SMALL}$, $S_{i,t}^{MEDIUM}$ and $S_{i,t}^{LARGE}$. As the samples shrink and each subsample only contains around one-third of the initial observations, the significance levels decline. However, for overeducation, OE_{ito}, we observe a positive and significant effect for medium and large payments, while small payments show no significant increase. Conversely, for undereducation, UE_{ito}, we find no significant impact of windfalls of any size.

The dynamic effects displayed in Figure D.14a do not show a significant increase in overeducation, OE_{ito}, in response to small windfalls in any of the post periods. Large and medium-sized payments (Figure D.14c and Figure D.14e), on the other hand, significantly increase the likelihood of being overeducated by up to 6pp (8 years after the treatment). This amounts to an increase of approximately a third compared to the share of overeducation in the 2020-cohort (see Table 5.1). Furthermore, the figures for overeducation (Figure D.14a, Figure D.14c, and Figure D.14e) consistently show no pre-trends. For undereducation (Figure D.14b, Figure D.14d, and Figure D.14f), the positive treatment effect for small windfalls displays significant pre-trends, which makes it less feasible to interpret.

[101]The sample is restricted to nominal values above EUR 2,500, which are deflated using the CPI of 2010.

Table 5.5: Effect of windfalls on educational mismatch - Size of windfalls

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:		OE_{ito}			UE_{ito}	
Amount	$S^{SMALL}_{i,t}$	$S^{MEDIUM}_{i,t}$	$S^{LARGE}_{i,t}$	$S^{SMALL}_{i,t}$	$S^{MEDIUM}_{i,t}$	$S^{LARGE}_{i,t}$
Windfall gain	0.017	0.046**	0.036*	0.021	0.007	-0.006
	(0.013)	(0.018)	(0.022)	(0.015)	(0.015)	(0.012)
Num. obs.	10900	11161	11032	10900	11161	11032
R^2	0.745	0.774	0.716	0.656	0.709	0.685
Within R^2	0.047	0.053	0.045	0.047	0.048	0.047

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between the size of the payment. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To account for the possibility that raw windfall amounts do not accurately reflect the financial benefit for the household, in an alternate specification, we normalise the raw amounts by the number of persons in the household at the time of the windfall gain. The results in Table D.3 correspond with the results reported in this section. Small payments have no significant effect, whereas medium and high payments increase the likelihood of overeducation. Again, undereducation appears to be positively impacted by small payments, but pre-trends diminish these conclusions.[102]

Overall, the findings align with the notion that the size of the windfall gain V increases the effect on overeducation, while the results for undereducation are inconclusive.

5.4.4 Heterogeneity

In this section, we will discuss various sources of heterogeneity in the main treatment effects by analysing whether the effects depend on financial and individual characteristics observed in the last period prior to the treatment. As our initial findings did not show a significant treatment effect on the probability of undereducation and the pre-treatment periods violated the parallel trends assumption, the analysis will mostly focus on overeducation as an outcome. To obtain subsamples based on non-categorical variables, we divide the sample using a median split. However, we also provide results using tertile splits for the variables in Table D.6.

[102]Furthermore, in an alternative specification, we normalise the size of the windfall by the household net income. In this specification, for overeducation, we observe increasing treatment effects with size, but even the smallest windfall gains show a significant (but smaller) effect with no violation of the parallel trends assumption. Additionally, in specifications restricted to singles, where windfalls should have a more substantial financial impact, we observe larger treatment effects (0.07) that are significant at the 11%-level. The results can be obtained upon request.

Financial situation

In line with the formalisation, the previous section shows that the effect of windfall gains comes into play only after a certain threshold. But, the windfall gain itself may not be the only source of non-labour income or wealth that determines whether individuals can switch occupations afterwards or not. We hypothesised that windfalls increase the likelihood of moving from undereducated to matched (overeducated) positions by reducing the marginal utility of the wage premium, thus amplifying the disutility of higher-paid roles. This mechanism could be particularly relevant for wealthier individuals, as pre-existing financial security might facilitate transitions into lower-disutility jobs when wealth increases further. If V in Equation 5.6 comprised windfalls v but also a stock of previously owned wealth v^*, the utility advantage of matched over undereducated positions arises at lower v amplifying the effect for wealthier individuals. Consequently, the impact of a given windfall could be larger for wealthier individuals. To account for different facets of financial security indicators, we derive measures of household income, household income satisfaction, debt and home ownership from the SOEP.[103] The prevalence of debt can be linked to a higher financial burden, potentially lowering the possibility of moving toward overeducation where one likely faces a wage penalty (cf., Duncan and Hoffman, 1981). Homeownership, on the other hand, may also proxy an individual's wealth, but it can also indicate a higher pre-treatment debt burden.

The results displayed in Table D.4 do not align with expectations when using a median split of household income (Panel A). The coefficients are not statistically distinct from each other for lower and higher household incomes. However, when looking at the tertile split in Table D.6 (Panel B), we find stronger and significant effects only at the top tertile of the household income distribution in line with our expectations. Furthermore, the findings for household income satisfaction report a positive effect of windfall gains on the likelihood of overeducation for individuals reporting higher household income satisfaction. Moreover, Panel C reveals a significant association only for employees without pre-treatment debt, in line with our expectations. However, the coefficients do not differ statistically from employees who are in debt. Finally, we do not observe differences in the treatment effects for individuals with and without their own properties (Panel D). In sum, the heterogeneity concerning pre-windfall wealth is less pronounced than that concerning the size of the windfall itself.

[103]Household income is the monthly income of all members of the household deflated using the CPI with base year 2010.

Individual characteristics

Besides this, other sources of heterogeneity may arise in the context of windfall gains and their impact on the likelihood of overeducation. Several studies, for example, report heterogeneous effects of income shocks for females and males, while others report no differences by sex (Belloc et al., 2025; Doorley and Pestel, 2020; Flèche et al., 2021; Georgellis et al., 2005; Golosov et al., 2024; Henley, 2004; Malo and Sciulli, 2021). Additionally, older employees may be less responsive to unexpected financial gains as they likely have lower labour mobility and move or switch industries or occupations less often (see early evidence from Gallaway, 1969). Moreover, the impact of a windfall gain on individuals' overall wealth may be more distinctive among younger compared to older individuals. However, since age is positively associated with the likelihood of receiving inheritances (Carman, 2013; Crawford and Hood, 2016; Kindermann et al., 2020), it is also possible that older individuals are more likely to receive the treatment itself.[104] Finally, a windfall gain may alleviate financial stress, and married employees or those with children may be more likely to use that relief to prioritise family over career orientation. This could result in stronger treatment effects on overeducation for those individuals. In line with this supposition, Sila and Sousa (2014) report larger reactions to windfall gains on working hours among married individuals and individuals with children below age 6. Workers with children in the household, on the other hand, may also face a higher initial financial weight, which may lead to a smaller response to windfalls (for related evidence, see, e.g., Doorley and Pestel, 2020).

In line with Golosov et al. (2024) and Flèche et al. (2021), the findings in Table D.5 indicate no differences in the response to windfall gains by sex as coefficients and confidence bands overlap. However, they confirm the first suspicion regarding age, i.e., that the treatment effects are solely driven by younger workers aged 18-43. When looking at age tertiles in Table D.6, this is even more obvious as only the first tertile (ages 18-38) displays a significant treatment effect. The findings show no heterogeneity when comparing workers with and without children (see Table D.5, Panel C). Furthermore, stronger results for married employees (Panel D) align with the notion of prioritising family over work and results presented by Sila and Sousa (2014). However, the confidence

[104] Elinder et al. (2012) argue that older individuals should report larger earnings responses to windfall gains, as the expected period in which they profit from their gain is shorter. Joulfaian and Wilhelm (1994) report comparable evidence for lottery wins, where individuals aged 55 to 65 respond more to the windfall gain in terms of earnings than the younger ones.

bands slightly overlap, so we refrain from drawing strong conclusions based on that heterogeneity.

In short, we find the positive effect of windfall gains on the likelihood of overeducation to be mostly driven by younger individuals.

5.5 Conclusion

Although windfall gains, comprising lottery wins, inheritances, and gifts, may have a huge impact on societal income as well as private wealth, only a few studies examine their consequences for labour market outcomes. While the existing studies provide insights into the effects of windfall gains on employment, labour income, and working hours (Doorley and Pestel, 2020; Golosov et al., 2024; Holtz-Eakin et al., 1993; Imbens et al., 2001; Joulfaian and Wilhelm, 1994), nothing is known about the impact on job match quality.

The present study fills that gap by providing conceptual as well as empirical insights on the impact of windfall gains on educational mismatch. More precisely, we suppose that the utility offered by undereducation exceeds that of a matched position if the utility surplus gained through the wage premium associated with undereducation exceeds the disutility from working in such a position compared to an adequate position. The opposite holds true when comparing overeducation to a matched position. An unexpected increase in wealth, e.g., through a windfall gain, diminishes the marginal utility of labour income in this comparative static such that non-monetary aspects, i.e., working hours or leisure time, gain relative importance. From this, we hypothesise that the likelihood of undereducation (overeducation) will decline (rise) after a sufficiently large windfall gain.

We test the derived suppositions empirically by employing German data from the Socio-Economic Panel for the years 2000 to 2020. Using TWFE estimators and event-study analyses based on Sun and Abraham (2021), we find evidence in line with our expectations as the likelihood of overeducation rises after a windfall gain. Disentangling these results further reveals that this rise in the likelihood of overeducation is driven by job switchers changing from high-skill to low-skill occupations or within high-skill occupations. The switches out of high-skill occupations are accompanied by a reduction in working and overtime hours, supporting the notion of a potential trade-off between job match quality and leisure time. As expected, an increase in the likelihood of overeducation can only be observed for medium or larger windfalls and is particularly pronounced among individuals with higher pre-treatment household income satisfaction. Finally, younger individuals tend to respond more strongly to the income shock. In contrast, we find no evidence for a decrease in the likelihood of undereducation after a windfall gain, and results are consistently affected by pre-trends.

Our findings contribute to the literature in several ways: We provide the first conceptual and empirical evidence on the link between windfall gains and educational mismatch. This extends our knowledge concerning the impact of windfall gains on labour-related outcomes (Doorley and Pestel, 2020; Georgellis et al., 2005; Golosov et al., 2024; Holtz-Eakin et al., 1993; Imbens et al., 2001; Joulfaian and Wilhelm, 1994; Lindh and Ohlsson, 1996; Sila and Sousa, 2014; Taylor, 2001). Moreover, we add to the literature on the antecedents of educational mismatch, which has in many parts focused on personal or job-related factors (e.g., Aleksynska and Tritah, 2013; Belfield, 2010; Dolton and Silles, 2008; McGoldrick and Robst, 1996; Robst, 1995b; Santiago-Vela and Mergener, 2022; Turmo-Garuz and Bartual-Figueras, 2019). Our findings indicate that preferences for job-education-match quality can shift in response to unforeseen circumstances, such as windfall gains. This dynamic, which has received limited attention in the existing literature, underscores the importance of considering external shocks in understanding the emergence of educational mismatch.

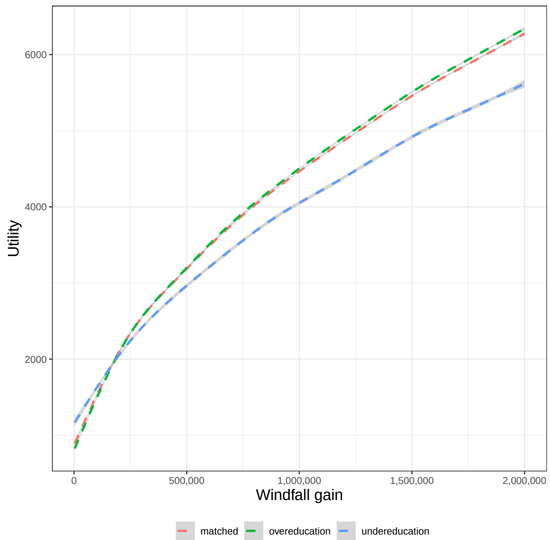
From a policy perspective, windfall gains, in the German case, particularly inheritances, can affect the labour market to a considerable extent. If individuals inherit a sum that reverses their choices regarding job match quality and shift, for example, to overeducation, this implies an inefficient distribution of a well-educated workforce. Besides an over-expenditure in education, such developments may, on the one hand, additionally foster the shortage of appropriately skilled labour that modern labour markets face (European Commission, 2023; OECD, 2022a). On the other hand, such mechanisms may lead to a crowding-out of middle or lower-skilled workers due to an oversupply of highly educated labour for jobs that do not require that level of education (Dolado et al., 2000; Lazear et al., 2018). Still, the results of this study hint at the possibility that individuals actively select overeducation under certain circumstances, which proposes that overeducation is not *per se* evaluated negatively at the individual level, in contrast to the perception in the majority of the existing literature (cf., Sloane et al., 1996). This paper also contributes to the continuing debate about inheritance taxes. According to our findings, increased tax allowances, or even the abolition of inheritance taxes, may enhance the risk of overeducation since individuals profit from increased post-tax earnings.

Finally, the following back-of-the-envelope calculation helps to illustrate the economic impact of these findings from an individual's perspective. A regression of net earnings (logarithm) on overeducation incidence, years of education, and socio-economic, industry, and occupation controls reveals a 5% wage penalty for overeducation (for similar evidence, see, e.g., Chu Ng, 2003). For the average worker, this translates to an annual income reduction of approximately EUR 1,084. Given a median windfall gain of EUR 12,036, this financial buffer could offset the wage penalty of overeducation for around 12 years. Since the average individual in the dataset is 43 years old, this suggests that by age 55, the entire windfall would be depleted by the cumulative wage penalty. As the legal retirement age varies between ages 65 and 67 in Germany, the shift induced by an unexpected increase in wealth actually ultimately leads to a reduction in overall lifetime wealth.

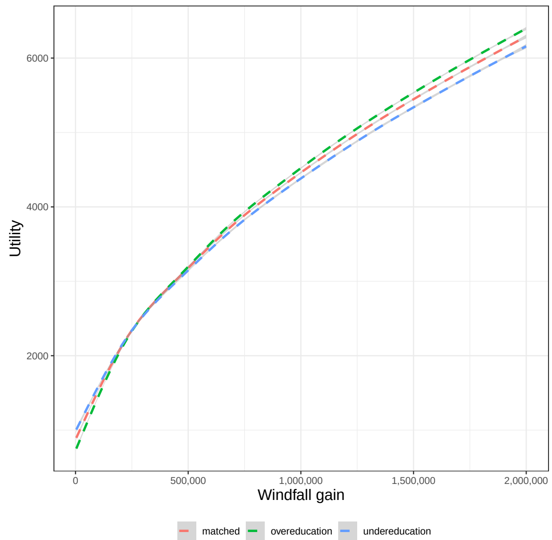
5.6 Appendix D

5.6.1 Simulations

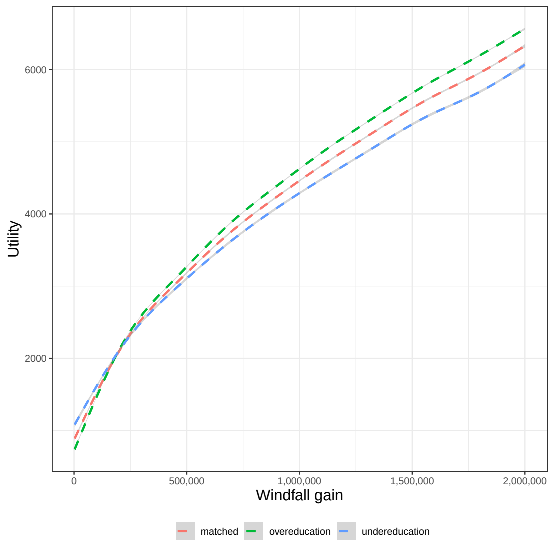
Figure D.1: Simulations



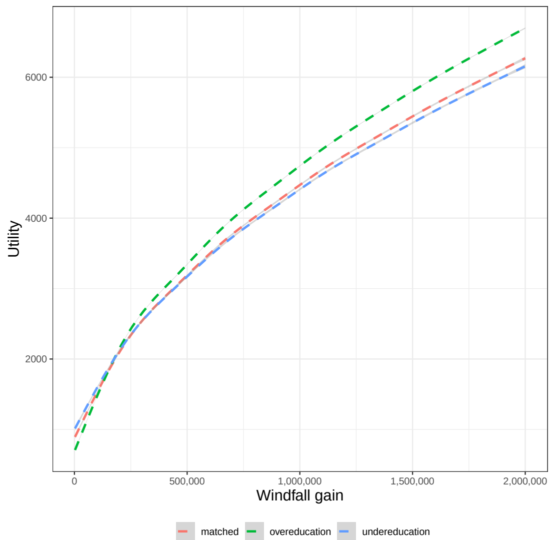
(a) First iteration



(b) Second iteration



(c) Third iteration



(d) Fourth iteration

Note: Figure D.1 displays four exemplary iterations simulating the utility function formalised in Section 5.2.

In order to simulate the utility function formalised in Section 5.2, we assume equal preferences for leisure and consumption, i.e., $a = b = 0.5$ for all iterations.[105] We use 20 iterations based on 10,000 individuals with 3 observations each. This

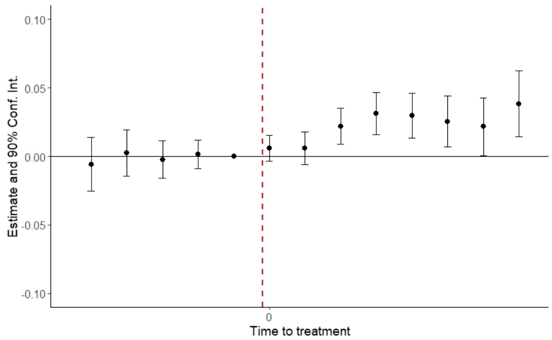
[105]Simulations with varying preferences between individuals can be obtained from the authors.

yields 30,000 observations in which we observe each individual thrice, once in overeducation, undereducation, and a matched position. To compare the impact of educational (mis)match on utility, we set base hours, wages, and windfalls as constant within individuals but allow them to vary between individuals. The base hours worked h_i are a random number between 1 and 8 hours, while the base wage w_i is a random number between 2,200 and 10,000. The windfall of each individual V_i is a random number between 2,500 and 2,000,000. We simulate a job in overeducation, undereducation, and a matched position for each individual. In overeducation, the base hours are reduced by a random proportion of the base hours such that $h_{oe} = h_i - h_i \times x$, where x randomly varies between 0.1 and 0.9. The hours worked in an undereducated position are defined as $h_{ue} = h_i + h_i \times x$. Similarly, the distinct wages are defined as $w_{oe} = w_i - w_i \times x$ and $w_{ue} = w_i - w_i \times x$. The total time available T is set to 24. The utility is estimated as defined in Formula 5.4. Figure D.1 shows smoothed conditional means relating V_i on the x-axis to U_i on the y-axis.[106]

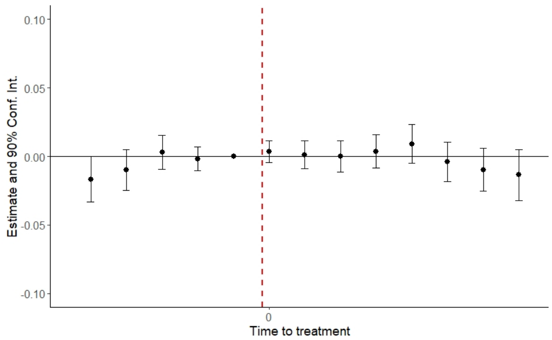
[106]Simulations are pursued with R 4.2.3. All remaining simulations and a version displaying the raw data points are available from the authors.

5.6.2 Robustness

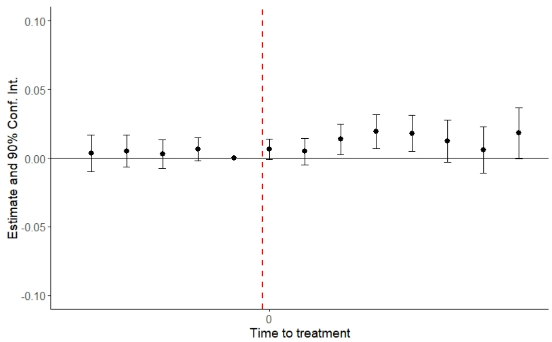
Figure D.2: Effect of windfalls on educational mismatch - Main event studies with adjusted control groups



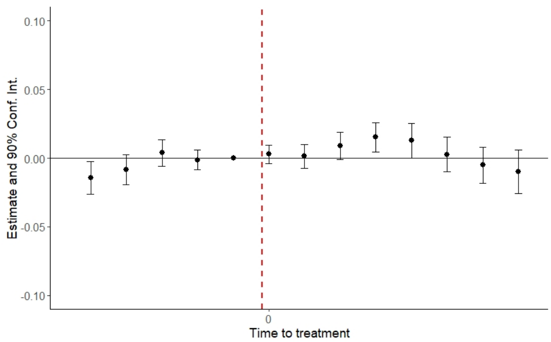
(a) Last five cohorts: $OE_{ito}(ATT : 0.026**)$



(b) Last five cohorts: $UE_{ito}(ATT : 0.000)$



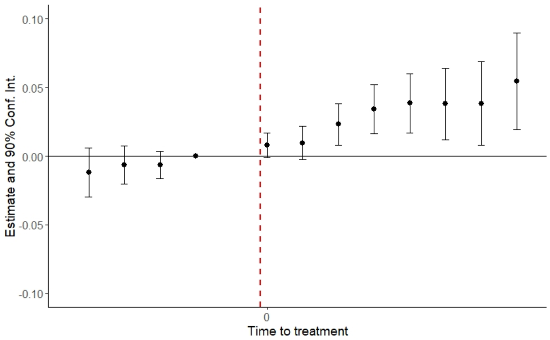
(c) Never-treated: $OE_{ito}(ATT : 0.013*)$



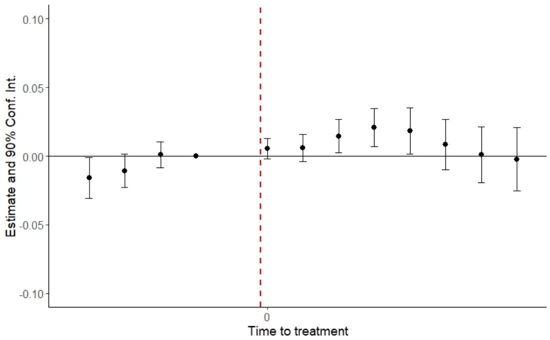
(d) Never-treated: $UE_{ito}(ATT : -0.004)$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.2a and Figure D.2c) and undereducation, UE_{ito}, (Figure D.2b and Figure D.2d) for each relative treatment period. We distinguish between specifications that are based on the last five cohorts (Figure D.2a and Figure D.2b) and never-treated individuals (Figure D.2c and Figure D.2d) as control group. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.3: Effect of windfalls on educational mismatch - Main event studies with adjusted reference category (-2 and -1)



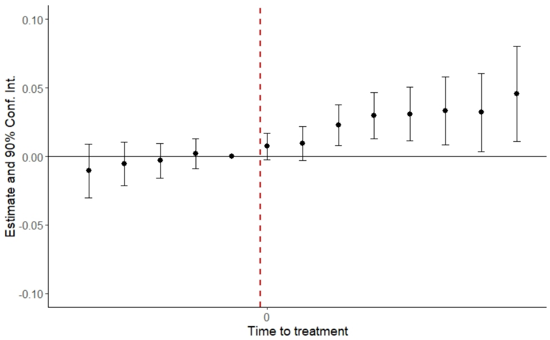
(a) $OE_{ito}(ATT: 0.032**)$



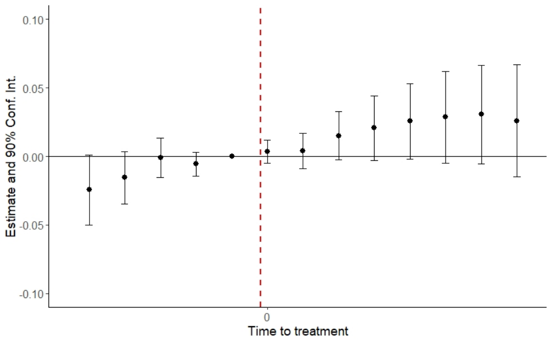
(b) $UE_{ito}(ATT: 0.010)$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.3a) and undereducation, UE_{ito}, (Figure D.3b) for each relative treatment period. In these specifications, we use $l = -2$ and -1 as the reference period. The sample includes the survey years between 2000 and 2020. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Figure D.4: Effect of windfalls on educational mismatch - Main event studies with more symmetric sample



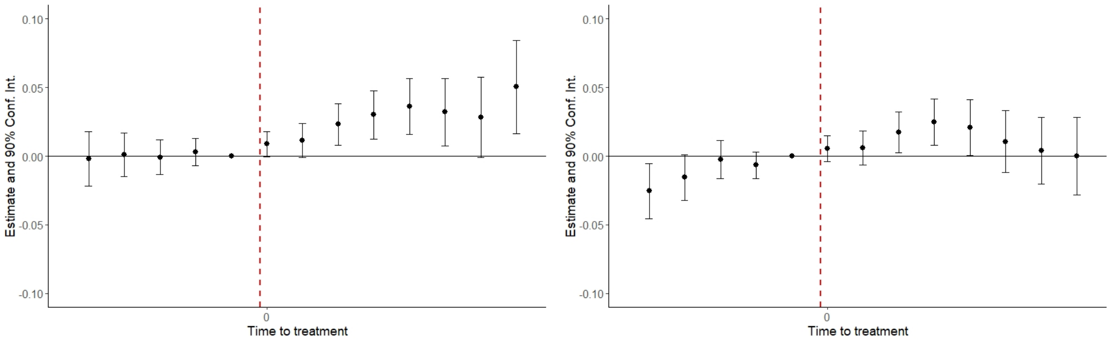
(a) $OE_{ito}(ATT: 0.028**)$



(b) $UE_{ito}(ATT: 0.021)$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.4a) and undereducation, UE_{ito}, (Figure D.4b) for each relative treatment period. The sample includes the survey years between 2000 and 2020 and is based on individuals observed before *and* after the treatment. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

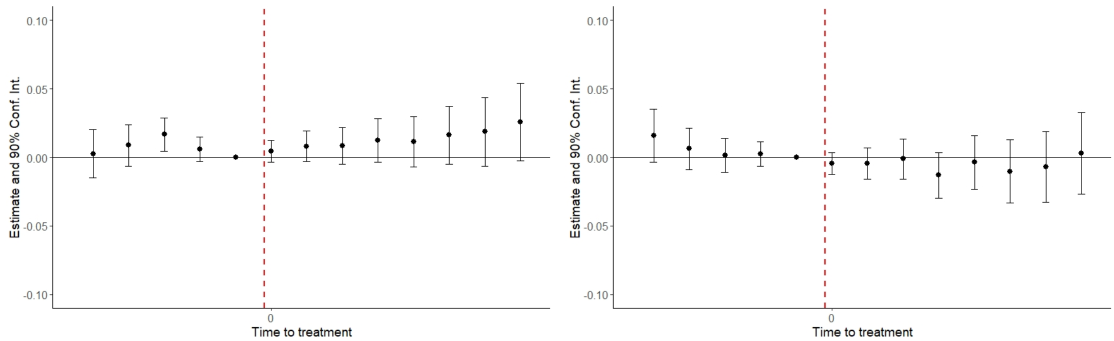
Figure D.5: Effect of windfalls on educational mismatch - Adjusted outcome specification



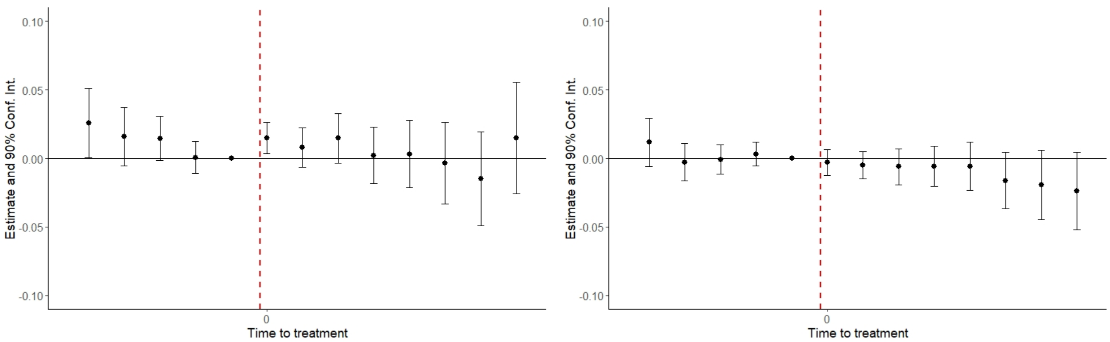
(a) $OE_{ito}(ATT : 0.029**)$ (b) $UE_{ito}(ATT : 0.011)$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.5a) and undereducation, UE_{ito}, (Figure D.5b) for each relative treatment period. Compared to the main specification, we only compare overeducation (undereducation) with being matched. The sample includes the survey years between 2000 and 2020. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Figure D.6: Effect of windfalls on educational mismatch - Main event studies with alternative outcome measures



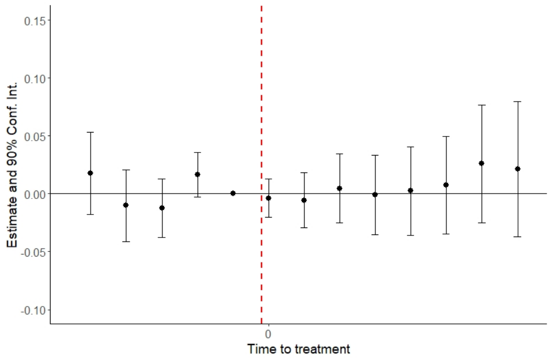
(a) JA: $OE_{ito}(ATT : 0.014, p - value : 0.11)$ (b) JA: $UE_{ito}(ATT : -0.005)$



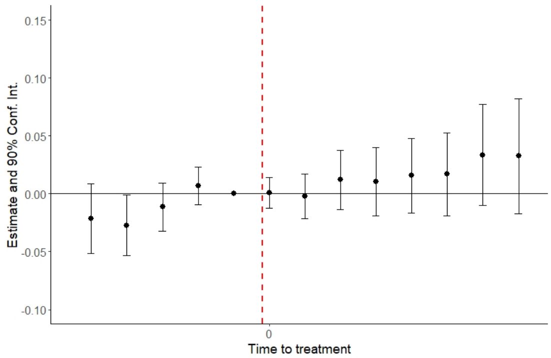
(c) ISA: $OE_{ito}(ATT : 0.010)$ (d) ISA: $UE_{ito}(ATT : -0.009)$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.6a and Figure D.6c) and undereducation, UE_{ito}, (Figure D.6b and Figure D.6d) for each relative treatment period. We distinguish between a job analyst measure (Figure D.6a and Figure D.6b) and indirect self assessment (Figure D.6c and Figure D.6d) for the educational mismatch variables. The sample includes the survey years between 2000 and 2020. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

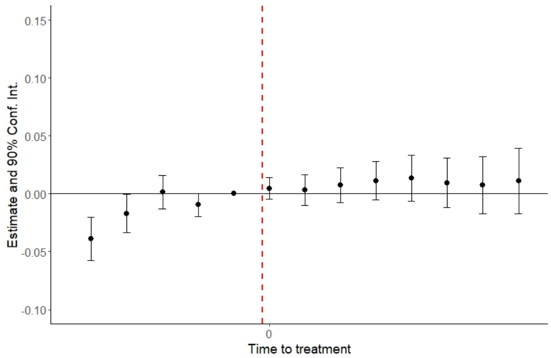
Figure D.7: Treatment effects on the years of mismatch - Dynamic estimates



(a) Effect on years of mismatch (ATT: 0.004)



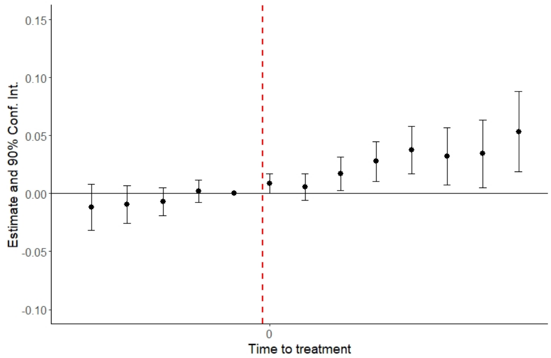
(b) Effect on years of OE_{ito} (ATT: 0.015)



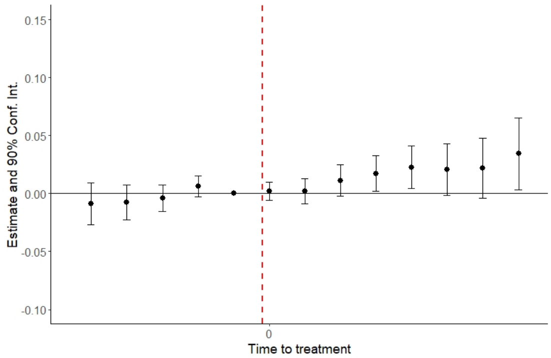
(c) Effect on years of UE_{ito} (ATT: 0.011)

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on years of mismatch (Figure D.7a), years of overeducation (Figure D.7b) and years of undereducation (Figure D.7c) for each relative treatment period. In Figure D.7b (Figure D.7c) positive years of undereducation (overeducation) are classified as zero. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

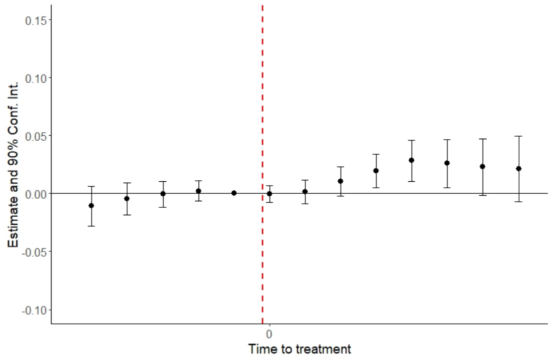
Figure D.8: Treatment effects on the years of overeducation - Linear probability models



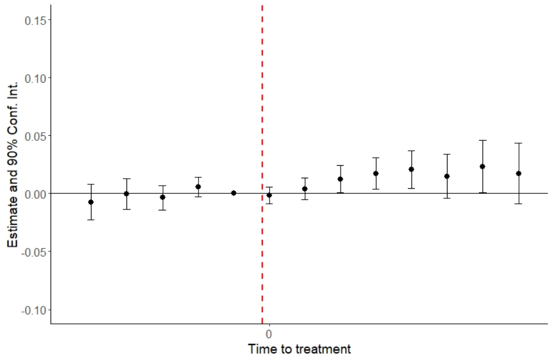
(a) Effect on years of $OE_{ito} > 0.1$ (ATT: 0.028***)



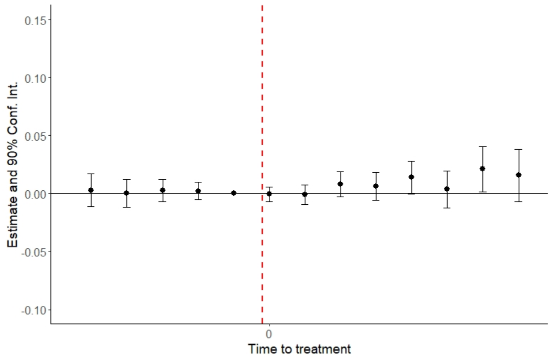
(b) Effect on years of $OE_{ito} > 0.3$ (ATT: 0.016*)



(c) Effect on years of $OE_{ito} > 0.5$ (ATT: 0.014)



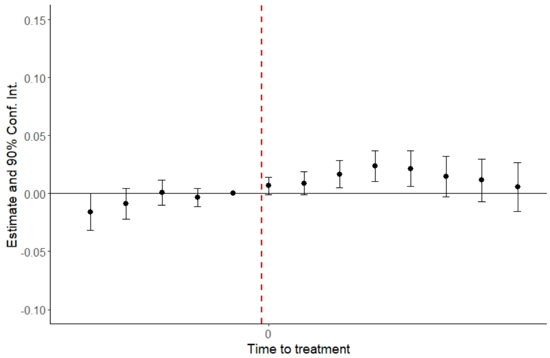
(d) Effect on years of $OE_{ito} > 0.7$ (ATT: 0.011)



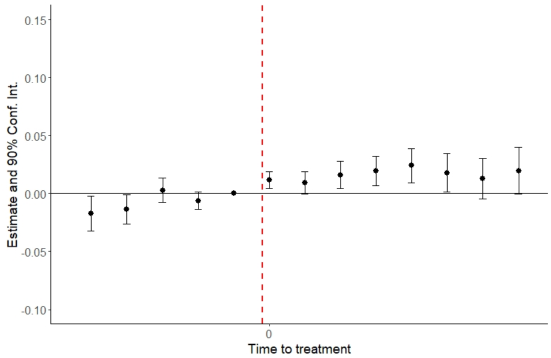
(e) Effect on years of $OE_{ito} > 0.9$ (ATT: 0.007)

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on the probability of being overeducated by at least x years for each relative treatment period. Positive years of undereducation are classified as zero. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

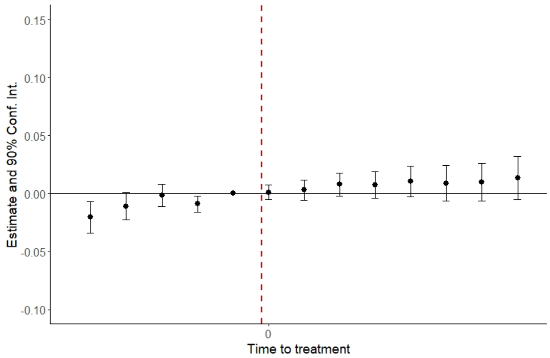
Figure D.9: Treatment effects on the years of undereducation - Linear probability models



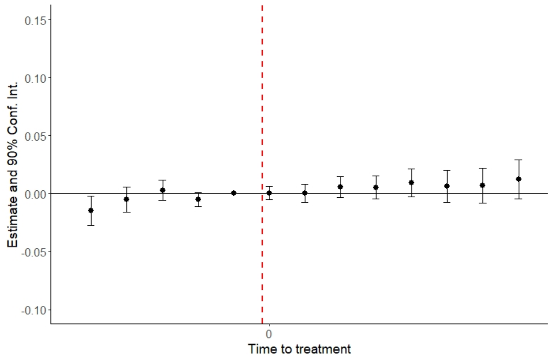
(a) Effect on years of $UE_{ito} > 0.1$ (ATT: 0.014**)



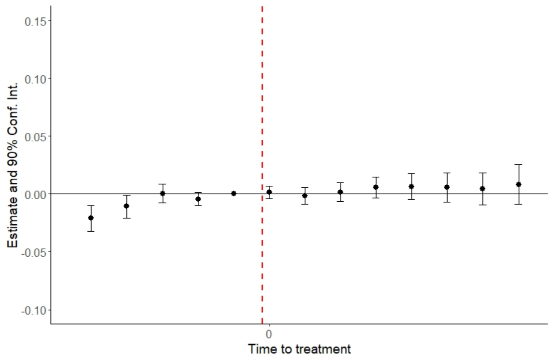
(b) Effect on years of $UE_{ito} > 0.3$ (ATT: 0.017**)



(c) Effect on years of $UE_{ito} > 0.5$ (ATT: 0.009)



(d) Effect on years of $UE_{ito} > 0.7$ (ATT: 0.006)



(e) Effect on years of $UE_{ito} > 0.9$ (ATT: 0.005)

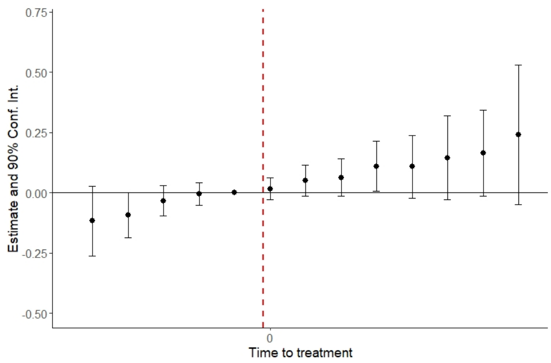
Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on the probability of being undereducated by at least x years for each relative treatment period. Positive years of overeducation are classified as zero. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.1: Effect of windfall gains on educational mismatch - Type of windfalls

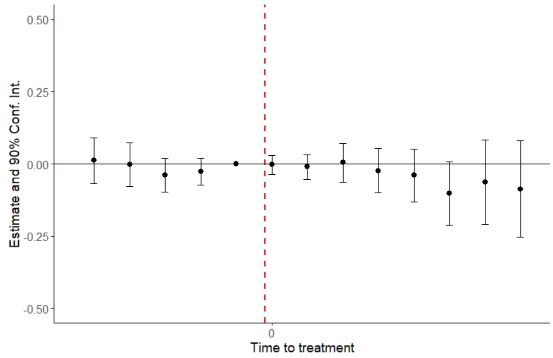
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:		OE_{ito}			UE_{ito}	
	Lottery	Inheritance	Gifts	Lottery	Inheritance	Gifts
Windfall gain	0.134	0.030**	0.024	-0.045	-0.001	0.023*
	(0.104)	(0.014)	(0.017)	(0.061)	(0.010)	(0.012)
Num. obs.	1565	17332	12792	1565	17332	12792
R^2	0.76075	0.73221	0.74702	0.75025	0.67751	0.68519
Within R^2	0.24056	0.02734	0.03703	0.27651	0.03317	0.04284

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between the type of the windfall. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

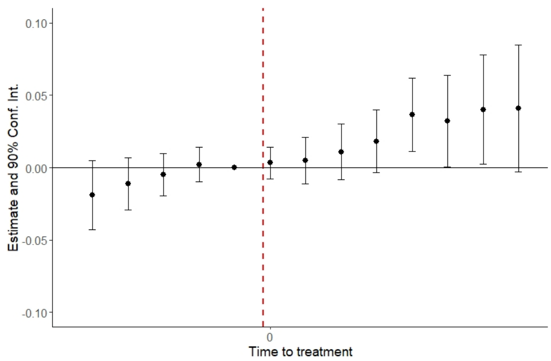
Figure D.10: Effect of windfalls on educational mismatch - Type of windfalls



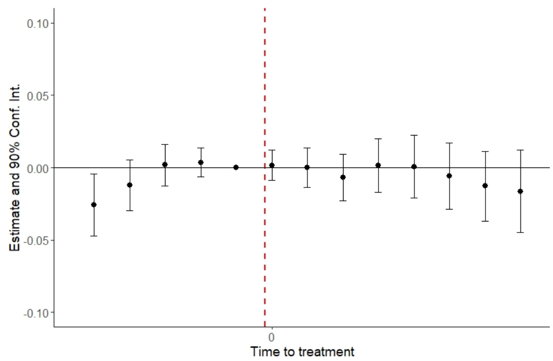
(a) Effect on OE_{ito} for lottery winnings



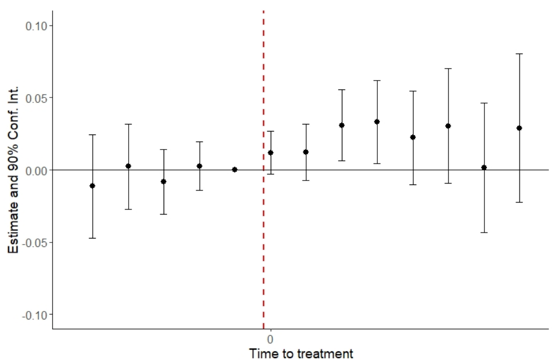
(b) Effect on UE_{ito} for lottery winnings



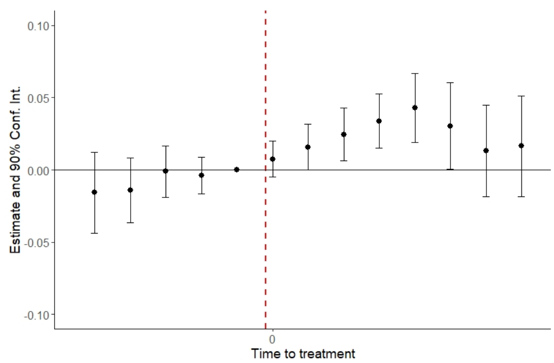
(c) Effect on OE_{ito} for inheritances



(d) Effect on UE_{ito} for inheritances



(e) Effect on OE_{ito} for gifts



(f) Effect on UE_{ito} for gifts

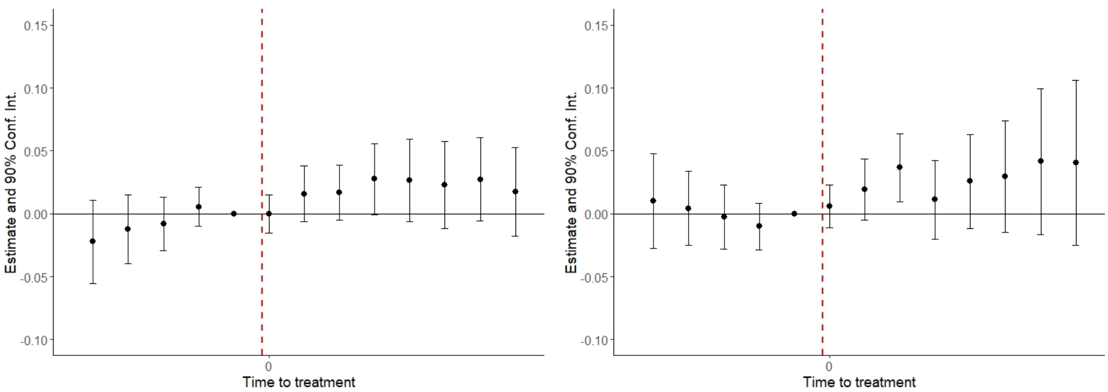
Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.10a, Figure D.10c and Figure D.10e) and undereducation, UE_{ito}, (Figure D.10b, Figure D.10d and Figure D.10f) for each relative treatment period. We distinguish between the type of windfall for each relative treatment period. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: Effect of windfall gains on educational mismatch - Anticipation

	(1)	(2)	(3)	(4)
Dependent variables:	OE_{ito}		UE_{ito}	
Inheritance/Gift in near future	Not expected	Expected/Do not know	Not expected	Expected/Do not know
Windfall gain	0.025*	0.035	0.018	0.008
	(0.014)	(0.024)	(0.018)	(0.013)
Num. obs.	6022	8441	6022	8441
R^2	0.805	0.751	0.682	0.667
Within R^2	0.088	0.050	0.088	0.054

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between individuals that expect and not expect an inheritance or gift in the near future. The question is asked in the 2001 survey. The sample includes the years between 2000 and 2020. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

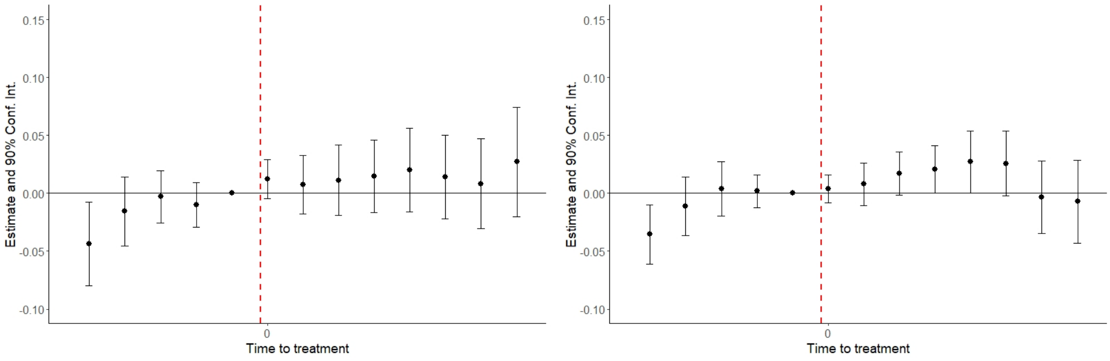
Figure D.11: Effect of windfalls on overeducation - Anticipation



(a) Unexpected (b) Expected + Do not know

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, for each relative treatment period. We distinguish between individuals that expect (Figure D.11b) and not expect (Figure D.11a) an inheritance or gift in the near future. The question is asked in the 2001 survey. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

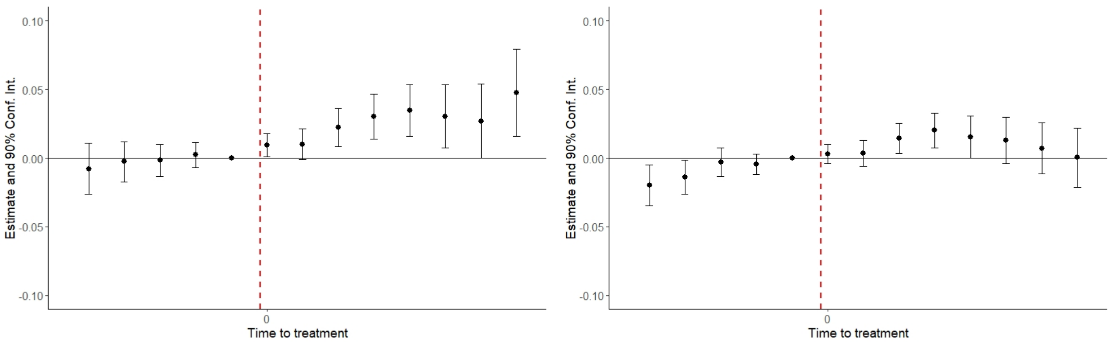
Figure D.12: Effect of windfalls on undereducation - Anticipation



(a) Unexpected (b) Expected + Do not know

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on undereducation, UE_{ito}, for each relative treatment period. We distinguish between individuals that expect (Figure D.12b) and not expect (Figure D.12a) an inheritance or gift in the near future. The question is asked in the 2001 survey. The sample includes the survey years between 2000 and 2020. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

Figure D.13: Effect of windfalls on educational mismatch - Main event studies with additional control variables

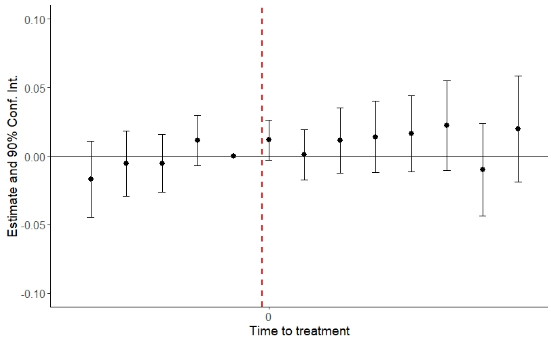


(a) $OE_{ito}(ATT : 0.028**)$ (b) $UE_{ito}(ATT : 0.010)$

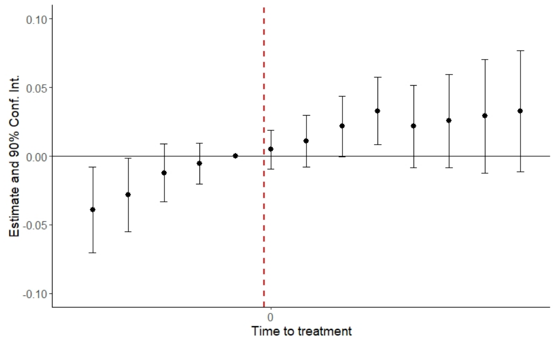
Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.13a) and undereducation, UE_{ito}, (Figure D.13b) for each relative treatment period. Compared to the main specification, we additionally control for industry, occupation, working in the public sector and education level. The sample includes the survey years between 2000 and 2020. * p $<$ 0.1, ** p $<$ 0.05, *** p $<$ 0.01.

5.6.3 Role of the size of the windfall gain

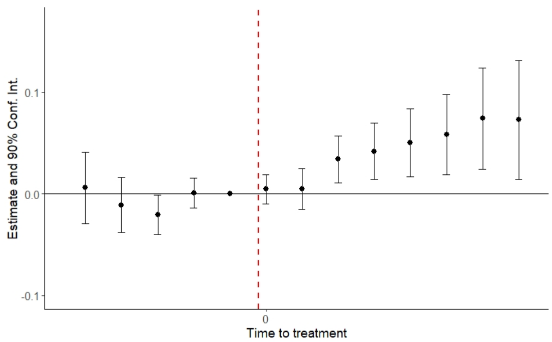
Figure D.14: Effect of windfalls on educational mismatch - Size of windfalls



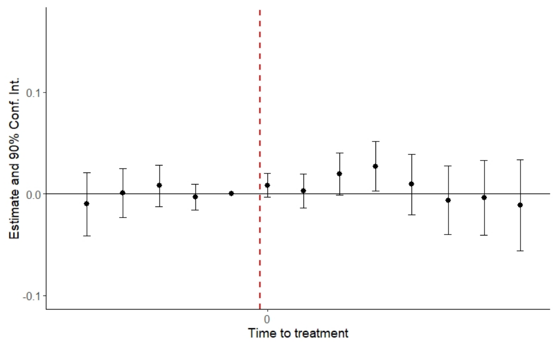
(a) Effect on OE_{ito} for $S_{i,t}^{SMALL}$



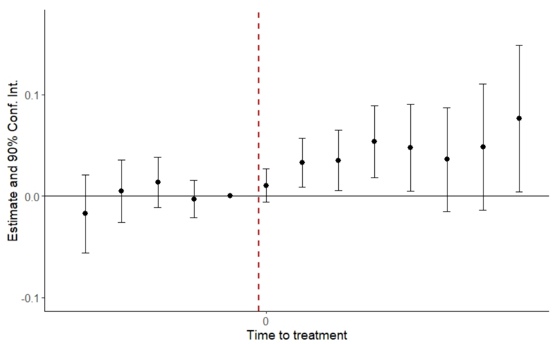
(b) Effect on UE_{ito} for $S_{i,t}^{SMALL}$



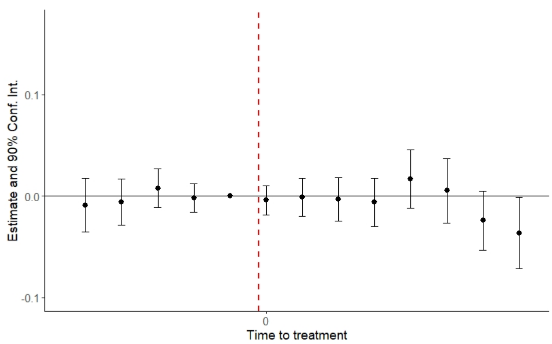
(c) Effect on OE_{ito} for $S_{i,t}^{MEDIUM}$



(d) Effect on UE_{ito} for $S_{i,t}^{MEDIUM}$



(e) Effect on OE_{ito} for $S_{i,t}^{LARGE}$



(f) Effect on UE_{ito} for $S_{i,t}^{LARGE}$

Note: Event study analyses based on Sun and Abraham (2021) showing treatment effects on overeducation, OE_{ito}, (Figure D.14a, Figure D.14c and Figure D.14e) and undereducation, UE_{ito}, (Figure D.14b, Figure D.14d and Figure D.14f) for each relative treatment period. We distinguish between the relative size of the windfall. The sample includes the survey years between 2000 and 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.3: Effect of windfalls on educational mismatch - Relative size of windfalls (normalised to household size)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:		OE_{it}			UE_{it}	
Amount	$S^{SMALL}_{i,t}$	$S^{MEDIUM}_{i,t}$	$S^{LARGE}_{i,t}$	$S^{SMALL}_{i,t}$	$S^{MEDIUM}_{i,t}$	$S^{LARGE}_{i,t}$
Windfall Gain	0.021	0.038**	0.040*	0.031*	0.003	-0.008
	(0.014)	(0.018)	(0.021)	(0.016)	(0.013)	(0.011)
Num. obs.	10684	11119	11290	10684	11119	11290
R^2	0.750	0.738	0.744	0.661	0.702	0.687
Within R^2	0.044	0.042	0.046	0.051	0.054	0.045

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between the relative size of the payment normalised by the number of persons in the household. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.6.4 Heterogeneity

Table D.4: Heterogeneity by financial situation

	(1)	(2)	(3)	(4)
Dependent variables:	OE_{ito}		UE_{ito}	
Panel A: Household income	**Lower**	**Higher**	**Lower**	**Higher**
Windfall gain	0.042*	0.057**	0.009	0.006
	(0.024)	(0.026)	(0.016)	(0.008)
Num. obs.	16,554	16,539	16,554	16,539
R^2	0.764	0.720	0.666	0.692
Within R^2	0.025	0.034	0.038	0.035
Panel B: Household income sat.	**Lower**	**Higher**	**Lower**	**Higher**
Windfall gain	0.025	0.081**	0.015	0.0006
	(0.016)	(0.040)	(0.012)	(0.010)
Num. obs.	17,019	15,546	17,019	15,546
R^2	0.760	0.722	0.670	0.687
Within R^2	0.028	0.037	0.035	0.036
Panel C: Debt	**No**	**Yes**	**No**	**Yes**
Windfall gain	0.050**	0.047	0.004	0.013
	(0.021)	(0.032)	(0.012)	(0.010)
Num. obs.	23,269	9,402	23,269	9,402
R^2	0.741	0.737	0.671	0.692
Within R^2	0.020	0.064	0.028	0.060
Panel D: Home owner	**No**	**Yes**	**No**	**Yes**
Windfall gain	0.050*	0.047**	-0.004	0.016
	(0.027)	(0.022)	(0.009)	(0.014)
Num. obs.	14,506	18,587	14,506	18,587
R^2	0.746	0.737	0.692	0.666
Within R^2	0.040	0.028	0.044	0.029

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between financial characteristics in the last observed period before the treatment. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.5: Heterogeneity by individual characteristics

	(1)	(2)	(3)	(4)
Dependent variables:	OE_{ito}		UE_{ito}	
Panel A: Gender	**Male**	**Female**	**Male**	**Female**
Windfall gain	0.045**	0.052*	0.002	0.011
	(0.021)	(0.028)	(0.010)	(0.015)
Num. obs.	16058	17035	16058	17035
R^2	0.760	0.720	0.694	0.662
Within R^2	0.034	0.029	0.031	0.040
Panel B: Age	**18-43**	**44-63**	**18-43**	**44-63**
Windfall gain	0.058***	0.009	0.009	0.001
	(0.021)	(80.2)	(0.011)	(71.7)
Num. obs.	16577	16516	16577	16516
R^2	0.717	0.765	0.655	0.704
Within R^2	0.034	0.030	0.040	0.031
Panel C: Children in household	**No**	**Yes**	**No**	**Yes**
Windfall gain	0.051**	0.036**	-0.012	0.041**
	(0.024)	(0.017)	(0.009)	(0.021)
Num. obs.	17752	15341	17752	15341
R^2	0.744	0.737	0.678	0.679
Within R^2	0.030	0.036	0.037	0.031
Panel D: Married	**No**	**Yes**	**No**	**Yes**
Windfall gain	0.028	0.078***	0.020	-0.010
	(0.020)	(0.030)	(0.013)	(0.012)
Num. obs.	21483	11610	21483	11610
R^2	0.749	0.725	0.688	0.666
Within R^2	0.023	0.046	0.025	0.058

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between individual characteristics in the last observed period before the treatment. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.6: Tertile splits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:		OE_{ito}			UE_{ito}	
Panel A: Age	**18-38**	**38-48**	**48-63**	**18-38**	**38-48**	**48-63**
Windfall gain	0.058**	-0.0002	0.016	0.0007	-0.004	0.010
	(0.024)	(0.013)	(85.6)	(0.012)	(0.016)	(72.9)
Num. obs.	11031	11031	11031	11031	11031	11031
R^2	0.711	0.749	0.776	0.666	0.643	0.729
Within R^2	0.050	0.042	0.047	0.060	0.047	0.044
Panel B: Household income	**392-2,741**	**2,741-3,899**	**$>$3,899**	**392-2,741**	**2,741-3,899**	**$>$3,899**
Windfall gain	0.029	0.032	0.101***	-0.006	0.034*	-0.005
	(0.026)	(0.025)	(0.039)	(0.013)	(0.018)	(0.010)
Num. obs.	11032	11030	11031	11032	11030	11031
R^2	0.759	0.743	0.732	0.689	0.640	0.721
Within R^2	0.037	0.046	0.056	0.054	0.054	0.050
Panel C: Household satisfaction	**0-7**	**7-8**	**8-10**	**0-7**	**7-8**	**8-10**
Windfall gain	0.006	0.051*	0.068	0.037	0.003	0.004
	(0.010)	(0.026)	(0.045)	(0.029)	(0.010)	(0.011)
Num. obs.	10856	10854	10855	10856	10854	10855
R^2	0.767	0.744	0.731	0.682	0.697	0.681
Within R^2	0.046	0.050	0.048	0.053	0.047	0.041

Note: The results display the ATT of windfall gains on overeducation, OE_{ito}, and undereducation, UE_{ito}, based on the Sun and Abraham (2021) dynamic estimator. We distinguish between age, household income and household income satisfaction tertiles in the last observed period before the treatment. People between the ages of 18 and 38 are included in the first age tertile, those between the ages of 38 and 48 in the second, and those between the ages of 48 and 63 in the third tertile. For household income, the tertiles range from EUR 392-2,741, from EUR 2,741-3,899 and above EUR 3,899. For household income satisfaction, the tertiles range from 0-7, 7-8 and 8-10. Note that the bands partly overlap to keep the subsamples at the same size. The sample includes the years between 2000 and 2020. The incidence of overeducation (OE_{ito}) and undereducation (UE_{ito}) are based on the measure defined in Section 5.3.1. Standard errors are clustered at the individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 6

What an (un)favourable match: Public sector employment and the reversal of the overeducation-job satisfaction penalty*

It is a well-documented phenomenon that individuals with higher education than required for their job report lower job satisfaction. However, whether this also applies to public sector employees remains unclear. The German case reveals a negative relationship between overeducation and job satisfaction in the private sector, which is reversed to positive for public sector employees. This holds robust across various empirical alterations. Furthermore, it is revealed that individuals with altruistic motives and a stronger-than-average family orientation drive this positive relationship.

Keywords: Job satisfaction; Overeducation; Public sector; Altruism; Family orientation

*This chapter is published in the Journal of Happiness Studies (Geißler, 2025). It is included under the terms of the Creative Commons Attribution License (CC BY 4.0). I thank the anonymous reviewers for their helpful comments and suggestions. Also, I am grateful for helpful comments by Anthony Lepinteur, Laszlo Goerke, Paul Matthias Gorny and Sven Hartmann, as well as the participants of the 4th Annual Meeting of the Berlin Network for Research on Well-being (BeWell), the 5th International Conference on European Studies (CefES), the 27th Annual Conference of the Society for Institutional and Organizational Economics (SIOE), the 2024 Annual Conference of the Royal Economic Society (RES) and the 2024 Annual Conference of the Scottish Economic Society (SES).

6.1 Introduction

Despite rising educational attainment, contemporary labour markets face a persistent shortage of skilled labour (European Commission, 2023; OECD, 2022a).[107] In Germany, for example, the number of university graduates doubled between 1997 and 2021. However, the skilled labour pool diminished due to a 15% decline in completed vocational training courses and apprenticeships between 2012 and 2022 (Statistisches Bundesamt, 2022a,b). This contrasts with a consistent demand for skilled labour, surpassing the demand for tertiary education degrees by almost threefold (Institut für Arbeitsmarkt- und Berufsforschung, 2024). A direct consequence of these developments is a qualification-job mismatch and, more precisely, overqualification, as competition for the limited number of jobs requiring higher education increases (e.g., Charlot and Decreuse, 2005; Dolado et al., 2000; Lazear et al., 2018). This induced a rise in overqualification from 11.8% to 18.8% in Germany between 2014 and 2019, slightly above the share in the US of 15.8% and the EU average of 15.3% (OECD, 2022b; Statistisches Bundesamt, nd).

Formal overeducation, the state in which the attained education exceeds the job requirements (e.g., Acemoglu, 1999; Feldman, 1996; Rumberger, 1981a; Sicherman and Galor, 1990), implies an overinvestment in formal attainment compared to the rewards of the job (e.g., Duncan and Hoffman, 1981; Korpi and Tåhlin, 2009; Sattinger and Hartog, 2013). Due to this imbalance, overeducation is related to a series of adverse outcomes (e.g., Nunley et al., 2017). In the context of job satisfaction, overeducated employees are likely to experience relative deprivation (Crosby, 1976) because their expectations, shaped through their educational attainment and comparisons with similarly educated peers, are not met in their current roles, resulting in lower job satisfaction (e.g., Bedemariam and Ramos, 2021; Burris, 1983; King and Hautaluoma, 1987; Peiró et al., 2010; Voces and Caínzos, 2021). Empirically, a large number of studies provide evidence for such a relationship (e.g., Burris, 1983; Chuang and Liang, 2022; Fleming and Kler, 2008, 2014; Sánchez-Sánchez and McGuinness, 2015). However, most of these studies do not distinguish between private and public sector employment (e.g., De Santis et al., 2021; Green and Zhu, 2010; McGuinness and Byrne, 2015; McGuinness and Sloane, 2011; Voces and Caínzos, 2021) or focus exclusively on private sector employees (e.g., Hersch, 1991). As a result, it remains unclear whether the negative relationship between overeducation and job satisfaction applies equally to both groups. This is particularly questionable, as previous

[107] For a detailed discussion of this paradox in the US, see, e.g., Cappelli (2015).

studies have found evidence in line with a trade-off perspective on overeducation. That is, individuals seem to accept overeducation, i.e., respond less negatively in terms of job satisfaction if they can trade this less favourable job match for other advantages (see e.g., Giuliano et al., 2024; McGuinness and Sloane, 2011).

Some of these advantages may be particularly relevant for public sector employees, as they differ from their private sector counterparts in many regards (for international evidence, see e.g., Bullock et al., 2015). Besides the fact that the majority of public sector employees are female in Germany, they are also more altruistically motivated, more risk-averse, and place a higher value on stability (Ayaita et al., 2019; Bellante and Link, 1981; Buurman et al., 2012; Chang, 2024; Dur and Van Lent, 2018; Dur and Zoutenbier, 2014, 2015; Hartog et al., 2002; Kim and Wiggins, 2011; Ko and Hur, 2014; Perry et al., 2010; Perry and Wise, 1990; Pfeifer, 2010; Tonin and Vlassopoulos, 2015). Many of these preferences are mirrored in the advantages of public sector employment, such as greater flexibility in working arrangements, job tasks aimed at serving society, and higher job security (e.g., Clark and Postel-Vinay, 2009; Deutscher Gewerkschaftsbund, 2022; Lewis and Frank, 2002). Following Person-Organization Fit Theory (P-O Theory), it can be expected that such alignment of preferences with the job conditions should induce higher levels of job satisfaction (e.g., Bright, 2008; Kim, 2012; Kristof, 1996; Steijn, 2008). If the benefit of these complementarities is traded against less attractive aspects, such as a less favourable job match (see McGuinness and Sloane, 2011), the relationship between overeducation and job satisfaction can be expected to differ for public and private sector employees.

Germany presents an interesting case for the investigation of this pattern as the public sector represents an essential cornerstone of the economy, which distinguishes between two types of public sector employment: Civil servants and public sector employees, who jointly represented a sixth of employees in Germany in 2021 (Bundeszentrale für Politische Bildung, Statistisches Bundesamt, Wissenschaftszentrum Berlin für Sozialforschung, and Bundesinstitut für Bevölkerungsforschung, 2021). Civil servants, who form the smaller part of the public sector, are usually appointed for life and have sovereign duties that help to safeguard public life. Their remuneration, ranks, and duties are highly formalised and settled by the Federal Civil Service Act ("Bundesbeamtengesetz") and the State Civil Service Act ("Landesbeamtengesetz"). In contrast, public sector employees have standard employment contracts, and collective wage agreements regulate their remuneration and duties (Bundesministerium des Inneren und für Heimat, nd). In both cases, the fields of activity cover a broad range of occupations, including public administration, military, police, health, education, science, or supply and disposal.

To explore the relationship between overeducation and job satisfaction, distinguishing between sectors, this article uses detailed information on individuals' traits, personalities, education, social backgrounds, and employment relations provided by the German Socio-Economic Panel (SOEP, Goebel et al., 2019). While the empirical strategy starts with a standard regression controlling for individual time-invariant unobservable heterogeneity, subsequent robustness checks consider selection into public sector employment through matching, weighting, and formal selection models. In line with previous studies examining only private sector employees or not differentiating between sectors (e.g., De Santis et al., 2021; Fleming and Kler, 2008, 2014; Hersch, 1991; Jones et al., 2014), the results show a negative relationship between overeducation and job satisfaction for private sector employees. In contrast, overeducation is associated with higher levels of job satisfaction among public sector employees. Finally, the rich information of the SOEP reveals that public sector employees with high levels of altruistic motivation and family-related life goals primarily drive the main results, possibly because their preferences are more in line with the work environment.

This article contributes to two main strands of the literature. First, it is the first to show that public sector employment moderates the relationship between overeducation and job satisfaction, reversing the relationship to positive. This adds to previous research investigating whether the negative relationship between overeducation and job satisfaction applies to all employees (Bedemariam and Ramos, 2021; Belfield, 2010; Danzer, 2019; Giuliano et al., 2024; Peiró et al., 2010; Verhaest and Omey, 2009) and enhances the understanding of the consequences of overeducation regarding job satisfaction (e.g., Battu et al., 1999; Fleming and Kler, 2008, 2014; Iseke, 2014; Jones et al., 2014; McGuinness and Byrne, 2015; Verhaest and Verhofstadt, 2016). Second, by delving into the role of motives and job conditions, this article adds to the current research landscape that evaluates differences in preferences between public and private sector employees and how their potential reflection in job conditions may affect job satisfaction (e.g., Ayaita et al., 2019; Banuri and Keefer, 2016b; Bellante and Link, 1981; Bullock et al., 2015; Buurman et al., 2012; Chang, 2024; Dur and Van Lent, 2018; Dur and Zoutenbier, 2014; Hartog et al., 2002; Kim and Wiggins, 2011; Ko and Hur, 2014; Lewis and Frank, 2002; Tonin and Vlassopoulos, 2015).

6.2 Previous literature and expectations

Previous research investigating factors related to job satisfaction established a negative correlation between educational attainment and job satisfaction (see e.g., Clark, 1997; Clark and Oswald, 1996; Clark et al., 1996; Gazioglu and Tansel, 2006). Still, some studies report contrasting results. For example, Fabra and Camisó (2009) find a positive link between education and job satisfaction caused by access to jobs that provide higher satisfaction levels. Moreover, Vila and García-Mora (2005) show that the relationship between education and job satisfaction is ambiguous and partly depends on the facet of job satisfaction considered. In contrast, they identify the mismatch between individuals' qualifications and the job requirements as a relevant determinant of job satisfaction and report a negative link with job satisfaction. Aligning with this, the majority of studies evaluating the relationship between overeducation and job satisfaction in the private sector or without differentiating between sectors report a significant negative correlation (Allen and van der Velden, 2001; Bedemariam and Ramos, 2021; Burris, 1983; Chuang and Liang, 2022; De Santis et al., 2021; Fleming and Kler, 2008, 2014; Giuliano et al., 2024; Green and Zhu, 2010; Hersch, 1991; Iseke, 2014; Jones et al., 2014; Lillo-Bañuls and Casado-Díaz, 2015; Mavromaras et al., 2012; McGuinness and Byrne, 2015; McGuinness and Sloane, 2011; Peiró et al., 2010; Sam, 2020; Sánchez-Sánchez and McGuinness, 2015; Tsang, 1987; Verhaest and Omey, 2006, 2009; Verhaest and Verhofstadt, 2016; Voces and Caínzos, 2021, see Table E.1).[108] Building on this, I expect a negative baseline correlation between overeducation and job satisfaction.

Still, there is preliminary evidence that certain factors mitigate this relationship. Peiró et al. (2010) employ data from the Occupation Observatory of Youth covering individuals living in Valencia, Barcelona and Madrid between 1996 and 2002. They show that the negative correlation between overeducation and extrinsic job satisfaction is moderated by work experience.[109] Similar evidence is reported by Verhaest and Verhofstadt (2016). Employing data for Flemish young workers, they report a negative relationship between overeducation and job satisfaction that reduces with years of experience. Additionally, their results reveal that the relationship between overeducation and job satisfaction is positively

[108]Note that Büchel (2002), Groot and Maassen van den Brink (1999) and Sohn (2010) do not report a negative relationship between overeducation and job satisfaction. Mavromaras et al. (2013) report a negative correlation with satisfaction with hours, only, while overeducation is negatively related to all other facets of job satisfaction only if the individual is also overskilled.

[109]Extrinsic job satisfaction refers to factors that are not associated with the job content itself, such as the wage, while intrinsic job satisfaction covers, for example, job autonomy (Peiró et al., 2010).

moderated by job autonomy. This is in line with studies arguing that, in some cases, educational mismatch, in particular overeducation, may contain voluntary elements as a result of trading unfavourable aspects of a job for favourable, desired aspects (see e.g., McGuinness and Sloane, 2011). Giuliano et al. (2024) provide evidence in line with such a trade-off perspective. Using data from the CEDEFOP European Skills and Jobs Survey from 2014, they show that the negative correlation between overeducation and job satisfaction is particularly pronounced among temporarily employed individuals while being lower among individuals with permanent employment contracts. Another factor that might moderate the relationship but has not yet been explored in full detail is public sector employment.[110]

The public sector, particularly in Germany and other European countries, offers some favourable working conditions, such as flexible working arrangements resulting in a better work-life-balance, higher job security, less tight deadlines, or a feeling of contributing to society (e.g., Ayaita et al., 2019; Clark and Postel-Vinay, 2009; Deutscher Gewerkschaftsbund, 2022; Ellguth et al., 2019; Fenizia et al., 2024; Lewis and Frank, 2002). Simultaneously, previous research provided evidence that public sector employees are more altruistically and intrinsically motivated, more risk averse, seem to select jobs with a better work-life balance by placing high emphasis on the compatibility of family and work, and tend to prefer jobs with lower required effort (Ayaita et al., 2019; Bellante and Link, 1981; Buelens and Van den Broeck, 2007; Buurman et al., 2012; Chang, 2024; Dur and Van Lent, 2018; Dur and Zoutenbier, 2014, 2015; Ehlert and García-Morán, 2022; Ezra and Deckman, 1996; Hartog et al., 2002; Kim and Wiggins, 2011; Ko and Hur, 2014; Perry et al., 2010; Perry and Wise, 1990; Pfeifer, 2010; Saltzstein et al., 2001; Tonin and Vlassopoulos, 2015). In line with P-O Theory, such a match (supplementary congruence) between the characteristics of the job and the individual's traits is related to higher job satisfaction (e.g., Bright, 2008; Kim, 2012; Kristof, 1996; Steijn, 2008).

If individuals choose public sector employment because it better fulfils their personal preferences, for example, regarding the compatibility of family and job responsibilities, it may be that their job satisfaction is less negatively affected due to a potential trade-off between overeducation and the benefits of supplementary congruence (cf., Giuliano et al., 2024; McGuinness and Sloane, 2011). From this, I

[110]Belfield (2010) studies overeducation and job satisfaction by providing split sample analyses. While the correlation is negative and significant in both sectors, the predicted percentages of satisfied individuals show differences. Danzer (2019) studies the private-public sector differential in job satisfaction and uses skill mismatch as an additional covariate.

expect that the negative relationship between overeducation and job satisfaction will be mitigated by public sector employment.

6.3 Data and methodology

To examine whether the relationship between overeducation and job satisfaction differs across sectors, representative survey data from the SOEP from 1985 to 2020 are used.[111]

6.3.1 Core variables

The dependent variable is individual job satisfaction. The variable is constructed using the yearly questions summarised under the central question: "How satisfied are you today with the following areas of your life?", of which one sub-question regards one's job. Individuals are asked to rate their satisfaction on a scale ranging from "0 - Low" to "10 - High". This measure is standardised in the regressions following prior satisfaction literature (e.g., D'Ambrosio et al., 2020; Lepinteur, 2019).

Overeducation is defined using the statistical measure (Clogg and Shockey, 1984; Verdugo and Verdugo, 1989), which defines the required education for one's job based on the average level of education attained by the reference group. Following, e.g., Blásquez and Budría (2012), the three-digit ISCO codes are used as reference groups. To account for credential and educational inflation over time, the mean years of education within these groups are estimated for each survey year. In accordance with Clogg and Shockey (1984) and Verdugo and Verdugo (1989), individuals with education exceeding the mean years of education within the occupation by more than one standard deviation in each survey year are considered overeducated.

To differentiate between individuals working in the public and the private sector, the survey asks: "Do you work for a public sector employer?". The dummy indicator equals one if the individual answered with "yes". Notably, the dummy includes regular employees of the public sector as well as civil servants, as defined earlier. As the goal of the present study is to distinguish the correlation between overeducation and job satisfaction in the private and public sector,

[111] SOEP-core, version 39 (DIW, 2024). Information from the first survey wave, 1984, is not used because of missing information on some confounding variables.

a sufficient range of occupations in both sectors is necessary to define and observe overeducation. Figure E.1 and Figure E.2 in the Appendix display the distribution of occupations at the three-digit ISCO level in both sectors. Although the distribution of occupations is slightly denser in the public sector, the graphical evidence suggests that variation in occupational dispersion is not a major concern when comparing sectors. This is also reflected in the fact that both sectors allow a distinction to be made between 155 occupations at the three-digit ISCO level.

6.3.2 Confounders

In the estimations, a range of individual and job-related variables is controlled for.[112] These include the civil status assessed by dummies, equal to one if the respondent is in any partnership (married or registered partnership) or has previously been in a partnership but no longer is (divorced or widowed). Continuous linear and non-linear (squared and divided by 100) measures of individuals' age supplement the individual confounders. The job and work-related controls contain dummies for working part-time, being a white-collar worker, and having a permanent contract. Furthermore, the firm size is controlled for by including three dummies equal to one if the firm has 20 to 199 employees, 200 to 1999 employees, or 2000 and more employees. In addition, the estimations include controls for one's contractual and overtime hours. Occupation fixed effects based on the one-digit ISCO classification are contained to negate the variation between occupations.[113] Moreover, heterogeneity across ten industries is accounted for. Finally, differences across years, seasons, and regional heterogeneity are captured by including year, month, and federal-state dummies.

6.3.3 Sample

Overeducation, job satisfaction, and public sector employment are constructs relevant exclusively to the working population. Thus, the sample is restricted

[112] Please note that dummies for migration background, sex, and the number of children are dropped in the fixed effects estimations due to the lack of variation over time.

[113] Keeping occupations constant implies that overeducation compares individuals working in the same occupation but having different education. In contrast, keeping education constant and excluding occupations would assess overeducation by comparing individuals having the same education but working in different occupations. Note that this distinction can greatly impact the interpretation of results (e.g., Duncan and Hoffman, 1981; Korpi and Tåhlin, 2009; Sattinger and Hartog, 2013). Using both indicators in parallel implies comparing individuals working in the same major occupational group and having the same education, narrowing the reference groups tremendously. It is, therefore, used only as a robustness check.

to observations of full- or part-time employed individuals. Observations of individuals older than their legal retirement age are dropped.[114] Moreover, the sample is cut to observations of individuals with working contracts of a maximum of 60 hours.[115] Furthermore, as self-employed individuals might differ from employed individuals in terms of educational requirements and, thus, overeducation, as well as regarding their satisfaction, the sample consists of observations of individuals in paid employment only (compare e.g., Acemoglu, 1999). Finally, all observations with missing information on the core variables and confounders are excluded. The final sample includes 181,900 observations of 39,701 individuals.

Table 6.1 provides the summary statistics for the entire sample and the descriptive statistics for both employment sectors, along with a comparison of means. 13% of the sample's observations are considered overeducated (cf., Bauer, 2002; Blásquez and Budría, 2012) and the average job satisfaction stands at 7.08 (cf., Cornelissen, 2009; Goerke and Huang, 2022; Khalil et al., 2023). Slightly above a quarter of the sample works in the public sector (cf., Becker et al., 2024; Görlitz and Tamm, 2020). The sex distribution is nearly equal, with the sample being almost evenly divided between males and females.[116] The average age is 42 years. Regarding civil status, 63.3% are married or in a partnership, 24.3% are single, and 12.4% are widowed or live separately from their former partners. Moreover, on average, respondents have 1.4 children, and most observations belong to individuals without a migration background, while 16.7% and 4.4% have direct or indirect migration backgrounds, respectively. Regarding employment characteristics, 62.1% of the participants work in white-collar positions, 5% are civil servants, and 90.8% hold permanent contracts. 76.4% work full-time. While the average weekly contractual working hours amount to 35, participants work an average of 2 overtime hours. Employment is almost evenly distributed among different firm sizes.

Several differences emerge when comparing the two employment sectors. While overeducation does not significantly differ between the groups, job satisfaction is, on average, significantly higher in the public sector.[117] Moreover,

[114] I compute the legal retirement age using the information provided in §235 of the German Social Security Code ("Sozialgesetzbuch") to adjust for the stepwise increase in the legal retirement age to 67 years since 2012 for the cohorts born since 1947.

[115] See §3 of the German Occupational Health and Safety Act ("Arbeitsschutzgesetz").

[116] Please note that gender has been assessed as a trichotomous variable (offering the options "male", "female" and "diverse") since 2019 only. This data limitation allows for a dichotomous distinction of biological sex exclusively.

[117] Overeducation rates do not significantly differ within the public sector between civil servants and public sector employees.

Table 6.1: Summary statistics and t-test by sector

	Full sample				Private		Public		T-test	
	Mean	Sd	Min	Max	Mean	Sd	Mean	Sd	Diff.	SE
Overeducation	0.130	0.336	0.000	1.000	0.130	0.336	0.130	0.336	0.000	0.002
Job satisfaction	7.077	1.977	0.000	1.000	7.036	1.994	7.192	1.923	0.156***	0.010
Job satisfaction (z)	-0.009	0.935	-3.357	1.374	-0.028	0.943	0.046	0.910	0.074***	0.005
Public	0.262	0.440	0.000	1.000						
Female	0.487	0.500	0.000	1.000	0.444	0.497	0.610	0.488	0.167***	0.003
Age	42.148	10.750	17.000	64.000	41.499	10.739	43.976	10.57	2.476***	0.057
Single	0.243	0.429	0.000	1.000	0.256	0.436	0.207	0.405	-0.049***	0.002
Married	0.633	0.482	0.000	1.000	0.625	0.484	0.654	0.476	0.029***	0.003
Separate/widowed	0.124	0.330	0.000	1.000	0.119	0.324	0.139	0.346	0.020***	0.002
Number of kids	1.409	1.183	0.000	15.000	1.381	1.186	1.486	1.168	0.104***	0.006
No migration	0.790	0.408	0.000	1.000	0.763	0.425	0.864	0.343	0.101***	0.002
Direct migration	0.167	0.373	0.000	1.000	0.189	0.392	0.104	0.305	-0.085***	0.002
Indirect migration	0.044	0.204	0.000	1.000	0.048	0.213	0.032	0.175	-0.016***	0.001
White-collar	0.621	0.485	0.000	1.000	0.603	0.489	0.669	0.470	0.066***	0.003
Civil servant	0.050	0.218	0.000	1.000	0.003	0.050	0.185	0.388	0.182***	0.002
Permanent	0.908	0.289	0.000	1.000	0.916	0.278	0.887	0.317	-0.029***	0.002
Full-time	0.764	0.425	0.000	1.000	0.787	0.409	0.698	0.459	-0.089***	0.002
Overtime	2.145	3.351	0.000	4.000	2.260	3.473	1.820	2.955	-0.440***	0.017
Working hours	35.182	7.943	2.000	60.000	35.559	7.688	34.119	8.531	-1.440***	0.044
Firm size: < 20	0.210	0.408	0.000	1.000	0.256	0.436	0.082	0.274	-0.174***	0.002
Firm size: 20-199	0.291	0.454	0.000	1.000	0.304	0.460	0.253	0.435	-0.050***	0.002
Firm size: 200-1999	0.238	0.426	0.000	1.000	0.213	0.410	0.307	0.461	0.093***	0.002
Firm size: ≥ 2000	0.256	0.437	0.000	1.000	0.223	0.416	0.351	0.477	0.129***	0.002
Num. obs.			181900			134298	47602			

Note: The sample is based on SOEP data from 1985 to 2020. (z) indicates that the variable is standardised. SOEP weights were applied.

the share of females is around 17 percentage points larger than in the private sector, and public sector employees are almost 2.5 years older, on average. Additionally, they are more likely to be married, to live separately, or to be widowed, and have more children. Besides that, the share of natives is 10 percentage points larger in the public sector.

The shares of white-collar workers and, by construction, that of civil servants are larger in the public sector. In contrast, the share of public sector employees who have permanent contracts is smaller than in the private sector, and a smaller share works full-time. Concerning working hours, public sector employees work, on average, 1.4 contractual hours and 0.4 overtime hours less than private sector employees. Finally, public sector employees work less often in firms with fewer than 200 employees and more often in firms above this threshold.

Table 6.2 further compares the means of job satisfaction in the private and public sector by overeducation status. The difference in means between the overeducated and non-overeducated is significant in the private sector, indicating that overeducated individuals in the private sector report lower job satisfaction than their non-overeducated counterparts. No such difference can be observed in the public sector, aligning with the expectations.

Table 6.2: Job satisfaction by sector and educational match

	No overeducation		Overeducation		T-test	
	Mean	Sd	Mean	Sd	Diff.	SE
Private						
Job satisfaction	7.051	1.996	6.934	1.973	-0.117***	0.016
Job satisfaction (z)	-0.021	0.944	-0.076	0.933	-0.055***	0.008
Num. obs.	134298					
Public						
Job satisfaction	7.192	1.933	7.194	1.861	0.002	0.025
Job satisfaction (z)	0.046	0.914	0.047	0.880	0.001	0.012
Num. obs.	47602					

Note: The sample is based on SOEP data from 1985 to 2020. (z) indicates that the variable is standardised. SOEP weights were applied.

6.3.4 Empirical approach

To test whether overeducation is differently related to job satisfaction in the private and public sector, the main estimation strategy follows previous literature (e.g, Clark et al., 2010; Cornelissen, 2009; Ferrer-i-Carbonell and Frijters, 2004; Song and Gao, 2020) and estimates:[118]

$$\begin{aligned} JS_{it} = \beta_0 + \beta_1 OE_{ito} + \beta_2 Pub_{it} + \beta_3 OE_{ito} \times Pub_{it} + \beta_4 X'_{it} \\ + \lambda_t + \kappa_m + \gamma_o + \eta_s + \phi_r + \alpha_i + \epsilon_{it} \end{aligned} \tag{6.1}$$

for i=1,...,I and t=1985,...,2020. JS_{it} is the standardised job satisfaction level of individual i at time t. OE_{ito} is equal to one if an individual i in occupation o at time t possesses more education than required by the current job. β_1 indicates the relationship between overeducation and job satisfaction and is expected to be negative. Pub_{it} contains whether individual i is employed in the public sector at time t, and β_2 is expected to be positive. β_3 is the coefficient of interest and contains the moderating effect of being overeducated and working in the public versus the private sector, which is also expected to be positive. X'_{it} is a vector containing the individual and job-related controls described above. λ_t, κ_m, γ_o, η_s, and ϕ_r comprise year, month, occupation, industry, and region dummies, respectively. α_i accounts for identification concerns based on individual unobserved time-invariant heterogeneity. This is particularly valuable when studying the impact of overeducation as it allows implicit control for factors such as one's unobserved abilities and skills (e.g., Agopsowicz et al., 2020). Finally, ϵ_{it} is the error term.

[118]All estimations are pursued with the software R in version 4.2.3. All fixed effects estimations are pursued with "fixest" by Bergé (2018).

6.4 Results

6.4.1 Main results

Table 6.3 presents the regression results based on Equation 1 (Table E.2 displays an extended set of results, including estimates for the job and demographic controls). Column (1) exclusively presents the raw correlations, including individual, month, and year fixed effects. In line with previous studies, the base correlation between overeducation and job satisfaction is negative (e.g., Burris, 1983; Chuang and Liang, 2022; Fleming and Kler, 2008, 2014; Hersch, 1991; Jones et al., 2014; Verhaest and Omey, 2006, 2009), while public sector employment is positively associated with job satisfaction (e.g., Green and Tsitsianis, 2005). In line with expectations, the interaction between overeducation and public sector employment is positive and significant. Column (2) adds federal-state fixed effects. Still, the coefficients remain unchanged. Column (3) includes occupational and industry fixed effects, distinguishing between ten industries and nine occupational groups and attenuating the negative correlation between overeducation and job satisfaction in the private sector. Additionally, the correlation between public sector employment and job satisfaction is reduced by around one-fifth through the addition of occupation and industry fixed effects. In contrast, the positive interaction between overeducation and public sector employment only reduces marginally by two in the third decimal place. Finally, adding demographics and job-related controls affects the coefficients for overeducation and the interaction term only marginally, while the public sector coefficient reduces again by approximately one-fifth.[119]

In sum, the negative correlation between overeducation and job satisfaction reported in prior research (e.g., Allen and van der Velden, 2001; Bedemariam and Ramos, 2021; Burris, 1983; Chuang and Liang, 2022; Fleming and Kler, 2008, 2014; Giuliano et al., 2024; Green and Zhu, 2010; Hersch, 1991; Iseke, 2014; McGuinness and Byrne, 2015; McGuinness and Sloane, 2011; Peiró et al., 2010; Sam, 2020; Sánchez-Sánchez and McGuinness, 2015; Tsang, 1987; Verhaest and Omey, 2006) holds only for private sector employees in the present study. In

[119] The results are qualitatively the same when re-estimating the regressions without fixed effects with standard OLS regressions (see Table E.3). Moreover, results are attenuated only marginally if civil servants are excluded from the sample (see Table E.4).

Table 6.3: Overeducation and job satisfaction

	(1)	(2)	(3)	(4)
Overeducation	-0.070***	-0.070***	-0.046**	-0.045**
	(0.015)	(0.015)	(0.015)	(0.015)
× public	0.069**	0.069**	0.067**	0.066**
	(0.023)	(0.024)	(0.023)	(0.023)
Public	0.082***	0.082***	0.067***	0.055***
	(0.014)	(0.014)	(0.014)	(0.014)
Individual FE	X	X	X	X
Month & year FE	X	X	X	X
Federal state FE		X	X	X
Industry & occ FE			X	X
Demographics & job controls				X
Num. obs.	181900	181900	181900	181900
R2 Adj.	0.386	0.387	0.387	0.388
R2 Within Adj.	0.010	0.011	0.012	0.014

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (4) contain fixed effects results. All columns contain individual, month, and year fixed effects. Columns (2) to (4) add federal state fixed effects, and columns (3) and (4) control for industry and occupation fixed effects. Column (4) includes demographic and job-related controls consisting of civil servant status, age (nominal and non-linear), civil status dummies, white-collar worker dummies, permanent contract ownership, full-time employment, and firm size dummies. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < .10, *p <.05, ** p <.01, *** p < .001.

contrast, conditional on the included covariates, the relationship is significantly attenuated and reversed to marginally positive in the public sector.[120]

6.4.2 Robustness

Covariates

Although industry and occupational fixed effects seem to be of large importance in the estimations presented above, the results may suffer from bias induced by limiting the variation in occupations to nine groups. Particularly, such bias arises as task variety within occupations may be large and would remain uncaptured in such a broad assessment of occupations. For this reason, columns (1) to (3) of Table E.6 include more granular occupational fixed effects based on the two-, three, and four-digit ISCO levels, distinguishing between 38, 155 and 574 occupations, respectively. Notably, even including the most granular occupational fixed effects (column (3) of Table E.6) reduces the coefficient of the interaction term by only two in the third decimal place compared to the base estimation (see Table 6.3).

[120] Using alternative measures of overeducation (indirect self-assessment and job analyst measure), the results obtained in column (4) can be qualitatively replicated. Anyhow, the interaction terms are imprecisely estimated and quantitatively smaller than the base coefficient in both cases (see Table E.5).

While the remaining set of covariates is apparently of minor relevance concerning the size of the overeducation coefficients and the model fit (see Table 6.3), it may be argued that the regressions miss three critical aspects, namely the individual's education, career stage, proxied by tenure with the firm, and wage. All three variables are not included in the main specification because tenure and wage may induce a bad control problem, while conditioning on education alters the reference group for the assessment of overeducation tremendously (e.g., Duncan and Hoffman, 1981; Korpi and Tåhlin, 2009; Sattinger and Hartog, 2013). Columns (4) to (6) of Table E.6 present the results containing the extended set of covariates. The results are qualitatively robust compared to Table 6.3, and coefficient sizes change only marginally in all cases.[121]

Job satisfaction scale

As explained in Section 6.3, job satisfaction is standardised in the main regressions. According to Bond and Lang (2019), the transformation of job satisfaction scales can lead to biased estimates, the sign of which likely depends on the chosen transformation. To account for such biases, two alternative transformations are proposed by Kaiser and Vendrik (2022), more specifically, a convex and a concave transformation of the form: $e^{jobsatisfaction_{it} \times 0.4}$ and $-e^{jobsatisfaction_{it} \times (-0.4)}$. As expected, due to the transformations, the coefficient size varies largely in columns (1) and (2) of Table E.8. Still, both columns provide evidence that aligns with the main results.

Balancing and matching

α_i accounts for identification concerns based on individual unobserved time-invariant heterogeneity in Equation 6.1. However, this approach cannot rule out selection into employment sectors based on observed heterogeneity although this may be relevant when considering the moderating impact of public sector employment on the correlation between overeducation and job satisfaction. A doubly robust entropy balancing and a propensity score matching approach, shown in columns (3) and (4) of Table E.8, allow to account for structural differences between groups more precisely than a data-rich covariate-based approach. This is because they estimate ex-ante weights, capturing the likelihood

[121] As overeducation is most likely among individuals with the highest educational attainment, while those may do structurally different tasks, and thus, report higher levels of job satisfaction, at the same time, I additionally estimated the base specification on a reduced sample excluding individuals with tertiary degrees. Results are qualitatively the same (see Table E.7).

of being observed in a certain group (e.g., Hainmueller, 2012). To construct the weights, the standard variables sex, age, civil status, number of children, migration background, being a white-collar worker, and education are used.[122] As discussed above, the inclusion of education has no huge impact on the baseline estimates, while previous studies linked education to the likelihood of public sector employment (e.g., Maczulskij, 2017). While including the estimated weights attenuates the coefficient sizes marginally, the evidence is qualitatively in line with the main regression.

Standard errors

Table 6.3 presents clustered standard errors at the individual level. While this enables accounting for within-individual correlations, it fails to adjust the estimations for the occupation-based definition of overeducation.[123] Thus, column (5) of Table E.8 adjusts the standard errors by clustering them at the three-digit ISCO level used to define the overeducation variable. While the standard errors increase marginally, the evidence still aligns with the base results.

6.4.3 Split sample validation

The results present robust evidence of a differential correlation between overeducation and job satisfaction in both employment sectors. Still, Table 6.1 reveals that public and private sector employees differ in almost all regards. Therefore, the relationship between overeducation and job satisfaction is also considered in the subsamples of private and public sector employees, accounting for differences between groups regarding all covariates. Table E.9 contains the baseline regression results without the interaction term for the private sector sample in column (1) and the public sector sample in column (2). The provided evidence is in line with the results of the interaction model, i.e., a negative correlation between overeducation and job satisfaction is found in the private sector. In contrast, the correlation is positive in the public sector.

[122]The weighting is run by using the "WeightIt"-package by Greifer (2023) with the options "ebal" and "glm" for entropy balancing and propensity score matching, respectively. The binary treatment is public. The Average Treatment effect on the Treated (ATT) is targeted to achieve doubly robust estimates (for more details see Greifer, 2023).

[123]For a detailed discussion on clustered standard errors see Abadie et al. (2023).

Heckman correction

As before, it should be noted that selection into public and private sector employment is non-random. A Heckman correction can account for this in the split sample analysis (Heckman, 1976, 1979). The Heckman correction uses a probit model to estimate the probability of being observed in the respective sample. From this estimation, the inverse mills ratio is predicted and used as an additional covariate in the regressions to account for non-random selection into public and private sector employment. As in the matching and balancing procedures, the standard variables sex, age, civil status, number of children, migration background, being a white-collar worker, and education are used in the probit model. Importantly, education is not included in the main regression as explained above.[124] Additionally, the probit model conditions on federal state, month, and year fixed effects. The results are presented in columns (3) and (4) of Table E.9 and are attenuated only marginally compared to the main results.

Endogeneity in overeducation

Finally, overeducation itself suffers from endogeneity. To provide additional evidence on robustness, information on the level of the father's education available in the SOEP is used to create an instrument equal to one if the respondent's father obtained A-levels or comparable. The rationale for this instrument is that father's education is likely linked with the overeducation probability of the child through the intergenerational transmission of educational attainment and its related impact on child's occupational choice (for a discussion of the intergenerational transmission of educational attainment, see e.g., Agüero and Ramachandran, 2020; Grönqvist et al., 2017; Huang, 2013; Maczulskij, 2017). However, the exclusion restriction necessary for a valid IV might be violated if the education of the father affects the job satisfaction of the child through paths other than the child's overeducation. For example, concerns may arise if the father's education is associated with the father's occupation, which could be related to the occupation of the child. Moreover, a link between the father's education and the education of the child may threaten the IV approach.[125] While the exclusion restriction cannot be tested empirically, I acknowledge the potential

[124]Adding fathers' education and fathers' occupation as exclusion restriction variables to the first stage of the Heckman correction does not alter the results qualitatively.

[125]Excluding these channels by incorporating a dummy equaling one if the father worked in a high-skill occupation or by controlling for the educational attainment of the child reveals similar results (see Table E.10). Importantly, an impact via the child's occupation is already excluded due to the occupational fixed effects in the base specification.

weaknesses of the instrument and refer to the Instrumental Variables Two-Stage Least Squares (IV2SLS) estimations only as supportive evidence of the previously discussed correlation.

As the instrument is time-invariant, columns (5) and (6) of Table E.9 present IV2SLS estimates without individual fixed effects. The first stage results are positive and significant in both sectors, indicating a higher likelihood of overeducation if one's father obtained A-levels or comparable, irrespective of the child's employment sector. In the second stage, the positive relationship between overeducation and job satisfaction is observed in the public sector sample, which aligns with the prior findings. In contrast, the negative relationship between overeducation and job satisfaction in the private sector fades in the IV2SLS setting.

In sum, the above-presented robustness checks reveal that the negative correlation between overeducation and job satisfaction is, if at all, only observable in the private sector. When considering public sector employees, this negative correlation is reversed to positive in both the interaction and split-sample framework.

6.5 Discussion

While previous research on the relationship between overeducation and job satisfaction focused on the private sector or did not differentiate between sectors in large parts (e.g., Lillo-Bañuls and Casado-Díaz, 2015; Sam, 2020), the results presented above reveal a differential relationship for private and public sector employees. As the positive relationship in the public sector diverges largely from the previous findings, it is particularly interesting to investigate which groups drive the premium in this sector. The first potential group covers females. Not only are they known to report higher satisfaction with their jobs (e.g., Clark, 1997), but the share of females working in the public sector is also more prominent than that of males (see Table 6.1; Statistisches Bundesamt, 2023a). Moreover, the theory of differential overqualification (Frank, 1978) and related studies (e.g., Büchel and Battu, 2003; McGoldrick and Robst, 1996) argue that females are at more considerable risk of overeducation due to more extensive limitations in regional mobility compared to males when searching for jobs.

In addition, the expectation that the relationship between overeducation and job satisfaction might be less pronounced among public than private sector employees was based on a potential trade-off between overeducation and a better fit of working conditions and individual preferences. Thus, it could be expected that the positive link between overeducation and job satisfaction in the public sector is particularly driven by individuals whose personality traits or preferences align better with the public sector or by those who benefit in particular from the more favourable working conditions. For example, prior literature (Ayaita et al., 2019; Banuri and Keefer, 2016a,b; Buurman et al., 2012; Chang, 2024; Dur and Van Lent, 2018; Dur and Zoutenbier, 2015; Tonin and Vlassopoulos, 2015) showed that public sector employees report higher levels of altruistic motivation for their job and are, on average, more risk-averse than private sector employees. Moreover, individuals valuing higher security are more likely to select the public sector (Bellante and Link, 1981; Lewis and Frank, 2002). While the desire to serve society is, by definition, mirrored in the tasks and duties of the public sector,[126] the preference for a lower level of risk is also reflected. In fact, the higher level of job security is often claimed to be among the main advantages of working in the public sector (e.g., Ayaita et al., 2019; Banuri and Keefer, 2016a,b; Tonin and Vlassopoulos, 2015). Still, not everyone in the public sector enjoys high job security as, for example, the share of temporary contracts is large (see Table 6.1; Statistisches Bundesamt, 2020a). This variation among public sector employees provides an opportunity to test whether job security is one of the drivers behind the main findings. If public sector employees value job security, but some do not have secure jobs, those with higher perceived job security may drive the positive relationship as their preferences better align with the job conditions (e.g., Bright, 2008; Kim, 2012; Kristof, 1996; McGuinness and Sloane, 2011; Steijn, 2008).

Moreover, it has been reported that reconciling family and work is particularly relevant for job satisfaction and turnover intentions among public sector employees (for a detailed review, see e.g., Ezra and Deckman, 1996; Kim and Wiggins, 2011; Ko and Hur, 2014; Saltzstein et al., 2001). Matching to these preferences, public sector employees often profit from flexible working time arrangements, such as working time accounts (Ellguth et al., 2019). Such flexible working time arrangements might themselves be positively related to job satisfaction due to an improved work-life balance (McNall et al., 2009) and more leisure time (Deng and Gao, 2017). Moreover, since public sector employees prioritise work-life balance and family compatibility, the negative effects of overeducation may be offset by the better realisation of their preferences in their

[126] For more information on the tasks of the public sector, see for example, the website of the German Federal Ministry of the Interior and Community (Bundesministerium des Innern und für Heimat): `https://www.bmi.bund.de/DE/themen/oeffentlicher-dienst/oeffentlicher-dienst-node.html`.

working conditions (e.g., Bright, 2008; Kim, 2012; Kristof, 1996; McGuinness and Sloane, 2011; Steijn, 2008).

The rich information of the SOEP allows for the inclusion of proxies for the mentioned individual preferences, as well as indicators capturing whether these preferences are currently fulfilled. Altruism and family orientation are assessed based on the items "Importance: To be socially and politically active" combined with "Importance: To help others" and "Importance: Owning a house", "Importance: Having a happy marriage/relationship", and "Importance: Having children", respectively (e.g., Headey, 2008; Richter et al., 2017). The scales, ranging from "1 - Very important" to "4 - Unimportant", are reversed and standardised to calculate the average across all individuals. A dummy variable is created, equalling one if an individual's altruism (family) value is above the overall mean, indicating a higher altruistic (family) orientation than in the full sample. Risk preferences are assessed by the question "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" rated on a scale ranging from "0 - Not at all willing to take risks" to "10 - Very willing to take risks". Proceeding in the same manner as before, the item is standardised. To proxy an individual's work-life-balance, I incorporate satisfaction with leisure time, which is measured based on an 11-point satisfaction scale using the question: "How satisfied are you right now with the following areas of life? - Your leisure time". Finally, information on perceived job security is assessed using the question: "Do you worry about losing your job?". The dummy equals one if the individual indicated no worries regarding job security.

Table 6.1 has already provided descriptive evidence in line with the expectations for females, i.e., the share of females in the public sector exceeds that in the private sector. Additionally, Table E.11 presents descriptive evidence on the newly introduced sources of heterogeneity. The evidence aligns with the anticipations in all cases except for family orientation. That is, the share of individuals having more altruistic motives than the average is significantly larger in the public sector than in the private sector. Moreover, the willingness to take risks is significantly lower among public sector employees, while average leisure satisfaction is larger. Finally, the share of individuals having no worries regarding their job security is 18.3 percentage points larger in the public sector. In contrast to the anticipations, no statistical differences in family orientation are discernible.

As almost all of the abovementioned attributes particularly suit public sector employees, Table E.12 only reports results for the subsample of public sector employees. Each column investigates one of the described aspects by including an interaction between overeducation and the respective source of heterogeneity. While the overeducation coefficient does not qualitatively vary by sex (column (1)), risk preferences (column (4)), leisure satisfaction (column (5)), or perceived job security (column (6)), columns (2) and (3) align with the expectations. In more detail, the positive correlation between overeducation and job satisfaction in the public sector is driven by individuals with altruistic motivation and family-related life goals.[127] This association suggests that overeducated individuals in the public sector could be particularly satisfied with their jobs because their preferences to serve society or to have working conditions facilitating the balance between family and work are better fulfilled by their jobs. In that regard, the evidence on family orientation is especially informative. Although family orientation is similar across sectors, public sector employees may be better able to reconcile their work with family obligations, which seems to mitigate the negative correlation between overeducation and job satisfaction, aligning with the trade-off perspective introduced above (cf., Giuliano et al., 2024; McGuinness and Sloane, 2011).

6.6 Conclusion

The number of individuals working in positions that do not fully utilise their educational qualifications is rising (OECD, 2022b). As overeducation remains persistent (Baert et al., 2013; Vera-Toscano and Meroni, 2021), understanding its consequences for individuals and society is becoming increasingly relevant. Moreover, it is crucial to evaluate whether the negative consequences of overeducation apply uniformly to all employees (e.g., Giuliano et al., 2024; Peiró et al., 2010). Following the premise that public sector employment might moderate the negative relationship, this article examines the relationship between overeducation and job satisfaction in both employment sectors.

Using data from the SOEP from 1985 to 2020, the results align with the expectation that overeducation relates differently to job satisfaction in the private and public sector. Specifically, private sector employees face lower job satisfaction compared to their non-overeducated counterparts, while overeducated public sector employees report higher satisfaction. This finding remains robust through various empirical alterations. The heterogeneity analysis reveals that this premium in the public sector is particularly pronounced for

[127] Despite the missing descriptive evidence for a family-orientation differential between sectors, results for family orientation are not discernible in the private sector sample, i.e., in the private sector, the coefficients of overeducation, and the interaction with the family dummy are negative, though insignificant.

individuals who arguably benefit most from the employment conditions and job content of the public sector, namely those who are altruistically motivated and those who are strongly family-oriented.

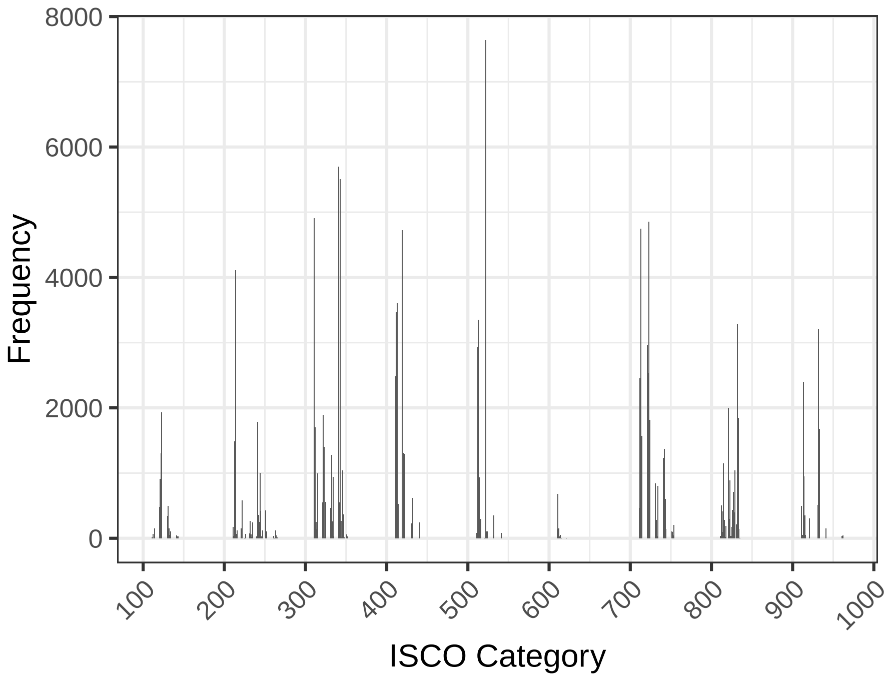
Whilst this paper offers insights, it also suffers from limitations that open avenues for future research. Firstly, the study identifies correlations instead of causal effects due to lacking sources of exogenous variation. Secondly, the study focuses on formal vertical overeducation, relying solely on attained and required levels of education. While this is insightful from an educational policy perspective, the measure does not enable an assessment of the full extent of task complexity within occupations. Moreover, the data do not allow for the measurement of skill mismatch where individuals' skills are compared to the skills required by their jobs (see e.g., McGuinness et al., 2018; Vera-Toscano and Meroni, 2021).[128] Future studies may follow research by, for example, Chevalier (2003) or Mavromaras et al. (2012), differentiating between overeducation and overskilling as well as the combination of the two to study their correlation with job satisfaction in the public and private sector. Thirdly, the study focuses on the German public and private sector, which have some particularities compared to other countries. Replicating this article in other contexts is necessary to determine whether the positive relationship between overeducation and job satisfaction is unique to the German public sector or applies to other labour markets, too. Fourthly, while the findings suggest that family-oriented and altruistically motivated individuals drive the findings in the public sector, it remains unclear whether private sector jobs that similarly meet their specific employees' needs and desires can also reverse the relationship. In any case, the findings of this research suggest that aligning working conditions and job content to individual preferences could be important in designing jobs that attract talented labour, particularly in times of a shrinking pool of applicants. Specifically, the public sector could benefit from emphasising the sense of purpose in the job and the compatibility of work with family life. Finally, both researchers and practitioners, especially in the public sector, should recognise that overeducation is not always evaluated negatively. Instead, it may be an active choice or the best available trade-off, in some cases.

[128] Note that the SOEP provides a variable measuring the usage of skills in the current job. But this variable is unsuitable to assess skill mismatch in the current job, as it measures the usage of skills compared to the last job. Moreover, the question was discontinued after 2009.

6.7 Appendix E

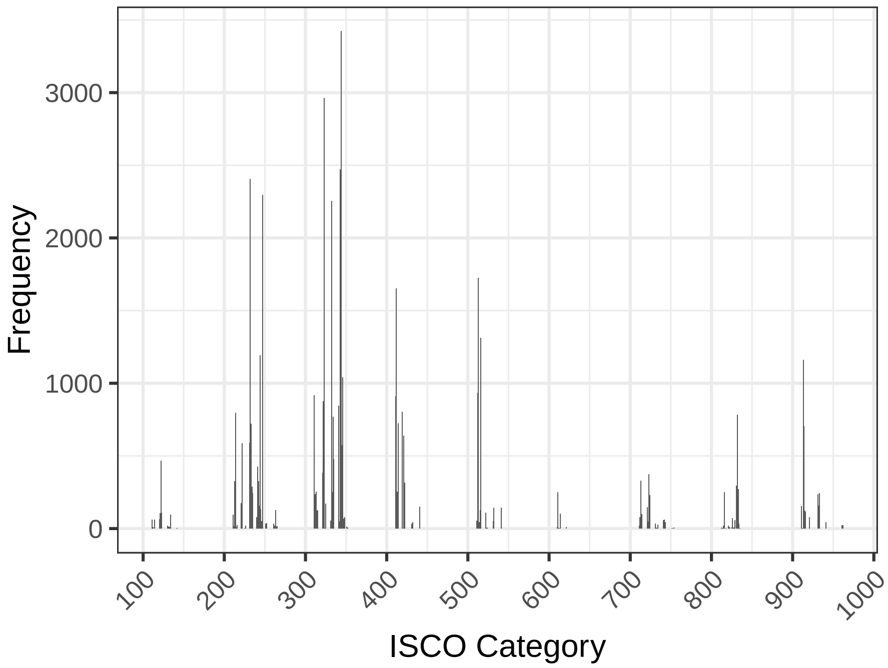
6.7.1 Figures

Figure E.1: Distribution of occupations (three-digit ISCO) in the private sector



Note: This figure is based on SOEP data from 1985 to 2020 on private sector employees. It displays the distribution of occupations at the three-digit ISCO level in the private sector.

Figure E.2: Distribution of occupations (three-digit ISCO) in the public sector



Note: This figure is based on SOEP data from 1985 to 2020 on public sector employees. It displays the distribution of occupations at the three-digit ISCO level in the public sector.

6.7.2 Tables

Table E.1: Comprehensive literature overview on overeducation and job satisfaction

Study	Country	Sectors	Measure	Result
Allen and van der Velden (2001)	The Netherlands	Not explained	ISA	-
Bedemariam and Ramos (2021)	Spain	Not explained	DSA	-
Belfield (2010)	UK	Both: split samples	DSA	-
Büchel (2002)	West Germany	Not explained	ISA	
Burris (1983)	US	Not explained	JA	-
Chuang and Liang (2022)	Taiwan	Not explained	ISA	-
De Santis et al. (2021)	Argentina	Not differentiated	ISA	-
Fleming and Kler (2008)	Australia	Not differentiated	JA	-
Fleming and Kler (2014)	Australia	Not differentiated	JA	-
Giuliano et al. (2024)	28 countries	Not explained	ISA	-
Green and Zhu (2010)	UK	Not differentiated	ISA	-
Groot and Maassen van den Brink (1999)	The Netherlands	Not differentiated	ISA	
Hersch (1991)	US	Private	ISA	-
Iseke (2014)	Germany	Not explained	ISA	-
Jones et al. (2014)	Australia	Not differentiated	MEAN	-
Lillo-Bañuls and Casado-Díaz (2015)	Spain	Not differentiated	DSA	-
Mavromaras et al. (2012)	Australia	Not differentiated	MODE	-
Mavromaras et al. (2013)	Australia	Not differentiated	MODE	
McGuinness and Byrne (2015)	11 countries	Not differentiated	DSA	-
McGuinness and Sloane (2011)	11 countries	Not differentiated	DSA	-
Peiró et al. (2010)	Spain	Not explained	ISA	-
Sam (2020)	Cambodia	Not differentiated	JA	-
Sánchez-Sánchez and McGuinness (2015)	15 countries	Not differentiated	DSA	-
Sohn (2010)	US	Not Explained	MEAN	
Tsang (1987)	US	Not differentiated	JA	-
Verhaest and Omey (2006)	Belgium	Not explained	DSA, ISA, JA, MEAN	-
Verhaest and Omey (2009)	Belgium	Not differentiated	JA	-
Verhaest and Verhofstadt (2016)	Belgium	Not differentiated	JA	-
Voces and Caínzos (2021)	Spain	Not differentiated	JA	-

Note: This literature overview is comprehensive, acknowledging that it does not assert absolute completeness. The measures distinguish the indirect self-assessment (ISA), where individuals are asked for the required level of education within their jobs; the direct self-assessment (DSA), asking individuals whether they are overeducated; the job analyst (JA) measure, relying on official classifications of occupations to assess the required level of education; and the statistical measure (MEAN or MODE). Allen and van der Velden (2001) do not report a significant negative correlation, as their significance threshold is 1%. Battu et al. (1999, 2000) use measures of overeducation slightly diverging from the standard measures, wherefore the studies are not included in the table, although in line with the negative correlation. Bedemariam and Ramos (2021) report a mitigating effect of career-enhancing strategies on the relationship between overeducation and job satisfaction. De Santis et al. (2021) consider the relationship between vertical and horizontal match and job satisfaction, differentiating different measures in the domains of satisfaction with pay, satisfaction with work environment and satisfaction with career prospects. Vertical match is not related to satisfaction with pay and is not related to satisfaction with the relationship with the supervisors for males. All remaining correlations hold for males and females. Fleming and Kler (2008) and Fleming and Kler (2014) estimate the relationship only for males and females, respectively. In the 2014 study, the negative relationship holds for females without children. Giuliano et al. (2024) use indicators of overeducation and overskilling, simultaneously differentiating apparent and genuine mismatches. The adverse effects on job satisfaction are largest for individuals being overeducated and overskilled. Moreover, they find this effect to be stronger for individuals with temporary contracts. Green and Zhu (2010) differentiate between formal and real overqualification, with the latter having a larger impact. Mavromaras et al. (2013, 2012) additionally use indicators of skill mismatch by combining the statistical method to assess overeducation with a direct assessment of the question, "I use many of my skills and abilities." The negative correlation between overeducation and job satisfaction holds for males except for male university graduates and females. In contrast, the job satisfaction facets are not significantly affected by overeducation for the female subsample. Mavromaras et al. (2013) report a significant relationship between overeducation and job satisfaction for the facet hours satisfaction only. All other facets of job satisfaction are negatively affected by overeducation only if the individual is overskilled, too. McGuinness and Sloane (2011) find a robust negative relationship between overeducation and job satisfaction only if overskilling is not controlled for. Peiró et al. (2010) differentiate extrinsic, intrinsic, and social job satisfaction and report a mitigation of the relationship between overeducation and extrinsic job satisfaction by work experience. Verhaest and Omey (2006) report the negative relationship between overeducation and job satisfaction in estimations conditioning on attained education only. Verhaest and Omey (2009) report a mitigation of the negative consequences of overeducation with raising years of experience. Verhaest and Verhofstadt (2016) show that autonomy moderates the negative relationship between overeducation and job satisfaction.

Table E.2: Overeducation and job satisfaction - Extended results

	(1)	(2)	(3)	(4)
Overeducation	-0.070***	-0.070***	-0.046**	-0.045**
	(0.015)	(0.015)	(0.015)	(0.015)
× public	0.069**	0.069**	0.067**	0.066**
	(0.023)	(0.024)	(0.023)	(0.023)
Public	0.082***	0.082***	0.067***	0.055***
	(0.014)	(0.014)	(0.014)	(0.014)
Civil servant				0.145***
				(0.044)
Age				-0.015
				(0.018)
Age (sq/100)				-0.012**
				(0.004)
Married				0.019
				(0.015)
Separate/ widowed				0.039+
				(0.021)
White-collar				0.083***
				(0.011)
Permanent				-0.045***
				(0.011)
Full-time				0.054***
				(0.014)
Firm size: 20-199				0.009
				(0.011)
Firm size: 200-1999				0.035**
				(0.013)
Firm size: ≥ 2000				0.058***
				(0.014)
Overtime				-0.002*
				(0.001)
Working hours				-0.002**
				(0.001)
Individual FE	X	X	X	X
Month & year FE	X	X	X	X
Federal state FE		X	X	X
Industry & occ FE			X	X
Num. obs.	181900	181900	181900	181900
R2 Adj.	0.386	0.387	0.387	0.388
R2 Within Adj.	0.010	0.011	0.012	0.014

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (4) contain fixed effects results. All columns contain individual, month, and year fixed effects. Columns (2) to (4) add federal state fixed effects, and columns (3) and (4) control for industry and occupation fixed effects. Column (4) includes demographic and job-related controls. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.3: OLS results

	(1)	(2)	(3)	(4)
Overeducation	-0.056***	-0.046***	-0.063***	-0.071***
	(0.012)	(0.012)	(0.012)	(0.012)
× public	0.056*	0.055*	0.065**	0.069**
	(0.023)	(0.023)	(0.023)	(0.023)
Public	0.072***	0.073***	0.022*	0.044***
	(0.010)	(0.010)	(0.011)	(0.012)
Month & year FE	X	X	X	X
Federal state FE		X	X	X
Industry & occ FE			X	X
Demographics & job controls				X
Num. obs.	181900	181900	181900	181900
R2 Adj.	0.006	0.01	0.014	0.024

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (4) contain OLS results. All columns contain month, and year fixed effects. Columns (2) to (4) add federal state fixed effects, and columns (3) and (4) control for industry and occupation fixed effects. Column (4) includes demographic and job-related controls. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.4: Results without civil servants

	(1)	(2)	(3)	(4)
Overeducation	-0.068***	-0.068***	-0.044**	-0.043**
	(0.015)	(0.015)	(0.015)	(0.015)
× public	0.054*	0.055*	0.053*	0.053*
	(0.025)	(0.025)	(0.025)	(0.025)
Public	0.082***	0.082***	0.067***	0.056***
	(0.014)	(0.014)	(0.014)	(0.014)
Individual FE	X	X	X	X
Month & year FE	X	X	X	X
Federal state FE		X	X	X
Industry & occ FE			X	X
Demographics & job controls				X
Num.Obs.	172774	172774	172774	172774
R2 Adj.	0.384	0.384	0.385	0.386
R2 Within Adj.	0.010	0.011	0.012	0.014

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Civil servants are dropped from the sample. All columns contain individual, month, and year fixed effects. Columns (2) to (4) add federal state fixed effects, and columns (3) and (4) control for industry and occupation fixed effects. Column (4) includes demographic and job-related controls. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < .10, *p <.05, ** p <.01, *** p < .001.

Table E.5: Alternative overeducation measures

	(1) Job analyst	(2) Indirect self-assessment
Overeducation	-0.033+	-0.068***
	(0.018)	(0.011)
× public	0.038	0.032
	(0.025)	(0.021)
Public	0.059***	0.057***
	(0.013)	(0.014)
Individual FE	X	X
Month & year FE	X	X
Federal state FE	X	X
Industry & occ FE	X	X
Demographics & job controls	X	X
Num.Obs.	181900	181900
R2 Adj.	0.388	0.388
R2 Within Adj.	0.013	0.014

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (2) contain fixed effects results and all covariates of column (4) in Table 6.3. Column (1) defines overeducation based on the job analyst approach, while column (2) employs the indirect self-assessment. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < .10, *p <.05, ** p <.01, *** p < .001.

Table E.6: Alternative covariates

	(1)	(2)	(3)	(4)	(5)	(6)
	Occupational fixed effects			Covariates		
	Two-dig	Three-dig	Four-dig	Education	Tenure	Income
Overeducation	-0.043**	-0.035*	-0.035*	-0.048**	-0.048**	-0.046**
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
× public	0.069**	0.068**	0.064**	0.067**	0.068**	0.066**
	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)
Public	0.052***	0.051***	0.048***	0.055***	0.051***	0.054***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)
Individual FE	X	X	X	X	X	X
Month & year FE	X	X	X	X	X	X
Federal state FE	X	X	X	X	X	X
Industry & occ FE	X	X	X	X	X	X
Demographics & job controls	X	X	X	X	X	X
Num. obs.	181900	181900	181900	181900	181900	181900
R2 Adj.	0.388	0.389	0.390	0.388	0.397	0.390
R2 Within Adj.	0.014	0.015	0.017	0.014	0.027	0.016

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (6) contain fixed effects results and all covariates of column (4) in Table 6.3. Columns (1) to (3) use more fine-grained occupational fixed effects using the two-, three-, and four-digit ISCO levels, respectively. Column (4) adds education dummies, while column (5) controls for tenure (level and quadratic term). Column (6) adds the deflated net income (log., base year 2015) to the covariates. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.7: Results without tertiary graduates

	(1)
Overeducation	-0.064**
	(0.020)
× public	0.071+
	(0.037)
Public	0.047**
	(0.015)
Individual FE	X
Month & year FE	X
Federal state FE	X
Industry & occ FE	X
Demographics & job controls	X
Num.Obs.	147208
R2 Adj.	0.393
R2 Within Adj.	0.015

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Tertiary graduates are dropped from the sample. Column (1) contains fixed effects results and all covariates of column (4) in Table 6.3. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < .10, *p <.05, ** p <.01, *** p < .001.

Table E.8: Alternative job satisfaction scale, covariate balance and standard errors

	(1)	(2)	(3)	(4)	(5)
	Job satisfaction scale		Covariate balance		Standard errors
	Concave	Convex	EB	PSM	Three-dig
Overeducation	-0.343*	-0.008**	-0.035*	-0.038*	-0.045*
	(0.139)	(0.003)	(0.017)	(0.017)	(0.018)
× public	0.421+	0.013**	0.063**	0.071**	0.066*
	(0.227)	(0.004)	(0.024)	(0.024)	(0.026)
Public	0.597***	0.007**	0.053***	0.039**	0.055***
	(0.130)	(0.003)	(0.014)	(0.014)	(0.014)
Individual FE	X	X	X	X	X
Month & year FE	X	X	X	X	X
Federal state FE	X	X	X	X	X
Industry & occ FE	X	X	X	X	X
Demographics & job controls	X	X	X	X	X
Num. obs.	181900	181900	181900	181900	181900
R2 Adj.	0.414	0.288	0.389	0.398	0.388
R2 Within Adj.	0.018	0.005	0.013	0.013	0.014

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) to (5) contain fixed effects results and all covariates of column (4) in Table 6.3. Columns (1) and (2) alter the job satisfaction scale by applying concave and convex transformations as proposed by Kaiser and Vendrik (2022). Columns (3) and (4) weight the regressions based on entropy balancing and propensity score weights, respectively. Job satisfaction is z-standardised in columns (3) to (5). In columns (1) to (4), standard errors are clustered at the individual level and are presented in parentheses. Column (5) adjusts the estimation of the standard errors and clusters them on the three-digit ISCO level. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.9: Split sample validation

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline		Heckman		IV	
	Private	Public	Private	Public	Private	Public
Overeducation	-0.041*	0.057*	-0.042**	0.056*	-0.049	0.454*
	(0.016)	(0.023)	(0.016)	(0.023)	(0.093)	(0.227)
Individual FE	X	X	X	X		
Month & year FE	X	X	X	X	X	X
Federal state FE	X	X	X	X	X	X
Industry & occ FE	X	X	X	X	X	X
Demographics & job controls	X	X	X	X	X	X
Num. obs.	134298	47602	134298	47602	134298	47602
R2 Adj.	0.391	0.435	0.391	0.435		
R2 Within Adj.	0.014	0.018	0.014	0.018		
First stage:					0.171***	0.099***
					(0.010)	(0.014)
F-stat.					2652.8	428.8

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1), (3), and (5) present results for the private sector sample, while columns (2), (4), and (6) contain results for the public sector sample. Columns (1) to (4) contain fixed effects results. All columns include the covariates of column (4) in Table 6.3. Columns (3) and (4) additionally include the inverse mills ratio to account for selection into the respective subsample. Columns (5) and (6) apply an instrumental variable approach using a dummy equalling one if one's father obtained A-levels as the instrument. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.10: IV specifications with additional covariates

	(1)	(2)	(3)	(4)
	Father occupation		Education	
	Private	Public	Private	Public
Overeducation	0.103	0.780*	0.020	0.933+
	(0.147)	(0.368)	(0.184)	(0.563)
Father high skill	-0.042*	-0.046		
	(0.021)	(0.033)		
Secondary			-0.050+	-0.153*
			(0.028)	(0.074)
Tertiary			-0.087	-0.404
			(0.111)	(0.251)
Month & year FE	X	X	X	X
Federal state FE	X	X	X	X
Industry & occ FE	X	X	X	X
Demographics & job controls	X	X	X	X
Num.Obs.	104026	38758	134298	47602
First stage:	0.127***	0.078***	0.087***	0.044***
	(0.012)	(0.015)	(0.009)	(0.012)
F-stat.	1049.5	191.7	865.6	97.3

Note: The sample is based on private and public sector employees observed in the survey waves from 1985 to 2020. Columns (1) and (3) present results for the private sector sample, while columns (2) and (4) contain results for the public sector sample. All columns include the covariates of column (4) in Table 6.3. Columns (1) to (4) apply an instrumental variable approach using a dummy equaling one if one's father obtained A-levels as the instrument. Columns (1) and (2) additionally control for having a father who worked in a high-skill occupation, while columns (3) and (4) control for the education of the individual. Job satisfaction is z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Table E.11: Sources of heterogeneity - T-test

	Private		Public		T-test	
	Mean	Sd	Mean	Sd	Diff	SE
Altruistic	0.326	0.469	0.405	0.491	0.079***	0.009
Family	0.484	0.500	0.494	0.500	0.010	0.009
Risk (z)	0.076	0.914	-0.031	0.888	-0.107***	0.010
Leisure sat (z)	-0.159	0.938	-0.091	0.938	0.069***	0.007
No worries	0.449	0.497	0.632	0.482	0.183***	0.004

Note: The sample is based on SOEP data from 1985 to 2020. Information on altruistic motives and family orientation is provided in 1990, 1992, 1995, 2004, 2008, 2012, and 2016. Willingness to take risks was assessed in 2004, 2006, and from 2008 to 2020. Information on leisure satisfaction is available from 1985 to 1989, 1991 to 1994, and since 1996. Information on worries regarding job security is contained in all waves. (z) indicates that the variable is standardised. SOEP weights were applied.

Table E.12: Heterogeneity among public sector employees

	(1) Sex	(2) Altruism	(3) Family	(4) Risk	(5) Leisure sat.	(6) Job security
Overeducation	0.058+	0.041	0.004	0.075**	0.052*	0.067*
	(0.032)	(0.092)	(0.111)	(0.029)	(0.022)	(0.032)
× female	-0.003					
	(0.044)					
× altruistic		0.215*				
		(0.096)				
× family			0.236*			
			(0.109)			
× risk (z)				0.013		
				(0.023)		
× leisure sat. (z)					-0.007	
					(0.019)	
× no worries						-0.016
						(0.032)
Altruistic		0.017				
		(0.038)				
Family			-0.038			
			(0.037)			
Risk (z)				0.029**		
				(0.010)		
Leisure sat (z)					0.189***	
					(0.008)	
No worries						0.141***
						(0.013)
Individual FE	X	X	X	X	X	X
Month & year FE	X	X	X	X	X	X
Federal state FE	X	X	X	X	X	X
Industry & occ FE	X	X	X	X	X	X
Demographics & job controls	X	X	X	X	X	X
Num. obs.	47602	7776	7776	27404	45758	47069
R2 Adj.	0.435	0.429	0.429	0.465	0.456	0.439
R2 Within Adj.	0.018	0.031	0.031	0.007	0.049	0.022

Note: The sample is based on public sector employees observed in the survey waves 1985 to 2020 in columns (1) and (6), 2004, 2008, 2012 and 2016 in columns (2) and (3), 2004, 2006 and 2008 to 2020 in column (4) and 1985 to 1989, 1991 to 1994 and 1996 to 2020 in column (5). Columns (1) to (6) contain fixed effects results and all covariates of column (4) in Table 6.3. Job satisfaction, leisure satisfaction, and risk aversion are z-standardised. Standard errors, clustered at the individual level, are presented in parentheses. + p < 0.10, *p <.05, ** p <.01, *** p < 0.001.

Chapter 7

Concluding remarks

Numbers from the OECD (nd) reveal that the number of tertiary graduates and the average educational attainment have risen consistently over the last decades. Considering even the relatively short period between 2011 and 2019, the share of individuals with tertiary degrees increased by more than a quarter in Germany, while the share of individuals with apprenticeships or vocational degrees declined (Statistisches Bundesamt, 2020b). This pattern contrasts with the requirements posted in job advertisements, which most frequently demand completed apprenticeships or training, and least often tertiary degrees (Institut für Arbeitsmarkt- und Berufsforschung, 2024). Building on previous conceptualisations combined with Signalling and Job Competition Theory, these developments can be expected to heighten the likelihood of educational mismatch, particularly of overeducation, within the population (e.g., Lazear et al., 2018; Spence, 1973; Thurow, 1975). While a large number of studies have aimed to shed light onto individual and job-related characteristics that can be linked to higher (lower) likelihoods of educational mismatch (e.g., Akgüç and Parasnis, 2023; Fleming and Kler, 2008), certain aspects have not been considered yet. Especially, perspectives on institutional features in the educational as well as the industrial relations system are underexplored. Moreover, while many studies have investigated the link between educational mismatch and wages (e.g., Duncan and Hoffman, 1981; Jacobs et al., 2023), it remains unclear how individuals' wealth affects the likelihood of an educational mismatch occurring at all.

The first part of this thesis approaches these gaps in the literature by investigating links between the educational system, the industrial relations system, individuals' wealth and educational mismatch using data for Germany from the SOEP. Chapter 2 evaluates individuals' entry to the educational system by examining the relationship between school-starting age and the likelihood of over- and undereducation. We exploit the exogenous variation in school-entry cutoffs and apply fuzzy regression discontinuity designs to account for endogeneity in the decision on when children start school. Doing so reveals a negative link between starting school one month older and the likelihood of undereducation in adulthood. Suggestive evidence indicates that occupational choice and educational attainment could explain this link, but the data do not allow for the identification of significant mediations. In summary, this chapter provides evidence that even decisions in early childhood may have long-lasting effects on the likelihood of educational mismatch, which manifest only several years later.

Similarly related to the educational system but focusing on early adulthood, Chapter 3 examines the link between educational costs and the likelihood of overeducation. The identification strategy relies on a quasi-experimental setting in Germany between 2006 and 2014, where tuition fees were introduced in seven federal states. The results show that treated individuals are significantly more likely to be overeducated after graduation than comparable individuals in the control group. Preliminary evidence suggests that the effect may be less pronounced among individuals in high-skill occupations and those treated for a shorter period. Aside from this, the effect of tuition fees is smaller among individuals with one parent possessing upper secondary education. Applying a dynamic perspective furthermore reveals that the positive link is not limited to the short run but can be observed even up to ten years after individuals graduated.

Changing the perspective from the educational system to labour market institutions, Chapter 4 investigates the relationship between trade union membership and the likelihood and extent of educational mismatch. Trade union members enjoy advantages in terms of bargaining power and information in Germany, which could help them in avoiding unfavourable educational (mis)matches. Aligning with this assumption, we find evidence that trade union membership negatively correlates with the likelihood and extent of overeducation, and is positively linked to the likelihood of being matched. However, there is no systematic association with undereducation. Estimating heterogeneity analyses distinguishing between groups with typically high and low union density, we conclude that these correlations likely arise from an improved bargaining power among union members, as informational advantages should not depend on the share of union members.

Finally, considering individuals' monetary resources instead of educational or labour market institutions, Chapter 5 examines the impact of a sudden increase in individuals' wealth through inheritances, gifts, or lottery winnings. As there is no empirical evidence on this link, we create a simple utility function approach to define our hypotheses. These predict that individuals transition into overeducation and out of undereducation after a windfall gain that exceeds a certain threshold. Empirically, we observe an increase in the likelihood of overeducation after the windfall gain. Further analyses reveal that this increase is driven by individuals who switch within or out of high-skill occupations. Moreover, these job changes are accompanied by larger amounts of leisure time and decreased working and overtime hours. This aligns with the model predicting that leisure time and overeducation, as an inferior job match, are traded when wealth rises. Additionally, in line with the formalisation, these effects are realised only for medium and large windfall gains.

In sum, the first part of this thesis reveals that the likelihood of undereducation is higher among individuals starting school older, while the likelihood of overeducation is increased by higher educational costs among graduates. Furthermore, the evidence suggests that the risk of overeducation can be reduced by support through trade unions and, thus, improved employee representation. Partially counter to previous evidence, the evaluation of windfall gains reveals that the perception of overeducation as an unfavourable and undesired match should not be generalised. This is as individuals select into overeducation once their wealth allows it, trading wage and job-match quality for non-monetary goods such as leisure time.

Taking this last aspect into account, one of the predominant findings in the educational mismatch literature may be questioned. In particular, if overeducation is indeed not always negatively evaluated, it might be that the previously established negative relationship between overeducation and individuals' job satisfaction might be too short-sighted. Chapter 6 focuses on this question by evaluating two mutually different groups of employees in Germany and distinguishes the relationship between overeducation and job satisfaction for public and private sector employees. The results align with the findings of the previous chapter as the established job satisfaction penalty materialises, if at all, only among private sector employees, while public sector employees report a premium. This positive link is driven by individuals with higher-than-average altruistic motivation and family orientation. Hence, the results align with the notion that the negative link between overeducation and job satisfaction may not be evident if the job characteristics correspond well with individuals' preferences and traits.

Bibliography

Abadie, A., Athey, S., Imbens, G., and Wooldridge, J. (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1):1–35.

Acemoglu, D. (1999). Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review*, 89(5):1259–1278.

Agopsowicz, A., Robinson, C., Stinebrickner, R., and Stinebrickner, T. (2020). Careers and mismatch for college graduates: College and noncollege jobs. *Journal of Human Resources*, 55(4):1194–1221.

Agüero, J. and Ramachandran, M. (2020). The intergenerational transmission of schooling among the education-rationed. *Journal of Human Resources*, 55(2):504–538.

Akgüç, M. and Parasnis, J. (2023). Occupation–education mismatch of immigrant women in Europe. *Social Indicators Research*, 170:75–98.

Alba-Ramirez, A. (1993). Mismatch in the Spanish labor market: Overeducation? *Journal of Human Resources*, 28(2):259–278.

Aldén, L. and Hammarstedt, M. (2016). Discrimination in the credit market? Access to financial capital among self-employed immigrants. *Kyklos*, 69(1):3–31.

Aleksynska, M. and Tritah, A. (2013). Occupation-education mismatch of immigrant workers in Europe: Context and policies. *Economics of Education Review*, 36:229–244.

Allen, J. and van der Velden, R. (2001). Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job-search. *Oxford Economic Papers*, 53(3):434–452.

Anderson, P., Butcher, K., Cascio, E., and Schanzenbach, D. (2011). Is being in school better? The impact of school on children's BMI when starting age is endogenous. *Journal of Health Economics*, 30(5):977–986.

Andrews, R. and Stange, K. (2019). Price regulation, price discrimination, and equality of opportunity in higher education: Evidence from Texas. *American Economic Journal: Economic Policy*, 11(4):31–65.

Angrist, J., Bettinger, E., Bloom, E., King, E., and Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review*, 92(5):1535–1558.

Angrist, J. and Krueger, A. (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association*, 87(418):328–336.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton.

Artz, B. and Welsch, D. (2021). Overeducation and wages revisited: A two-cohort comparison and random coefficients approach. *Southern Economic Journal*, 87(3):909–936.

Attar, I. and Cohen-Zada, D. (2018). The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth. *Journal of Economic Behavior & Organization*, 153:38–57.

Ayaita, A., Gülal, F., and Yang, P. (2019). Where does the good shepherd go? Civic virtue and sorting into public sector employment. *German Economic Review*, 20(4):571–599.

Bachmann, A. and Boes, S. (2014). Private transfers and college students' decision to work. *Economics of Education Review*, 42:34–42.

Badillo-Amador, L. and Vila, L. (2013). Education and skill mismatches: Wage and job satisfaction consequences. *International Journal of Manpower*, 34(5):416–428.

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

Baert, S., Marx, I., Neyt, B., Van Belle, E., and Van Casteren, J. (2018). Student employment and academic performance: An empirical exploration of the primary orientation theory. *Applied Economics Letters*, 25(8):547–552.

Bahrs, M. and Schumann, M. (2020). Unlucky to be young? The long-term effects of school starting age on smoking behavior and health. *Journal of Population Economics*, 33(2):555–600.

Bahrs, M. and Siedler, T. (2019). University tuition fees and high school students' educational intentions. *Fiscal Studies*, 40(2):117–147.

Balestra, S., Eugster, B., and Liebert, H. (2020). Summer-born struggle: The effect of school starting age on health, education, and work. *Health Economics*, 29(5):591–607.

Banuri, S. and Keefer, P. (2016a). Mellowing with tenure? Socialization increases prosocial behavior in public organizations. *Research in Experimental Economics*, 19:127–140.

Banuri, S. and Keefer, P. (2016b). Pro-social motivation, effort and the call to public service. *European Economic Review*, 83:139–164.

Barabasch, A., Cygan-Rehm, K., and Leibing, A. (2024). Timing of school entry and personality traits in adulthood. IZA Discussion Paper No. 17387.

Baran, J. (2024). Overeducation in the EU: Gender and regional dimension. *Labour Economics*, 90:102603.

Baron, R. and Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182.

Barr, A. (2019). Fighting for education: Financial aid and degree attainment. *Journal of Labor Economics*, 37(2):509–544.

Barth, E., Bryson, A., and Dale-Olsen, H. (2020). Union density effects on productivity and wages. *Economic Journal*, 130(631):1898–1936.

Basiglio, S., Rossi, M., and van Soest, A. (2023). Subjective inheritance expectations and economic outcomes. *Review of Income and Wealth*, 69(4):1088–1113.

Battu, H., Belfield, C., and Sloane, P. (1999). Overeducation among graduates: A cohort review. *Education Economics*, 7(1):21–38.

Battu, H., Belfield, C., and Sloane, P. (2000). How well can we measure graduate over-education and its effects? *National Institute Economic Review*, 171:82–93.

Bauer, T. (2002). Educational mismatch and wages: A panel analysis. *Economics of Education Review*, 21(3):221–229.

Becker, B., Goerke, L., and Huang, Y. (2024). Trade unions and life satisfaction in Germany. *British Journal of Industrial Relations*. Forthcoming.

Becker, G. (1975). *Human Capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press, Chicago, second edition.

Bedard, K. and Dhuey, E. (2012). School-entry policies and skill accumulation across directly and indirectly affected individuals. *Journal of Human Resources*, 47(3):643–683.

Bedemariam, R. and Ramos, J. (2021). Over-education and job satisfaction: The role of job insecurity and career enhancing strategies. *European Review of Applied Psychology*, 71(3):100632.

Behrens, M. (2009). Still married after all these years? Union organizing and the role of works councils in German industrial relations. *ILR Review*, 62(3):275–293.

Belfield, C. (2010). Over-education: What influence does the workplace have? *Economics of Education Review*, 29(2):236–245.

Bellante, D. and Link, A. (1981). Are public sector workers more risk averse than private sector workers? *ILR Review*, 34(3):408–412.

Belloc, I., Molina, J., and Velilla, J. (2025). Unexpected inheritances and household labor supply: Does the identity of the recipient matter? *Review of Income and Wealth*, 71(1):e12723.

Berglund, T. and Furåker, B. (2016). Employment protection regulation, trade unions and tenure of employment: An analysis in 23 European countries. *Industrial Relations Journal*, 47(5-6):492–512.

Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed effects: The R package FENmlm. CREA Discussion Papers No. 13.

Bernal, G. and Penney, J. (2019). Scholarships and student effort: Evidence from Colombia's Ser Pilo Paga program. *Economics of Education Review*, 72:121–130.

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., and Stevens, M. (2019). The long-run impacts of financial aid: Evidence from California's Cal Grant. *American Economic Journal: Economic Policy*, 11(1):64–94.

Bietenbeck, J., Leibing, A., Marcus, J., and Weinhardt, F. (2023). Tuition fees and educational attainment. *European Economic Review*, 154:104431.

Black, S., Denning, J., Dettling, L., Goodman, S., and Turner, L. (2023). Taking it to the limit: Effects of increased student loan availability on attainment, earnings, and financial well-being. *American Economic Review*, 113(12):3357–3400.

Black, S., Devereux, P., and Salvanes, K. (2011). Too young to leave the nest? The effects of school starting age. *Review of Economics and Statistics*, 93(2):455–467.

Blásquez, M. and Budría, S. (2012). Overeducation dynamics and personality. *Education Economics*, 20(3):260–283.

Bodvarsson, Ö. and Walker, R. (2004). Do parental cash transfers weaken performance in college? *Economics of Education Review*, 23(5):483–495.

Body, K., Bonnal, L., and Giret, J.-F. (2014). Does student employment really impact academic achievement? The case of France. *Applied Economics*, 46(25):3061–3073.

Boll, C., Leppin, J., and Schömann, K. (2016). Who is overeducated and why? Probit and dynamic mixed multinomial logit analyses of vertical mismatch in East and West Germany. *Education Economics*, 24(6):639–662.

Bonaccolto-Töpfer, M. and Schnabel, C. (2023). Is there a union wage premium in Germany and which workers benefit most? *Economies*, 11(50):1102005.

Bond, T. and Lang, K. (2019). The sad truth about happiness scales. *Journal of Political Economy*, 127(4):1629–1640.

Bracke, P., Pattyn, E., and von dem Knesebeck, O. (2013). Overeducation and depressive symptoms: Diminishing mental health returns to education. *Sociology of Health & Illness*, 35(8):1242–1259.

Bright, L. (2008). Does public service motivation really make a difference on the job satisfaction and turnover intentions of public employees? *American Review of Public Administration*, 38(2):149–166.

Bruckmeier, K., Fischer, G.-B., and Wigger, B. (2013). The willingness to pay for higher education: Does the type of fee matter? *Applied Economics Letters*, 20(13):1279–1282.

Bruckmeier, K. and Wigger, B. (2014). The effects of tuition fees on transition from high school to university in Germany. *Economics of Education Review*, 41:14–23.

Bruffaerts, R., Mortier, P., Kiekens, G., Auerbach, R., Cuijpers, P., Demyttenaere, K., Green, J., Nock, M., and Kessler, R. (2018). Mental health problems in college freshmen: Prevalence and academic functioning. *Journal of Affective Disorders*, 225:97–103.

Brugger, P., Blaeschke, F., Brackmann, T., Fußmann, M., and Leiste, M. (2022). Bildungsfinanzbericht 2022: Im Auftrag des Bundesministeriums für Bildung und Forschung und der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00074904/1023206227004_korr13122022.pdf, Accessed: 2024-02-14.

Brunello, G., Fort, M., Schneeweis, N., and Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviors? *Health Economics*, 25(3):314–336.

Bryson, A. (2014). Union wage effects: What are the economic implications of union wage bargaining for workers, firms, and society? *IZA World of Labor*, 35.

Büchel, F. (2002). The effect of overeducation on productivity in Germany – The firms' viewpoint. *Economics of Education Review*, 21(3):263–275.

Büchel, F. and Battu, H. (2003). The theory of differential overqualification: Does it work? *Scottish Journal of Political Economy*, 50(1):1–16.

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

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

Buelens, M. and Van den Broeck, H. (2007). An analysis of differences in work motivation between public and private sector organizations. *Public Administration Review*, 67(1):65–74.

Bullock, J., Stritch, J., and Rainey, H. (2015). International comparison of public and private employees' work motives, attitudes, and perceived rewards. *Public Administration Review*, 75(3):479–489.

Bundesministerium des Inneren und für Heimat (nd). The federal public service. https://www.bmi.bund.de/EN/topics/administrative-reform/federal-public-service/federal-publi-service-node.html, Accessed: 2024-06-10.

Bundeszentrale für Politische Bildung, Statistisches Bundesamt, Wissenschaftszentrum Berlin für Sozialforschung, and Bundesinstitut für Bevölkerungsforschung (2021). Datenreport 2021: Ein Sozialbericht für die

Bundesrepublik Deutschland. https://www.destatis.de/DE/Service/Statistik-Campus/Datenreport/Downloads/datenreport-2021.html, Accessed: 2024-06-14.

Burlando, A. (2023). Tuition fees and the intra-household allocation of schooling: Evidence from Uganda's free primary education reform. *Economics of Education Review*, 97:102466.

Burris, V. (1983). The social and political consequences of overeducation. *American Sociological Review*, 48(4):454–467.

Buurman, M., Delfgaauw, J., Dur, R., and Van den Bossche, S. (2012). Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior & Organization*, 83(3):279–291.

Cabanillas-Jiménez, G. (2024). Testing the permanent income hypothesis using the Spanish christmas lottery. *Applied Economic Analysis*, 32(96):232–259.

Callaway, B. and Sant'Anna, P. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Calonico, S., Cattaneo, M., and Farrell, M. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *Econometrics Journal*, 23(2):192–210.

Calonico, S., Cattaneo, M., Farrell, M., and Titiunik, R. (2023). Robust data-driven statistical inference in regression-discontinuity designs. 10.32614/CRAN.package.rdrobust.

Campbell, D. (2009). Civic engagement and education: An empirical test of the sorting model. *American Journal of Political Science*, 53(4):771–786.

Cappelli, P. (2015). Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the United States. *ILR Review*, 68(2):251–290.

Capsada-Munsech, Q. (2015). The role of social origin and field of study on graduates' overeducation: The case of Italy. *Higher Education*, 69(5):779–807.

Capsada-Munsech, Q. (2019). Measuring overeducation: Incidence, correlation and overlaps across indicators and countries. *Social Indicators Research*, 145(1):279–301.

Carman, K. (2013). Inheritances, intergenerational transfers, and the accumulation of health. *American Economic Review*, 103(3):451–455.

Caroleo, F. and Pastore, F. (2018). Overeducation at a glance: Determinants and wage effects of the educational mismatch based on AlmaLaurea data. *Social Indicators Research*, 137:999–1032.

Carroll, D. and Tani, M. (2015). Job search as a determinant of graduate over-education: Evidence from Australia. *Education Economics*, 23(5):631–644.

Castleman, B. and Long, B. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Chang, A. (2024). Risk aversion and public sector employment. *Public Administration Review*, 84(5):833–847.

Charlot, O. and Decreuse, B. (2005). Self-selection in education with matching frictions. *Labour Economics*, 12(2):251–267.

Chen, J. and Park, A. (2021). School entry age and educational attainment in developing countries: Evidence from China's compulsory education law. *Journal of Comparative Economics*, 49(3):715–732.

Chevalier, A. (2003). Measuring over-education. *Economica*, 70(279):509–531.

Chu Ng, Y. (2003). Earnings and "skill" allocation in the Canadian labor market. *International Journal of Manpower*, 24(8):964–980.

Chuang, Y.-C. and Liang, C.-Y. (2022). Overeducation and skill mismatch of university graduates in Taiwan. *Review of Development Economics*, 26(3):1693–1712.

Clark, A. (1997). Job satisfaction and gender: Why are women so happy at work? *Labour Economics*, 4(4):341–372.

Clark, A., Knabe, A., and Rätzel, S. (2010). Boon or bane? Others' unemployment, well-being and job insecurity. *Labour Economics*, 17(1):52–61.

Clark, A. and Oswald, A. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3):359–381.

Clark, A., Oswald, A., and Warr, P. (1996). Is job satisfaction u-shaped in age? *Journal of Occupational and Organizational Psychology*, 69(1):57–81.

Clark, A. and Postel-Vinay, F. (2009). Job security and job protection. *Oxford Economic Papers*, 61(2):207–239.

Clogg, C. and Shockey, J. (1984). Mismatch between occupation and schooling: A prevalence measure, recent trends and demographic analysis. *Demography*, 21(2):235–257.

Cohn, E. and Khan, S. (1995). The wage effects of overschooling revisited. *Labour Economics*, 2(1):67–76.

Cook, P. and Kang, S. (2016). Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation. *American Economic Journal: Applied Economics*, 8(1):33–57.

Cornelissen, T. (2009). The interaction of job satisfaction, job search, and job changes: An empirical investigation with German panel data. *Journal of Happiness Studies*, 10(3):367–384.

Cornelissen, T. and Dustmann, C. (2019). Early school exposure, test scores, and noncognitive outcomes. *American Economic Journal: Economic Policy*, 11(2):35–63.

Crawford, C., Dearden, L., and Greaves, E. (2014). The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 177(4):829–860.

Crawford, R. and Hood, A. (2016). Lifetime receipt of inheritances and the distribution of wealth in England. *Fiscal Studies*, 37(1):55–75.

Crosby, F. (1976). A model of egoistical relative deprivation. *Psychological Review*, 83(2):85–113.

D'Ambrosio, C., Jäntti, M., and Lepinteur, A. (2020). Money and happiness: Income, wealth and subjective well-being. *Social Indicators Research*, 148:47–66.

Danzer, N. (2019). Job satisfaction and self-selection into the public or private sector: Evidence from a natural experiment. *Labour Economics*, 57:46–62.

Darolia, R. (2014). Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students. *Economics of Education Review*, 38:38–50.

Davia, M., McGuinness, S., and O'Connell, P. (2017). Determinants of regional differences in rates of overeducation in Europe. *Social Science Research*, 63:67–80.

De Haan, M. (2011). The effect of parents' schooling on child's schooling: A nonparametric bounds analysis. *Journal of Labor Economics*, 29(4):859–892.

De Santis, M., Florensa, M., Gáname, M., and Moncarz, P. (2021). Job satisfaction of recent university graduates in economics sciences: The role of the match between formal education and job requirements. *Journal of Happiness Studies*, 22:3157–3197.

Dee, T. and Sievertsen, H. (2018). The gift of time? School starting age and mental health. *Health Economics*, 27(5):781–802.

Deng, S. and Gao, J. (2017). The mediating roles of work–family conflict and facilitation in the relations between leisure experience and job/life satisfaction among employees in Shanghai banking industry. *Journal of Happiness Studies*, 18:1641–1657.

Denning, J. (2019). Born under a lucky star: Financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources*, 54(3):760–784.

Deutscher Bundestag (2019a). Kurzinformation: Gehaltsoffenlegungspflichten in Stellenanzeigen und Tarifbindung. Aktenzeichen WD 6 - 3000 - 018/19. https://www.bundestag.de/resource/blob/662622/944ff8369885a598a7856d0d5a4794ca/WD-6-108-19-pdf-data.pdf, Accessed: 2024-06-24.

Deutscher Bundestag (2019b). Sachstand: Stellenausschreibungen im öffentlichen Dienst. Aktenzeichen WD 6 - 3000 - 020/19. https://www.bundestag.de/resource/blob/645592/a2da84d5bda635dc8d194028e0d95fc5/WD-6-020-19-pdf-data.pdf, Accessed: 2024-06-24.

Deutscher Gewerkschaftsbund (2014). Unterwertige Beschäftigung - Beleuchtung eines am Arbeitsmarkt vernachlässigten Problems. *Arbeitsmarktaktuell*, 02:1–21.

Deutscher Gewerkschaftsbund (2022). Personalreport öffentlicher Dienst 2021. https://www.dgb.de/fileadmin/download_center/Studien/DGB_Personalreport_%C3%96ffentlicher_Dienst_2021.pdf, Accessed: 2024-06-14.

Deutscher Gewerkschaftsbund (2023). Stellungnahme des Deutschen Gewerkschaftsbundes zum Referentenentwurf des Bundesministeriums für Arbeit und Soziales: Entwurf eines Gesetzes zur Stärkung der

Aus- und Weiterbildungsförderung und Einführung einer Bildungszeit. https://www.dgb.de/fileadmin/download_center/Stellungnahmen/Stellungnahme-zum-Referentenentwurf-Weiterbildungsgesetz.pdf, Accessed: 2025-01-27.

DGB Rechtsschutz (2024). Anteil der Streitgegenstände im Arbeitsrecht. https://www.dgbrechtsschutz.de/wir/publikationen/statistiken/, Accessed: 2025-01-26.

Dhuey, E., Figlio, D., Karbownik, K., and Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management*, 38(3):538–578.

Di Pietro, G. and Cutillo, A. (2006). University quality and labour market outcomes in Italy. *Labour*, 20(1):37–62.

Dicks, A. and Lancee, B. (2018). Double disadvantage in school? Children of immigrants and the relative age effect: A regression discontinuity design based on the month of birth. *European Sociological Review*, 34(3):319–333.

Dickson, M. (2013). The causal effect of education on wages revisited. *Oxford Bulletin of Economics and Statistics*, 75(4):477–498.

Dickson, M., Gregg, P., and Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes? *Economic Journal*, 126(596):F184–F231.

Diem, A. (2015). Overeducation among graduates from universities of applied sciences: Determinants and consequences. *Journal of Economic & Financial Studies*, 3(2):63–77.

Diem, A. and Wolter, S. (2014). Overeducation among Swiss university graduates: Determinants and consequences. *Journal of Labour Market Research*, 47:313–328.

DiNardo, J. and Lee, D. (2004). Economic impacts of new unionization on private sector employers: 1984–2001. *Quarterly Journal of Economics*, 119(4):1383–1441.

DIW (2022). Sozio-oekonomisches Panel (SOEP)-Version 37, Daten der Jahre 1984-2020, (SOEP-Core, v37, EU Edition). DIW-Deutsches Institut für Wirtschaftsforschung, Berlin. DOI: 10.5684/soep.core.v37eu.

DIW (2024). Sozio-oekonomisches Panel (SOEP)-Version 39, Daten der Jahre 1984-2022, (SOEP-Core, v39, EU Edition). DIW-Deutsches Institut für Wirtschaftsforschung, Berlin. DOI: 10.5684/soep.core.v39eu.

Dobkin, C. and Ferreira, F. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review*, 29(1):40–54.

Dodini, S., Salvanes, K., and Willén, A. (2021). The dynamics of power in labour markets: Monopolistic unions versus monopsonistic employers. CESifo Working Paper No. 9495.

Dodini, S., Salvanes, K., Willén, A., and Zhu, L. (2023a). The career effects of union membership. IZA Discussion Papers No. 16185.

Dodini, S., Stansbury, A., and Willén, A. (2023b). How do firms respond to unions? IZA Discussion Papers No. 16697.

Dolado, J., Felgueroso, F., and Jimeno, J. (2000). Youth labour markets in Spain: Education, training, and crowding-out. *European Economic Review*, 44(4-6):943–956.

Dolton, P. and Silles, M. (2008). The effects of over-education on earnings in the graduate labour market. *Economics of Education Review*, 27(2):125–139.

Dolton, P. and Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, 19(2):179–198.

Doorley, K. and Pestel, N. (2020). Labour supply after inheritances and the role of expectations. *Oxford Bulletin of Economics and Statistics*, 82(4):843–863.

Downey, D. (1995). When bigger is not better: Family size, parental resources, and children’s educational performance. *American Sociological Review*, 60(5):746–761.

Duncan, G. and Hoffman, S. (1981). The incidence and wage effects of overeducation. *Economics of Education Review*, 1(1):75–86.

Dur, R. and Van Lent, M. (2018). Serving the public interest in several ways: Theory and empirics. *Labour Economics*, 51:13–24.

Dur, R. and Zoutenbier, R. (2014). Working for a good cause. *Public Administration Review*, 74(2):144–155.

Dur, R. and Zoutenbier, R. (2015). Intrinsic motivations of public sector employees: Evidence for Germany. *German Economic Review*, 16(3):343–366.

Dwenger, N., Storck, J., and Wrohlich, K. (2012). Do tuition fees affect the mobility of university applicants? Evidence from a natural experiment. *Economics of Education Review*, 31(1):155–167.

Dynarski, S. (2008). Building the stock of college-educated labor. *Journal of Human Resources*, 43(3):576–610.

Dynarski, S., Libassi, C., Michelmore, K., and Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, 111(6):1721–1756.

Ehlert, A. and García-Morán, E. (2022). Workers' self-selection into public sector employment: A tale of absenteeism. *Kyklos*, 75(3):394–409.

Eisenberg, D., Golberstein, E., and Hunt, J. (2009). Mental health and academic success in college. *BE Journal of Economic Analysis & Policy*, 9(1):40.

Elinder, M., Erixson, O., and Ohlsson, H. (2012). The impact of inheritances on heirs' labor and capital income. *BE Journal of Economic Analysis & Policy*, 12(1):61.

Ellguth, P., Gerner, H., and Zapf, I. (2019). Working time accounts in the public - and private sector: Flexible working hours are becoming increasingly important. `https://www.iab-forum.de/en/working-time-accounts-in-the-public-and-private-sector-flexible-working-hours-are-becoming-increasingly-important/`, Accessed: 2024-06-14.

Elliott, C. and Soo, K. (2013). The international market for MBA qualifications: The relationship between tuition fees and applications. *Economics of Education Review*, 34:162–174.

Erdsiek, D. (2016). Overqualification of graduates: Assessing the role of family background. *Journal for Labour Market Research*, 49(3):253–268.

Eurofund (2024). Coding and classification standards. `https://www.eurofound.europa.eu/en/coding-and-classification-standards-0`, Accessed: 2025-06-25.

European Centre for the Development of Vocational Training (nd). Over-qualification rate (of tertiary graduates). `https://www.cedefop.europa.eu/en/tools/skills-intelligence/over-qualification-rate-tertiary-graduates?year=2022&country=DE#4`, Accessed: 2025-01-28.

European Commission (2023). Commission report finds labour and skills shortages persist and looks at possible ways to tackle them. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3704, Accessed: 2024-06-14.

Eurostat (2024). Government expenditure on education. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_on_education, Accessed: 2024-12-30.

Ezra, M. and Deckman, M. (1996). Balancing work and family responsibilities: Flextime and child care in the federal government. *Public Administration Review*, 56(2):174–179.

Fabra, E. and Camisó, C. (2009). Direct and indirect effects of education on job satisfaction: A structural equation model for the Spanish case. *Economics of Education Review*, 28(5):600–610.

Fack, G. and Grenet, J. (2015). Improving college access and success for low-income students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics*, 7(2):1–34.

Fang, T. and Hartley, J. (2022). Evolution of union wages and determinants. In Zimmermann, K. F., editor, *Handbook of Labor, Human Resources and Population Economics*. Springer-Verlag, Berlin and Heidelberg.

Federal Institute for Vocational Education and Training (2020). Young people study in the company and at school. https://www.bibb.de/en/77203.php, Accessed: 2025-05-19.

Feldman, D. (1996). The nature, antecedents and consequences of underemployment. *Journal of Management*, 22(3):385–407.

Fenizia, A., Checchi, D., and Lucifora, C. (2024). Public- and private-sector jobs: A cross-country perspective. *Oxford Economic Papers*, 76:759–779.

Ferrer-i-Carbonell, A. and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal*, 114(497):641–659.

Field, E. (2009). Educational debt burden and career choice: Evidence from a financial aid experiment at NYU law school. *American Economic Journal: Applied Economics*, 1(1):1–21.

Fischer, G.-B. and Wigger, B. (2016). Fiscal competition and higher education spending in Germany. *German Economic Review*, 17(2):234–252.

Fitzenberger, B., Ernst, M., and Haggeney, I. (1999). Wer ist noch Mitglied in Gewerkschaften? Eine Panelanalyse für Westdeutschland. *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 119(2):223–263.

Flavin, P., Pacek, A., and Radcliff, B. (2010). Labor unions and life satisfaction: Evidence from new data. *Social Indicators Research*, 98(3):435–449.

Flavin, P. and Shufeldt, G. (2016). Labor union membership and life satisfaction in the United States. *Labor Studies Journal*, 41(2):171–184.

Flèche, S., Lepinteur, A., and Powdthavee, N. (2021). The importance of capital in closing the entrepreneurial gender gap: A longitudinal study of lottery wins. *Journal of Economic Behavior & Organization*, 188:591–607.

Fleming, C. and Kler, P. (2008). I'm too clever for this job: A bivariate probit analysis on overeducation and job satisfaction in Australia. *Applied Economics*, 40:1123–1138.

Fleming, C. and Kler, P. (2014). Female overeducation, job satisfaction and the impact of children at home in Australia. *Economic Analysis and Policy*, 44(2):143–155.

Fletcher, J. and Kim, T. (2016). The effects of changes in kindergarten entry age policies on educational achievement. *Economics of Education Review*, 50:45–62.

Flisi, S., Goglio, V., Meroni, E., Rodrigues, M., and Vera-Toscano, E. (2017). Measuring occupational mismatch: Overeducation and overskill in Europe — evidence from PIAAC. *Social Indicators Research*, 131(3):1211–1249.

Frandsen, B. (2021). The surprising impacts of unionization: Evidence from matched employer-employee data. *Journal of Labor Economics*, 39(4):861–894.

Frank, R. (1978). Why women earn less: The theory and estimation of differential overqualification. *American Economic Review*, 68(3):360–373.

Fredriksson, P. and Öckert, B. (2014). Life-cycle effects of age at school start. *Economic Journal*, 124(579):977–1004.

Freeman, R. (1976). *The overeducated American*. Academic Press, New York.

Freeman, R. (1980). The exit-voice tradeoff in the labor market: Unionism, job tenure, quits, and separations. *Quarterly Journal of Economics*, 94(4):643–673.

Fricke, H. (2018). Tuition fees, student finances, and student achievement: Evidence from a differential raise in fees. *Journal of Human Capital*, 12(3):504–541.

Fumarco, L., Vandromme, A., Halewyck, L., Moens, E., and Baert, S. (2022). Does relative age affect speed and quality of transition from school to work? IZA Discussion Paper No. 14977.

Gallaway, L. (1969). Age and labor mobility patterns. *Southern Economic Journal*, 36(2):171–180.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Garibaldi, P., Giavazzi, F., Ichino, A., and Rettore, E. (2012). College cost and time to complete a degree: Evidence from tuition discontinuities. *Review of Economics and Statistics*, 94(3):699–711.

Gazioglu, S. and Tansel, A. (2006). Job satisfaction in Britain: Individual and job related factors. *Applied Economics*, 38(10):1163–1171.

Geißler, T. (2025). What an (un)favorable match: Public sector employment and the reversal of the overeducation-job satisfaction penalty. *Journal of Happiness Studies*, 26(94).

Geißler, T. and Goerke, L. (2025). Educational mismatch and trade union membership. *Industrial Relations*, 65(1):24–44.

Georgellis, Y., Sessions, J., and Tsitsianis, N. (2005). Windfalls, wealth, and the transition to self-employment. *Small Business Economics*, 25:407–428.

Gervais, M. and Ziebarth, N. (2019). Life after debt: Postgraduation consequences of federal student loans. *Economic Inquiry*, 57(3):1342–1366.

Giuliano, R., Mahy, B., Rycx, F., and Vermeylen, G. (2024). Overeducation, overskilling and job satisfaction in Europe: The moderating role of employment contracts. IZA Discussion Paper No. 16913.

Goebel, J., Grabka, M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2019). The German Socio-Economic Panel Study (SOEP). *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics*, 239(2):345–360.

Goerke, L. and Huang, Y. (2022). Job satisfaction and trade union membership in Germany. *Labour Economics*, 78:102238.

Goerke, L. and Pannenberg, M. (2004). Norm-based trade union membership: Evidence for Germany. *German Economic Review*, 5(4):481–504.

Goerke, L. and Pannenberg, M. (2011). Trade union membership and dismissals. *Labour Economics*, 18(6):810–821.

Golosov, M., Graber, M., Mogstad, M., and Novgorodsky, D. (2024). How Americans respond to idiosyncratic and exogenous changes in household wealth and unearned income. *Quarterly Journal of Economics*, 139(2):1321–1395.

Görlitz, K., Heß, P., and Tamm, M. (2025). Should states allow early school enrollment? An analysis of individuals' long-term labor market effects. *Empirical Economics*, 68:2383–2411.

Görlitz, K., Penny, M., and Tamm, M. (2022). The long-term effect of age at school entry on cognitive competencies in adulthood. *Journal of Economic Behavior & Organization*, 194:91–104.

Görlitz, K. and Tamm, M. (2020). Parenthood, risk attitudes and risky behavior. *Journal of Economic Psychology*, 79:102189.

Graham, B., Miller, M., and Strøm, K. (2017). Safeguarding democracy: Powersharing and democratic survival. *American Political Science Review*, 111(4):686–704.

Green, F. and McIntosh, S. (2007). Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics*, 39:427–439.

Green, F. and Tsitsianis, N. (2005). An investigation of national trends in job satisfaction in Britain and Germany. *British Journal of Industrial Relations*, 43(3):401–429.

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

Greifer, N. (2023). Package 'WeightIt'. Version: 0.14.2, `https://cran.r-project.org/web/packages/WeightIt/index.html`, Accessed: 2024-06-14.

Grönqvist, E., Öckert, B., and Vlachos, J. (2017). The intergenerational transmission of cognitive and noncognitive abilities. *Journal of Human Resources*, 52(4):887–918.

Groot, W. (1996). The incidence of, and returns to overeducation in the UK. *Applied Economics*, 28:1345–1350.

Groot, W. and Maassen van den Brink, H. (1999). Job satisfaction of older workers. *International Journal of Manpower*, 20(6):343–360.

Grunau, P. (2016). The impact of overeducated and undereducated workers on establishment-level productivity: First evidence for Germany. *International Journal of Manpower*, 37(2):372–392.

Grunau, P. and Pecoraro, M. (2017). Educational mismatch and promotions to managerial positions: A test of the career mobility theory. *Applied Economics*, 49(12):1226–1240.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

Hamilton, L. (2013). More is more or more is less? Parental financial investments during college. *American Sociological Review*, 78(1):70–95.

Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2):131–147.

Hartog, J., Ferrer-i-Carbonell, A., and Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos*, 55(1):3–26.

Hassani-Nezhad, L., Anderberg, D., Chevalier, A., Lührmann, M., and Pavan, R. (2021). Higher education financing and the educational aspirations of teenagers and their parents. *Economics of Education Review*, 85:102175.

Havranek, T., Irsova, Z., and Zeynalova, O. (2018). Tuition fees and university enrolment: A meta-regression analysis. *Oxford Bulletin of Economics and Statistics*, 80(6):1145–1184.

Haywood, L. (2016). Wealth effects on job preferences. *Labour Economics*, 38:1–11.

Headey, B. (2008). Life goals matter to happiness: A revision of set-point theory. *Social Indicators Research*, 86(2):213–231.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5(4):475–492.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Heijke, H., Meng, C., and Ris, C. (2003). Fitting to the job: The role of generic and vocational competencies in adjustment and performance. *Labour Economics*, 10(2):215–229.

Henley, A. (2004). House price shocks, windfall gains and hours of work: British evidence. *Oxford Bulletin of Economics and Statistics*, 66(4):439–456.

Herdeiro, R., Oliveira, A., and Menezes-Filho, N. (2025). The effect of age at school entry on educational and labor market outcomes in Brazil: A regression discontinuity analysis. *Economic Development and Cultural Change*. Forthcoming.

Hersch, J. (1991). Education match and job match. *Review of Economics and Statistics*, 73(1):140–144.

Ho, D., Imai, K., King, G., and Stuart, E. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8):1–28.

Hohendanner, C. and Kohaut, S. (2024). Tarifbindung und betriebliche Interessenvertretung: Ergebnisse aus dem IAB-Betriebspanel 2023. *WSI-Mitteilungen*, 77(4):289–295.

Holtz-Eakin, D., Joulfaian, D., and Rosen, H. (1993). The Carnegie conjecture: Some empirical evidence. *Quarterly Journal of Economics*, 108(2):413–435.

Horrace, W. and Oaxaca, R. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3):321–327.

Huang, C., Zhang, S., and Zhao, Q. (2020). The early bird catches the worm? School entry cutoff and the timing of births. *Journal of Development Economics*, 143:102386.

Huang, J. (2013). Intergenerational transmission of educational attainment: The role of household assets. *Economics of Education Review*, 33:112–123.

Hübner, M. (2012). Do tuition fees affect enrollment behavior? Evidence from a 'natural experiment' in Germany. *Economics of Education Review*, 31(6):949–960.

IG Metall (2012). Großer Erfolg für Auszubildende: Tarifergebnis unbefristete Übernahme. `https://www.igmetall.de/tarif/tarifrunden/metall-und-elektro/grosser-erfolg-fuer-auszubildende`, Accessed: 2025-01-29.

IG Metall (2018). Standortfaktor Fachkräfte: Herausforderungen und Handlungsfelder der Fachkräftesicherung. Arbeitspapier 3 2018 https://www.igmetall.de/download/20180112_Arbeitspapier_3_110118_Online_22e5fbdfdfdc665cce0687900fdf84634821b23b.pdf, Accessed: 2025-01-27.

Imbens, G., Rubin, D., and Sacerdote, B. (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review*, 91(4):778–794.

Institut für Arbeitsmarkt- und Berufsforschung (2024). IAB-Stellenerhebung/ IAB Job Vacancy Survey. https://iab.de/das-iab/befragungen/iab-stellenerhebung/, Accessed: 2025-05-15.

International Labour Office (2012). International standard classification of occupations: Structure, group definitions and correspondence tables. https://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/@publ/documents/publication/wcms_172572.pdf, Accessed: 2024-06-24.

Iseke, A. (2014). The part-time job satisfaction puzzle: Different types of job discrepancies and the moderating effect of family importance. *British Journal of Industrial Relations*, 52(3):445–469.

Ivlevs, A. and Veliziotis, M. (2017). What do unions do in times of economic crisis? Evidence from Central and Eastern Europe. *European Journal of Industrial Relations*, 23(1):81–96.

Jacobs, V., Mahy, B., Rycx, F., and Volral, M. (2021). Over-education among immigrants: The role of demographics, time, and firm characteristics. *Applied Economics*, 53(1):61–78.

Jacobs, V., Pineda-Hernández, K., Rycx, F., and Volral, M. (2023). Does over-education raise productivity and wages equally? The moderating role of workers' origin and immigrants' background. *Education Economics*, 31(6):698–724.

Jäger, S., Noy, S., and Schoefer, B. (2022). The German model of industrial relations: Balancing flexibility and collective action. *Journal of Economic Perspectives*, 36(4):53–80.

Jones, M., Mavromaras, K., Sloane, P., and Wei, Z. (2014). Disability, job mismatch, earnings and job satisfaction in Australia. *Cambridge Journal of Economics*, 38(5):1221–1246.

Joona, P., Gupta, N., and Wadensjö, E. (2014). Overeducation among immigrants in Sweden: Incidence, wage effects and state dependence. *Journal of Migration*, 3(9):1–23.

Joulfaian, D. and Wilhelm, M. (1994). Inheritance and labor supply. *Journal of Human Resources*, 29(4):1205–1234.

Kaiser, C. and Vendrik, M. (2022). How much can we learn from happiness data? https://raw.githubusercontent.com/CasparFriedrichKaiser/docs/main/happytransformations_v220622.pdf, Accessed: 2024-06-14.

Kalenkoski, C. and Pabilonia, S. (2010). Parental transfers, student achievement, and the labor supply of college students. *Journal of Population Economics*, 23:469–496.

Kampelmann, S. and Rycx, F. (2012). The impact of educational mismatch on firm productivity: Evidence from linked panel data. *Economics of Education Review*, 31(6):918–931.

Kane, T. (2007). Evaluating the impact of the DC tuition assistance grant program. *Journal of Human Resources*, 42(3):555–582.

Keane, L., Pacek, A., and Radcliff, B. (2012). Organized labor, democracy, and life satisfaction: A cross-national analysis. *Labor Studies Journal*, 37(3):253–270.

Khalil, S., Lietz, A., and Mayer, S. (2023). Overeducation as a moderator for the link between job change and job satisfaction among immigrants and natives in Germany. *Current Psychology*, 42:26692–26708.

Kiersztyn, A. (2013). Stuck in a mismatch? The persistence of overeducation during twenty years of the post-communist transition in Poland. *Economics of Education Review*, 32:78–91.

Kiker, B., Santos, M., and De Oliveira, M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2):111–125.

Kim, J. and Wiggins, M. (2011). Family-friendly human resource policy: Is it still working in the public sector? *Public Administration Review*, 71(5):728–739.

Kim, S. (2012). Does person-organization fit matter in the public sector? Testing the mediating effect of person-organization fit in the relationship between public service motivation and work attitudes. *Public Administration Review*, 72(6):830–840.

Kim, T. (2021). Age culture, school-entry cutoff, and the choices of birth month and school-entry timing in South Korea. *Journal of Demographic Economics*, 87(1):33–65.

Kindermann, F., Mayr, L., and Sachs, D. (2020). Inheritance taxation and wealth effects on the labor supply of heirs. *Journal of Public Economics*, 191:104127.

King, W. and Hautaluoma, J. (1987). Comparison of job satisfaction, life satisfaction, and performance of overeducated and other workers. *Journal of Social Psychology*, 127(5):421–433.

Ko, J. and Hur, S. (2014). The impacts of employee benefits, procedural justice, and managerial trustworthiness on work attitudes: Integrated understanding based on social exchange theory. *Public Administration Review*, 74(2):176–187.

Korpi, T. and Tåhlin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, 16(2):183–193.

Kristof, A. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49(1):1–49.

Landersø, R., Nielsen, H., and Simonsen, M. (2017). School starting age and the crime-age profile. *Economic Journal*, 127(602):1096–1118.

Larreguy, H. and Marshall, J. (2017). The effect of education on civic and political engagement in nonconsolidated democracies: Evidence from Nigeria. *Review of Economics and Statistics*, 99(3):387–401.

Lazear, E., Shaw, K., and Stanton, C. (2018). Who gets hired? The importance of competition among applicants. *Journal of Labor Economics*, 36(S1):133–181.

Lee, S. and Ohtake, F. (2021). How conscious are you of others? Further evidence on relative income and happiness. *Journal of Happiness Studies*, 22(8):3321–3356.

Lepinteur, A. (2019). The shorter workweek and worker wellbeing: Evidence from Portugal and France. *Labour Economics*, 58:204–220.

Lewis, G. and Frank, S. (2002). Who wants to work for the government? *Public Administration Review*, 62(4):395–404.

Li, X., Lou, X., and Zhang, J. (2022). Does the early bird catch the worm? The effect of school starting age on educational attainment and labor market outcomes:

Evidence from Chinese urban twins data. *Journal of Comparative Economics*, 50(3):832–848.

Lillo-Bañuls, A. and Casado-Díaz, J. (2015). Exploring the relationship between educational mismatch, earnings and job satisfaction in the tourism industry. *Current Issues in Tourism*, 18(4):361–375.

Lindh, T. and Ohlsson, H. (1996). Self-employment and windfall gains: Evidence from the Swedish lottery. *Economic Journal*, 106(439):1515–1526.

Liu, Y., Yin, L., and Guo, J. (2021). The quality of higher education and overeducation: Where should higher education funding go? *Finance Research Letters*, 41:101824.

Luo, W. (2022). Inequality and growth in the twenty-first century. *Scottish Journal of Political Economy*, 69(4):345–366.

Maczulskij, T. (2017). Who becomes a public sector employee? *International Journal of Manpower*, 38(4):567–579.

Malo, M. and Sciulli, D. (2021). Wealth transfers and labour supply: Impact of inheritances and gifts by gender in Europe. *International Journal of Manpower*, 42(8):1450–1478.

Martin, A. (2009). Age appropriateness and motivation, engagement, and performance in high school: Effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology*, 101(1):101–114.

Marx, B. and Turner, L. (2019). Student loan nudges: Experimental evidence on borrowing and educational attainment. *American Economic Journal: Economic Policy*, 11(2):108–141.

Matta, R., Ribas, R., Sampaio, B., and Sampaio, G. (2016). The effect of age at school entry on college admission and earnings: A regression-discontinuity approach. *IZA Journal of Labor Economics*, 5(9):1–25.

Mavromaras, K., McGuinness, S., and Fok, Y. (2009). Assessing the incidence and wage effects of overskilling in the Australian labour market. *Economic Record*, 85(268):60–72.

Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P., and Wei, Z. (2013). Job mismatches and labour market outcomes: Panel evidence on university graduates. *Economic Record*, 89(286):382–395.

Mavromaras, K., Sloane, P., and Wei, Z. (2012). The role of education pathways in the relationship between job mismatch, wages and job satisfaction: A panel estimation approach. *Education Economics*, 20(3):303–321.

McCloud, T. and Bann, D. (2019). Financial stress and mental health among higher education students in the UK up to 2018: Rapid review of evidence. *Journal of Epidemiology & Community Health*, 73(10):977–984.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

McEwan, P. and Shapiro, J. (2008). The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates. *Journal of Human Resources*, 43(1):1–29.

McGoldrick, K. and Robst, J. (1996). Gender differences in overeducation: A test of the theory of differential overqualification. *American Economic Review*, 86(2):280–284.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3):387–418.

McGuinness, S. and Byrne, D. (2015). Born abroad and educated here: Examining the impacts of education and skill mismatch among immigrant graduates in Europe. *IZA Journal of Migration*, 4(17):1–30.

McGuinness, S., Pouliakas, K., and Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, 32(4):985–1015.

McGuinness, S. and Sloane, P. (2011). Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review*, 30(1):130–145.

McKee-Ryan, F. and Harvey, J. (2011). "I have a job, but...": A review of underemployment. *Journal of Management*, 37(4):962–996.

McKee-Ryan, F., Virick, M., Prussia, G., Harvey, J., and Lilly, J. (2009). Life after the layoff: Getting a job worth keeping. *Journal of Organizational Behavior*, 30:561–580.

McNall, L., Masuda, A., and Nicklin, J. (2009). Flexible work arrangements, job satisfaction, and turnover intentions: The mediating role of work-to-family enrichment. *Journal of Psychology*, 144(1):61–81.

McPherson, M. and Schapiro, M. (1991). Does student aid affect college enrollment? New evidence on a persistent controversy. *American Economic Review*, 81(1):309–318.

McVicar, D. and McKee, B. (2002). Part–time work during post–compulsory education and examination performance: Help or hindrance? *Scottish Journal of Political Economy*, 49(4):393–406.

Minicozzi, A. (2005). The short term effect of educational debt on job decisions. *Economics of Education Review*, 24(4):417–430.

Minor, R. (2023). How tuition fees affected student enrollment at higher education institutions: The aftermath of a German quasi-experiment. *Journal for Labour Market Research*, 57(28):1–19.

Molina, T. and Rivadeneyra, I. (2021). The schooling and labor market effects of eliminating university tuition in Ecuador. *Journal of Public Economics*, 196:104383.

Monks, J. (2001). Loan burdens and educational outcomes. *Economics of Education Review*, 20(6):545–550.

Montalbán, J. (2023). Countering moral hazard in higher education: The role of performance incentives in need-based grants. *Economic Journal*, 133(649):355–389.

Montmarquette, C., Viennot-Briot, N., and Dagenais, M. (2007). Dropout, school performance, and working while in school. *Review of Economics and Statistics*, 89(4):752–760.

Mühlenweg, A. (2010). Young and innocent: International evidence on age effects within grades on victimization in elementary school. *Economics Letters*, 109(3):157–160.

Mühlenweg, A., Blomeyer, D., Stichnoth, H., and Laucht, M. (2012). Effects of age at school entry (ASE) on the development of non-cognitive skills: Evidence from psychometric data. *Economics of Education Review*, 31(3):68–76.

Mühlenweg, A. and Puhani, P. (2010). The evolution of the school-entry age effect in a school tracking system. *Journal of Human Resources*, 45(2):407–438.

Neill, C. (2009). Tuition fees and the demand for university places. *Economics of Education Review*, 28:561–570.

Neill, C. (2015). Rising student employment: The role of tuition fees. *Education Economics*, 23(1):101–121.

Nunley, J., Pugh, A., Romero, N., and Seals, R. (2017). The effects of unemployment and underemployment on employment opportunities: Results from a correspondence audit of the labor market for college graduates. *ILR Review*, 70(3):642–669.

OECD (2022a). Adult education level. https://www.oecd.org/en/data/indicators/adult-education-level.html, Accessed: 2025-03-25.

OECD (2022b). Mismatch by country. https://stats.oecd.org/Index.aspx?DataSetCode=S4J2022_MISMATCH, Accessed: 2024-06-24.

OECD (2022c). Skills for jobs 2022. https://www.oecdskillsforjobsdatabase.org/data/S4J2022_results.pdf, Accessed: 2025-01-09.

OECD (2024). Education at a glance 2024: OECD indicators. OECD Publishing, Paris, https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/09/education-at-a-glance-2024_5ea68448/c00cad36-en.pdf, Accessed: 2025-04-30.

OECD (nd). Population with tertiary education. https://www.oecd.org/en/data/indicators/population-with-tertiary-education.html?oecdcontrol-00b22b2429-var3=1998, Accessed: 2025-05-19.

OECD and AIAS (2023). Institutional characteristics of trade unions, wage setting, state intervention and social pacts. Version 1.1, https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/collective-bargaining-and-social-dialogue/oecd-aias-ictwss_v1.1.xlsx, Accessed: 2024-06-24.

Oettinger, G. (1999). Does high school employment affect high school academic performance? *ILR Review*, 53(1):136–151.

Oosterbeek, H., ter Meulen, S., and van Der Klaauw, B. (2021). Long-term effects of school-starting-age rules. *Economics of Education Review*, 84:102144.

Ordine, P. and Rose, G. (2009). Overeducation and instructional quality: A theoretical model and some facts. *Journal of Human Capital*, 3(1):73–105.

Ordine, P. and Rose, G. (2011). Inefficient self-selection into education and wage inequality. *Economics of Education Review*, 30(4):582–597.

Ordine, P. and Rose, G. (2017). Individual mismatch and aggregate overeducation: Evidence from a quasi-natural experiment. *Italian Economic Journal*, 3:1–23.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Page, L., Kehoe, S., Castleman, B., and Sahadewo, G. (2019). More than dollars for scholars: The impact of the Dell Scholars Program on college access, persistence, and degree attainment. *Journal of Human Resources*, 54(3):683–725.

Peiró, J., Agut, S., and Grau, R. (2010). The relationship between overeducation and job satisfaction among young Spanish workers: The role of salary, contract of employment, and work experience. *Journal of Applied Social Psychology*, 40(3):666–689.

Perry, J., Hondeghem, A., and Wise, L. (2010). Revisiting the motivational bases of public service: Twenty years of research and an agenda for the future. *Public Administration Review*, 70(5):681–690.

Perry, J. and Wise, L. (1990). The motivational bases of public service. *Public Administration Review*, 50(3):367–373.

Pfeifer, C. (2010). Risk aversion and sorting into public sector employment. *German Economic Review*, 12(1):85–99.

Pierse, T. and McHale, J. (2015). Unions and involuntary job separations. *Human Resource Management Journal*, 25(4):496–515.

Plesca, M. and Summerfield, F. (2023). The productivity benefits of overeducation. *Journal of Human Capital*, 17(4):463–502.

Ponzo, M. and Scoppa, V. (2014). The long-lasting effects of school entry age: Evidence from Italian students. *Journal of Policy Modeling*, 36(3):578–599.

Puhani, P. and Weber, A. (2008). Does the early bird catch the worm? In: Dustmann, C. and Fitzenberger, B. and Machin, S., editors. *The Economics of Education and Training*. Physica-Verlag, Heidelberg.

Radcliff, B. (2005). Class organization and subjective well-being: A cross-national analysis. *Social Forces*, 84(1):513–530.

Richter, D., Rohrer, J., Metzing, M., Nestler, W., Weinhardt, M., and Schull, J. (2017). SOEP scales manual. SOEP survey papers 423: Series C – Data documentation. https://www.diw.de/documents/publikationen/73/diw_01.c.571151.de/diw_ssp0423.pdf, Accessed: 2024-06-14.

Riphahn, R. (2012). Effect of secondary school fees on educational attainment. *Scandinavian Journal of Economics*, 114(1):148–176.

Riphahn, R. and Trübswetter, P. (2013). The intergenerational transmission of education and equality of educational opportunity in East and West Germany. *Applied Economics*, 45(22):3183–3196.

Robst, J. (1995a). Career mobility, job match, and overeducation. *Eastern Economic Journal*, 21(4):539–550.

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

Robst, J. (2007). Education, college major, and job match: Gender differences in reasons for mismatch. *Education Economics*, 15(2):159–175.

Roller, C., Rulff, C., and Tamminga, M. (2020). It's a mismatch! Overeducation and career mobility in Germany. *German Economic Review*, 21(4):493–514.

Romanov, D., Tur-Sinai, A., and Eizman, G. (2017). Overeducation, job mobility and earnings mobility among holders of first degrees. *Applied Economics*, 49(26):2563–2578.

Rosen, S. (1986). The theory of equalizing differences. In: Ashenfelter, O. and Card, D., editors. *Handbook of Labor Economics*, volume 1. North Holland, Amsterdam, Netherlands.

Ross, C. and Wu, C.-l. (1995). The links between education and health. *American Sociological Review*, 60(5):719–745.

Roth, J., Sant'Anna, P., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244.

Rothstein, J. and Rouse, C. (2011). Constrained after college: Student loans and early-career occupational choices. *Journal of Public Economics*, 95(1-2):149–163.

Rumberger, R. (1981a). The rising incidence of overeducation in the US labor market. *Economics of Education Review*, 1(3):293–314.

Rumberger, R. (1981b). The changing skill requirements of jobs in the US economy. *ILR Review*, 34(4):578–590.

Rumberger, R. (1987). The impact of surplus schooling on productivity and earnings. *Journal of Human Resources*, 22(1):24–50.

Saltzstein, A., Ting, Y., and Saltzstein, G. (2001). Work-family balance and job satisfaction: The impact of family-friendly policies on attitudes of federal government employees. *Public Administration Review*, 61(4):452–467.

Sam, V. (2020). Impacts of educational mismatches on job satisfaction: The case of university graduates in Cambodia. *International Journal of Manpower*, 41(1):84–99.

Santiago-Vela, A. and Mergener, A. (2022). Gender overeducation gap in the digital age: Can spatial flexibility through working from home close the gap? *Social Science Research*, 106:102727.

Sattinger, M. and Hartog, J. (2013). Nash bargaining and the wage consequences of educational mismatches. *Labour Economics*, 23:50–56.

Sauer, R. (2004). Educational financing and lifetime earnings. *Review of Economic Studies*, 71(4):1189–1216.

Savoca, E. (1990). Another look at the demand for higher education: Measuring the price sensitivity of the decision to apply to college. *Economics of Education Review*, 9(2):123–134.

Schäfer, D., Talavera, O., and Weir, C. (2011). Entrepreneurship, windfall gains and financial constraints: Evidence from Germany. *Economic Modelling*, 28(5):2174–2180.

Schaub, S. (2025). robomit: Robustness checks for omitted variable bias. Version: 1.0.7, `https://cran.r-project.org/package=robomit`, Accessed: 2025-06-06.

Schmidt, C. and Zimmermann, K. (1991). Work characteristics, firm size and wages. *Review of Economics and Statistics*, 73(4):705–710.

Sharma, S. and Sharma, P. (2017). Educational mismatch and its impact on earnings: Evidence from Indian labour market. *International Journal of Social Economics*, 44(12):1778–1795.

Shin, Y. (2023). School starting age policy and students' risky health behaviors. *Health Economics*, 32(11):2446–2459.

Sicherman, N. (1991). "Overeducation" in the labor market. *Journal of Labor Economics*, 9(2):101–122.

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

Sikhova, A. (2023). Understanding the effect of parental education and financial resources on the intergenerational transmission of income. *Journal of Labor Economics*, 41(3):771–811.

Sila, U. and Sousa, R. (2014). Windfall gains and labour supply: Evidence from the European household panel. *IZA Journal of Labor Economics*, 3:1–27.

Sloane, P., Battu, H., and Seaman, P. (1996). Overeducation and the formal education/experience and training trade-off. *Applied Economics Letters*, 3(8):511–515.

Sloane, P., Battu, H., and Seaman, P. (1999). Overeducation, undereducation and the British labour market. *Applied Economics*, 31:1437–1453.

Smith, A. (1776). *The Wealth of Nations*. The Modern Library, New York.

Sohn, K. (2010). The role of cognitive and noncognitive skills in overeducation. *Journal of Labor Research*, 31:124–145.

Solis, A. (2017). Credit access and college enrollment. *Journal of Political Economy*, 125(2):562–622.

Song, Y. and Gao, J. (2020). Does telework stress employees out? A study on working at home and subjective well-being for wage/salary workers. *Journal of Happiness Studies*, 21(7):2649–2668.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3):355–374.

Staiger, D. and Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586.

Statistisches Bundesamt (2020a). 2018 war mehr als jeder zehnte Erwerbstätige in Deutschland im öffentlichen Dienst beschäftigt. Pressemitteilung Nr. N 021 vom 29. April 2020. `https://www.destatis.de/DE/Presse/Pressemitteilungen/2020/04/PD20_N021_742.html`, Accessed: 2024-06-14.

Statistisches Bundesamt (2020b). Bevölkerung im Alter von 15 Jahren und mehr nach allgemeinen und beruflichen Bildungsabschlüssen nach Jahren. `https:`

//www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Bildungsstand/Tabellen/bildungsabschluss.html?nn=209544, Accessed: 2025-04-25.

Statistisches Bundesamt (2022a). Berufsbildungsstatistik: Bestandene Abschlussprüfung. https://www-genesis.destatis.de/genesis//online?operation=table&code=21211-0009&bypass=true&levelindex=0&levelid=1717590507114#abreadcrumb, Accessed: 2024-01-18.

Statistisches Bundesamt (2022b). Prüfungen an Hochschulen. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/pruefungen-hochschulen-2110420217005.xlsx?__blob=publicationFile, Accessed: 2025-07-02.

Statistisches Bundesamt (2023a). Öffentlicher Dienst: Beschäftigte nach Geschlecht und der Art des Dienst- oder Arbeitsverhältnisses, Stichtag 30. Juni. https://www.destatis.de/DE/Themen/Staat/Oeffentlicher-Dienst/Tabellen/beschaeftigte-geschlecht.html, Accessed: 2024-03-05.

Statistisches Bundesamt (2023b). Statistischer Bericht - Statistik der Prüfungen - Prüfungsjahr 2022. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/statistischer-bericht-pruefungen-2110420227005.xlsx?__blob=publicationFile, Accessed: 2024-6-24.

Statistisches Bundesamt (2024a). Bevölkerung nach Nationalität und Geschlecht 1970 bis 2023 in Deutschland. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/deutsche-nichtdeutsche-bevoelkerung-nach-geschlecht-deutschland.html, Accessed: 2025-05-15.

Statistisches Bundesamt (2024b). Studienanfänger: Bundesländer, Semester, Nationalität, Geschlecht. Table 21311-0014, https://www-genesis.destatis.de/genesis//online?operation=table&code=21311-0014&bypass=true&levelindex=1&levelid=1729073317940#abreadcrumb, Accessed: 2024-10-16.

Statistisches Bundesamt (2024c). Studierende: Bundesländer, Semester, Nationalität, Geschlecht. Table 21311-0005, https://www-genesis.destatis.de/datenbank/online/table/21311-0005/search/s/MjEzMTEtMDAwNQ==, Accessed: 2024-12-13.

Statistisches Bundesamt (2024d). Zusammengefasste Abschlussprüfungen nach Geschlecht, Nationalität und Durchschnittsalter. `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Tabellen/bestandenepruefungen-gruppen.html`, Accessed: 2024-11-20.

Statistisches Bundesamt (nd). Unter- und Überqualifikation. `https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Qualitaet-Arbeit/Dimension-6/unter-ueberqualifikationl.html`, Accessed: 2024-06-14.

Steijn, B. (2008). Person-environment fit and public service motivation. *International Public Management Journal*, 11(1):13–27.

Stinebrickner, R. and Stinebrickner, T. (2003). Working during school and academic performance. *Journal of Labor Economics*, 21(2):473–491.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Suziedelyte, A. and Zhu, A. (2015). Does early schooling narrow outcome gaps for advantaged and disadvantaged children? *Economics of Education Review*, 45:76–88.

Sánchez-Sánchez, N. and McGuinness, S. (2015). Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: An analysis using REFLEX data. *Education Economics*, 23(4):419–432.

Tarvid, A. (2015). The role of industry in the prevalence of overeducation in Europe. *Procedia Economics and Finance*, 30:876–884.

Taylor, M. (2001). Self–employment and windfall gains in Britain: Evidence from panel data. *Economica*, 68(272):539–565.

Thurow, L. (1975). *Generating inequality: Mechanisms of distribution in the U.S. economy*. Basic books, New York.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1):24–36.

Tonin, M. and Vlassopoulos, M. (2015). Are public sector workers different? Cross-European evidence from elderly workers and retirees. *IZA Journal of Labor Economics*, 4(11):1–21.

Tsang, M. (1987). The impact of underutilization of education on productivity: A case study of the U.S. Bell companies. *Economics of Education Review*, 6(3):239–254.

Tsang, M. and Levin, H. (1985). The economics of overeducation. *Economics of Education Review*, 4(2):93–104.

Turmo-Garuz, J. and Bartual-Figueras, M. (2019). Factors associated with overeducation among recent graduates during labour market integration: The case of Catalonia (Spain). *Social Indicators Research*, 144:1273–1301.

UNESCO (2006). International Standard Classification of Education (ISCED 1997). http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf, Accessed: 2023-01-25.

United Nations Development Programme (2024). Human development report 2023-2024. https://ourworldindata.org/grapher/average-years-of-schooling?tab=chart&country=~OWID_WRL, Accessed: 2025-01-07.

Vera-Toscano, E. and Meroni, E. (2021). An age-period-cohort approach to the incidence and evolution of overeducation and skills mismatch. *Social Indicators Research*, 153:711–740.

Verdugo, R. and Verdugo, N. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, 24(4):629–643.

Vereinte Dienstleistungsgewerkschaft (2014). Stellungnahme zum Entwurf eines Gesetzes zur Weiterentwicklung des Personalrechts der Beamtinnen und Beamten der früheren Deutschen Bundespost (PostPersWG-E). https://www.bundesfinanzministerium.de/Content/DE/Gesetzestexte/Gesetze_Gesetzesvorhaben/Abteilungen/Abteilung_V/18_Legislaturperiode/2015-06-05-Personalrecht-beamte-deutsche-bundespost/Stellungnahme-02-verdi.pdf?__blob=publicationFile&v=2, Accessed: 2025-01-27.

Verhaest, D. and Omey, E. (2006). The impact of overeducation and its measurement. *Social Indicators Research*, 77(3):419–448.

Verhaest, D. and Omey, E. (2009). Objective over-education and worker well-being: A shadow price approach. *Journal of Economic Psychology*, 30(3):469–481.

Verhaest, D. and Omey, E. (2010). The determinants of overeducation: Different measures, different outcomes? *International Journal of Manpower*, 31(6):608–625.

Verhaest, D. and Omey, E. (2012). Overeducation, undereducation and earnings: Further evidence on the importance of ability and measurement error bias. *Journal of Labor Research*, 33(1):76–90.

Verhaest, D. and Verhofstadt, E. (2016). Overeducation and job satisfaction: The role of job demands and control. *International Journal of Manpower*, 37(3):456–473.

Vila, L. and García-Mora, B. (2005). Education and the determinants of job satisfaction. *Education Economics*, 13(4):409–425.

Voces, C. and Caínzos, M. (2021). Overeducation as status inconsistency: Effects on job satisfaction, subjective well-being and the image of social stratification. *Social Indicators Research*, 153:979–1010.

Wald, S. and Fang, T. (2008). Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey. *Canadian Public Policy/ Analyse de Politiques*, 34(4):457–479.

Walsemann, K., Gee, G., and Gentile, D. (2015). Sick of our loans: Student borrowing and the mental health of young adults in the United States. *Social Science & Medicine*, 124:85–93.

Wen, L. and Maani, S. (2019). Job mismatches and career mobility. *Applied Economics*, 51(10):1010–1024.

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data*. The MIT Press, Cambridge, second edition.

World Bank Group (2024). World development indicators. `https://databank.worldbank.org/source/world-development-indicators/Series/SE.XPD.TOTL.GD.ZS`, Accessed: 2024-12-30.

Zhang, L. (2013). Effects of college educational debt on graduate school attendance and early career and lifestyle choices. *Education Economics*, 21(2):154–175.

Curriculum vitae

Education

2022-2025	Ph.D. in Economics and Business, Trier University
2019-2021	M.Sc. Business Administration, Trier University
2018	Erasmus exchange, HZ University of Applied Sciences (NL)
2016-2019	B.A. International Business, Cooperative State University Baden-Württemberg

Professional experience

2022-2025	Research assistant at Trier University & IAAEU
2021-2022	Project coordinator at Trier University
2020-2021	Student assistant at Trier University & IAAEU
2020	Intern at IT-Haus GmbH
2016-2019	Dual student at ALDI GmbH & Co. KG Wittlich

Honors and awards

2023	INFER PhD Paper Award
2021	Germany scholarship

Teaching

2025	Master research seminar
2023-2024	Master thesis supervision
2023	Master research seminar
	Colloquium for Bachelor theses
2022-2023	Colloquium for Bachelor theses
2022	Lecture in human resources management (VWA)
	Lecture in personnel management (ZFUW)

Publications in refereed journals

Geißler, T. What an (un)favourable match: Public sector employment and the reversal of the overeducation-job satisfaction penalty. *Journal of Happiness Studies, forthcoming*.

Geißler, T. & Goerke, L. Educational mismatch and trade union membership. *Industrial Relations, forthcoming*.