Essays on the Economics of Education

Inaugural dissertation

submitted by Stefanie Hof

in fulfillment of the requirements for the degree of

Doctor rerum oeconomicarum

at the Faculty of Business, Economics and Social Sciences

of the

University of Bern

Submitted by

STEFANIE HOF

from Zwingen (BL), Switzerland

2016

The faculty accepted this work as dissertation on 26th May 2016 at the request of the two advisors Prof. Dr. Stefan C. Wolter and Prof. Dr. Guido Schwerdt, without wishing to take a position on the view presented therein.

Contents

List of Figures

List of Tables

Acknowledgments

It would not have been possible to write this thesis without the help and support of the kind people around me. First and foremost, I would like to thank Prof. Dr. Stefan C. Wolter for the opportunity to write a PhD thesis under his supervision and for his advice, continuous support, interesting discussions and constructive feedback. I also thank Prof. Dr. Guido Schwerdt for being my co-advisor.

A special thanks goes to my co-authors Thomas Bolli, Christine Sälzer and Mirjam Strupler Leiser. Their knowledge and skills, as well as their guidance and eagerness to discuss my thesis, were invaluable to me. I also highly appreciate the comments I received in a number of conferences and seminars.

I further appreciate the great support I received from the Swiss Coordination Center for Research in Education, and in particular, I appreciate the possibility to attend the PhD course program in "Economics of Education" organized by Prof. Dr. Uschi Backes-Gellner and Prof. Dr. Stefan C. Wolter.

Finally, a huge thanks goes to my family and friends. In particular, I am very grateful to my husband, Alexis Seiler, for his personal support, patience and encouragement.

Chapter 1

Introduction

This dissertation consists of four chapters analyzing various empirical aspects of the economics of education. All chapters are based on different types of data and use different empirical strategies to answer the research questions. Within the first three chapters addressing student outcomes, the first two chapters use academic achievement as the outcome variable. While chapter 2 analyzes the impact of private tutoring based on the Swiss PISA data from 2009, chapter 3 analyzes the impact of truancy based on a national enhancement of the German PISA 2012 data. However, in both chapters, a non-parametric bounds method is applied, though each has a different focus and application of the method. Chapter 4 assesses the impact of work-based education (apprenticeship) compared to school-based education personality skills using a Swiss longitudinal study (TREE). This chapter applies different methods to identify the causal impact but is essentially based on a lagged dependent variable (LDV) approach and an instrument variable (IV) approach. Chapter 5 discusses job change into teaching in vocational education and training (VET) using a representative dataset of VET teachers in Switzerland. The empirical strategy combines matching and regression analyses.

1.1 Thesis Summary

Chapter 2 presents an alternative framework – a nonparametric bounds method – to analyze the effect of private tutoring on students' academic outcomes. The presented examination uses, for the first time, a large representative data set in a European setting to identify the causal effect of self-initiated private tutoring. Under relatively weak assumptions, the chapter suggests some evidence that private tutoring improves students' outcomes in reading. However, the results indicate a heterogeneous and nonlinear effect of private tutoring; i.e., a threshold may exist after which private tutoring becomes ineffective or even detrimental.

Chapter 3[1] investigates the causal impact of truancy on students' academic achievement in mathematics, addressing both the identification issues of non-random selection and measurement error regarding possible misreporting of truancy using self-reports. Applying a nonparametric

[1] This chapter was co-authored with Christine Sälzer.

approach instead of point estimates, as in chapter 2, this chapter uses bounds that are more reliable. The imposed assumptions concerning the nature of selection and measurement error are weak. The main finding is that regardless of measurement error, the relationship between truant behavior and student achievement cannot be assigned for the full sample. However, the impact of misreporting is profound. If as little as two percent of the sample misreport their truancy behavior, the direction of the average treatment effect (ATE), even under the condition of exogenous selection, is no longer negative.

Chapter 4[2] extends the analysis on the impact of education by the outcome "personality skills". This chapter analyzes how work-based upper secondary education affects personality skills in comparison to school-based upper secondary education. The identification strategy accounts for selection into education tracks by analyzing growth of personality skills over time by a lagged dependent variable (LDV) approach. In addition, the identification strategy indicates exploitation of the growth of governmentally defined differences in the relative weight of school- and work-based education across regions to tackle unobserved heterogeneity in personality skill growth. The results suggest that work-based upper secondary education permanently decreases emotion-centered coping.

Chapter 5[3] differs in its focus. While the first three chapters analyze students and their academic or personal outcomes, this chapter analyzes individuals who change their job to become a teacher in vocational education and training (VET) as a second career. Using a unique data set of career changers in teaching in a vocational subject, the applied matching method shows that individuals who change their careers to teaching in VET earned, on average, more in their first career than comparable workers in the same occupation. The results suggest that educational systems can attract highly qualified individuals as career changers into teaching. The findings also demonstrate that the average career changer still expects to earn significantly more as a teacher than in their former career. However, the study shows that one-third of the career changers expect a wage loss.

1.2 Data and Methodology

Using a representative dataset for Switzerland, chapter 2 analyzes the question of whether self-initiated private tutoring has a causal effect on students' academic achievement in mathematics and reading. The Program for International Student Assessment (PISA) 2009 data collection for Switzerland includes a large nationally representative sample of 15-year-old students and a supplementary study of grade-9 students from a selection of cantons. These surveys include a

[2] This chapter was co-authored with Thomas Bolli.

[3] This chapter was co-authored with Mirjam Strupler Leiser.

national option on the demand of private tutoring. This chapter applies an alternative method to overcome the selection bias and to identify the effect of self-initiated private tutoring on students' outcomes. Applying this nonparametric bounds method allows us to analyze the causal effect of private tutoring by relying on a set of relatively weak nonparametric assumptions. In particular, the method drops the likely unrealistic assumption of a linear and homogeneous effect; i.e., it allows for the effect of private tutoring to vary with the quality of the private tutoring. The chapter starts by making no assumptions and then successively imposes weak nonparametric assumptions to tighten the bounds step by step. First, a monotone treatment selection (MTS) assumption is imposed, which states that attending private tutoring classes is weakly monotonically related with poor academic outcomes. Second, it uses the parents' education as a monotone instrument variable (MIV). Third, this study applies the monotone treatment response (MTR), which means the effect of private tutoring is not negative.

Chapter 3 exploits a German national grade-nine oversample of PISA 2012 to analyze the impact of student self-reported truancy on academic achievement. Hence, as in chapter 2, the applied method is a nonparametric bounds method accounting for the problem of self-selection into the truancy status. Chapter 3 focuses on a potential measurement error because there is reason to assume that truancy is sometimes misreported in terms of exaggerated frequencies or denied behavior. In other words, students' responses to questionnaire items may involve both false positive and false negative classifications of students as truants or non-truants. In addition to using the technique of starting by making no assumptions and successively imposing weak nonparametric assumptions to tighten the bounds, this chapter allows for different misclassifications of truancy.

Chapter 4 exploits a dataset that follows the participants of the 2000 Swiss PISA examination at grade 9 up to the age of approximately 25. The Transition to Education and Employment (TREE) survey is administered each year between 2001 and 2007 and in 2010. The sample is representative of both the country as a whole and its three main language regions (German, French, and Italian). This unique database combines the variables in the standard PISA survey, such as parental background, PISA test scores and living conditions, with information on personality skills and employment/education status. This chapter aims to provide evidence on the causal effect of work-based secondary education compared to school-based secondary education on personality skills. As in the first two chapters, the analysis is based on a grade 9 student population. However, the data follow the students until the age of 25, and the applied methodology differs compared to that used in the first two chapters. To address the concerns regarding endogeneity due to selection and unobserved heterogeneity, three different strategies are applied. First, the chapter makes use of the panel structure of the data set to analyze changes over time, including the lagged dependent variable on the right-hand side of the equation accounting for selection in terms of the personality skill level. Second, an instrumental-variable approach is applied that exploits regional differences

in the relevance of general secondary education across Switzerland to account for potential selection in personality skill growth. We adresse potential endogeneity due to unobserved heterogeneity across regions correlated to both personality skills and general secondary education share by controlling for an extensive vector of control variables and in four ways. First, we include the lagged dependent variable on the right-hand side, thereby removing any unobserved heterogeneity in the level of personality skills across cantons. Second, we compare regions within relatively homogeneous areas, which limits potential endogeneity problems. Third, we exploit the small variation in the shares of general secondary education across time. Fourth, we apply a second instrument based on the relevance of work-based education in the students' country of origin, formally ensuring testing instrument validity.

Chapter 5 exploits a representative data set of VET teachers of the whole German- and French-speaking Switzerland and data on the workforce from the Swiss Labor Force Survey (SLFS). First, the matching method is applied to compare how much the teachers in vocational subjects earned relative to others with the same characteristics in their original occupation. We use their former wage position compared with the average wage of similar individuals in the former occupation of the teachers as an indicator for their performance in their former occupation as well as for teaching quality. Second, we investigate the wage prospects of career changers who have opted to become teachers. Those who opt to become teachers are unlikely to represent a random sample of all individuals who could theoretically become teachers. Thus, a simple comparison of teachers' wages with average alternative wages is not a useful method of learning whether the decision to change to teaching pays off financially. This study therefore explores the counterfactual situation to the decision to become a teacher by surveying teachers' expectations on both options.

1.3 Findings

In Chapter 2, the tightest MTS-MIV nonparametric bounds indicate a positive causal impact of self-initiated private tutoring in the intermediate school track on students' academic achievement in reading. In particular, estimates reveal that private tutoring increases outcomes by at least 5.8% of a standard deviation. Although these results suggest that private tutoring leads to better outcomes, all 95% confidence intervals of the estimates still fall in the negative range; thus, the hypothesis that self-initiated private tutoring is ineffective cannot be rejected. However, the results suggest a heterogeneous and nonlinear effect; i.e., there is some evidence of an inverse U-shaped function for the relation between the amount of private tutoring and student outcomes. Different types of tutors (e.g., retired teachers, students, older pupils) who may vary in the quality of their private tutoring, the type of settings (e.g., one-to-one, two-to-one) or the frequency of private tutoring (once a week or twice a week) may have a different impact on students; therefore, more research is needed to be

able to rate the different forms of private tutoring and to further tighten the bounds for different sub-samples.

In Chapter 3, even when ignoring the non-random selection or misclassification, the bounds include zero in all cases. They therefore fail to identify the direction of the ATE of truant behavior and the probability of being a high-performing student. Although the imposed assumptions are relatively weak and plausible, there is still a large ambiguity concerning the impact of truancy on students' academic achievement in mathematics. In this regard, the findings are meaningful because they indicate that playing truant is not always a harmful decision of the student resulting in low academic achievement. However, accounting for misreporting, the estimated effect of truancy on academic performance is largely inconclusive. If only two percent of the students declare their truancy behavior incorrectly, the ATE cannot be assigned even under the assumption of exogenous selection. This confirms the sensitivity of analyses involving self-reported data on truancy to misclassification. Further, the results suggest that skipping mathematics classes seems to have a less negative effect on student achievement in the PISA mathematics test in non-Gymnasium schools than in Gymnasium schools. This may indicate that mathematics lessons in Gymnasium schools are used more efficiently by the students who attend them because the students' achievement is considerably higher than that of students who skip classes on purpose.

The evidence in chapter 4 indicates that education can change personality skills. The findings show that work-based education decreases emotion-centered coping, i.e., increases emotional stability. It potentially increases contact-centered coping, indicating an improvement in interpersonal relationships, and potentially reduces intrinsic work motivation. No effect is found for task-centered coping. The effect sizes are economically significant. However, in the long-term, the initial impact of education on personality skills may diminish or even disappear, potentially because students start to work after the school-based upper secondary education. Therefore, the chapter analyzes whether the differences still exist in 2007 and 2010, i.e., approximately four to seven years after concluding secondary education. The results suggest that the impact on emotion-centered coping represents a permanent shift. Analyzing the heterogeneity of the effects between females and males reveals that work-based education compared to school-based education decreases emotion-centered coping for females more than for males.

Chapter 5 shows that those who change careers to teaching earned, on average, significantly more in their former occupations than comparison subjects, which supports the appeal for teaching. Because a positive correlation between productivity in the original occupation and aptitude for teaching in vocational teaching is likely, this result has positive implications for the quality of vocational schools. As to recruitment chances of vocational schools in the individual occupations, the higher the average wage level in an occupation is, the larger is the probability that individuals

recruited from that occupation will rank among the low earners. Teachers need to be recruited from sectors of the rest of the economy with extremely different wage levels, but there is no major wage differential in the educational system. Positive selection for individuals with a university degree applies only to those individuals (largely male) who did not have the option of working part-time in their former occupation, whereas the other teachers with a university degree constitute a negative selection in terms of their relative earnings in their former occupations. Second, although the average teacher tends to rank among the higher earners in their original career, the majority of career changers expect to earn more as a teacher than in their original careers. Again, however, substantial heterogeneity exists given that between one-quarter and one-third of career changers are prepared to accept a cut in wages after changing to teaching. One probable explanation is the very high relevance of non-monetary factors that make teaching a more attractive option, at least for some individuals.

Chapter 2

Does Private Tutoring Work? The Effectiveness of Private Tutoring: A Nonparametric Bounds Analysis

This chapter has been published in another version in Education Economics (2014), 22(4)

2.1 Introduction

Private tutoring – fee-based tutoring in academic subjects that is in addition to the provision by formal schooling – has become popular throughout the world (Baker & LeTendre, 2005; M. Bray, 2007; Mark Bray, 2011; Dang & Rogers, 2008; Jung & Lee, 2010; Mariotta & Nicoli, 2005; Southgate, 2009). Despite the widespread nature of private tutoring, to date, there is little quantitative research on the impact of private tutoring on students' academic performance.

Assessing the impact of privately paid tutoring faces fundamental identification problems. It is well known that a student's educational expenditures are not exogenous. Therefore, participation in private tutoring is endogenous and correlated with at least some unobservable personal and family characteristics. A difficulty is that one cannot observe the outcomes a person would experience under all treatments. At most, one can observe the outcome that a person experiences under the treatment he or she actually receives. Furthermore, the varying nature of private tutoring impairs the identification. Past research mainly focused on *remedial private tutoring programmes*, i.e., tutoring programmes explicitly targeting weaker students lagging behind in some academic skills. This is different to the *self-initiated tutoring*, i.e., private tutoring not initiated by the school or an official party but rather by the student's parents or the student him or herself, as analysed in this present study. Furthermore, it is important to clarify that only self-initiated and fee-based private tutoring in academic subjects is discussed here, in particular reading and mathematics.

Studies on remedial private tutoring programmes for underperforming students find positive effects of private tutoring on students' academic outcome.[1] For India, e.g., Banerjee, Cole, Duflo and Linden (2007) found that a randomised tutoring programme that was targeted toward the weakest

[1] Only studies that control in a credible way for the endogeneity of private tutoring are included in this literature review.

children improved student test scores. For the US, Jacob and Lefgren (2004) applied a regression-discontinuity approach to analyse the effect of summer schools for low-achieving students and found increased academic achievement in reading and mathematics. For Israel, Lavy and Schlosser (2005) found positive short-term effects from a remedial education programme that provided additional instruction to underperforming students. For Italy, De Paola and Scoppa (2014) applied a fuzzy regression discontinuity design and found positive effects of remedial courses as well. Only a few studies analyse the effect of self-initiated private tutoring, and they show mixed results. Ono (2007) explores students in Japan who spend an additional year or more preparing for college and finds positive effects on the quality of colleges in which students can enrol. Survadarma, Suryahadi, Sumarto and Rogers (2006) show, for Indonesia, no significant point estimate effects in the instrument variables approach implemented. Zhang (2013) finds mixed and heterogeneous effects of private tutoring on students' academic achievement in China; inter alia students with lower academic performance and urban students are more likely to benefit from private tutoring, which he attributes to the quality of private tutoring.

While the research focusing on the impact of self-initiated private tutoring is contradictory and sparse, to my knowledge, none of the credible studies provides evidence for a European setting. Different explanations might account for these diverging results on self-initiated private tutoring. One potential cause is that the findings present different local average treatment effects (LATE) and not an average treatment effect (ATE). The measured LATE of private tutoring equals the ATE only if the effect of the tutoring is linear and homogenous. Another potential cause for the diverging results might be that the assumptions made do not hold and lead to an invalid estimate.

Thus, the credibility of empirical analysis on the impact of self-initiated private tutoring on academic achievement depends on the strength of the underlying assumptions. Therefore, this study applies a nonparametric bounds method, introduced by Manski (1990, 1997) and developed by Manski and Pepper (2000, 2009), to calculate the lower and upper bounds of the treatment effect with as few assumptions as possible. This bounds method has been applied, for example, in different recent studies (Blundell, Gosling, Ichimura, & Meghir, 2007; Boes, 2013; De Haan, 2011; Gerfin & Schellhorn, 2006; Gundersen, Kreider, & Pepper, 2012; Kang, 2011; Kreider, Pepper, Gundersen, & Jolliffe, 2012; Manski & Pepper, 2011; Pinkovskiy, 2013). Although this approach produces a range instead of a point estimate, the bounds are informative because the true causal effect of private tutoring is somewhere between these estimated bounds. However, these bounds on the average treatment effect of private tutoring are an important step towards identifying the causal effect of private tutoring on students' academic achievement. However, no study exists that applies this method to identify the causal impact of private tutoring on students' achievement. Moreover, to my knowledge, this is the first study in a European setting analysing self-initiated tutoring, dealing with the correlation between student characteristics, which may affect academic achievement, and

investment in private tutoring and the possibility of a heterogeneous treatment effect (e.g., because of quality differences). Using a representative dataset for Switzerland, this chapter analyses the question of whether self-initiated private tutoring has a causal effect on students' academic achievement in mathematics and reading.

The partial identification approach developed in this research allows for the evaluation of bounds on the ATE of private tutoring under different assumptions, which allows one to successively layer stronger identification assumptions and therefore elucidate how assumptions shape inferences about the causal effect of private tutoring. The analysis starts by investigating the effect of private tutoring without imposing assumptions. Then, the analysis imposes weak nonparametric assumptions to tighten the bounds. First, it makes a monotone treatment selection (MTS) assumption that states that attending private tutoring classes is weakly monotonically related with poor academic outcome. Second, I use the parents' education as a monotone instrument variable (MIV). Third, this study applies the monotone treatment response (MTR), which means the effect of private tutoring is not negative.

The tightest bounds indicate a positive causal impact of self-initiated private tutoring in the intermediate school track on students' academic achievement in reading. Although these results suggest that private tutoring leads to a better outcome, I cannot reject the hypothesis that private tutoring is ineffective. However, the results suggest a heterogeneous and nonlinear effect, e.g., there is some evidence of an inverse U-shaped function for the relation between the amount of private tutoring and student outcome.

This chapter is structured as follows. Section 2.2 describes the Swiss education system with special focus on private tutoring and the data. Section 2.3 explains the identification problem and the nonparametric bounds method. Section 2.4 reports the results of the impact of self-initiated private tutoring on students' academic achievement, and section 2.5 presents the conclusion.

2.2 Swiss Education System and Data

2.2.1 Swiss Education System

Compulsory school in Switzerland comprises nine years of schooling: approximately five to six years of primary school and three to four years of lower secondary school. At the lower secondary school level, different school type models exist that vary from canton to canton[2]. The majority of school type models sort pupils into different school tracks according to their intellectual abilities. Although two to four different tracks exist, the majority of cantons apply a three-track model: an

[2] The equivalent of states in the US.

upper-level school track (Progymnasium), which teaches the more intellectually demanding courses; an intermediate level school track (Sekundarschule), and finally, one offering basic-level courses (Realschule).

After finishing compulsory schooling (9th grade), students can choose between two different possibilities: full-time educational school (Gymnasium or Fachmittelschule) or vocational track (apprenticeship training). In Switzerland, approximately 20% of school graduates attend a Baccalaureate school (Gymnasium), which prepares then for university. Approximately 60% of school graduates choose apprenticeship training. This so-called ”dual-education” provides them with formal and on-the-job training within a training firm and one to two days per week of formal schooling in a vocational school.

2.2.2 Private Tutoring in Switzerland

Private tutoring in Switzerland is completely unregulated and takes mainly two different forms. The first type and the lion’s share is private instruction by a privately paid teacher either at the teacher’s or at the student’s home (Hof & Wolter, 2012). Because of the unregulated market, the quality of the teachers is neither defined nor examined; thus, no information about the quality is available. The second type of private tutoring is undertaken by profit-oriented school-like organisations where professional teachers or students tutor in a classroom setting (for example, ´Kick Lernstudio´ or ´Studienkreis´). Such centres usually own or rent multi-story buildings in city centres. Students attend these centres in addition to formal school hours. These centres provide smaller class sizes (one-to-one, in groups of two or sometimes up to 10 students), special materials, e.g., workbooks, and improved student-teacher relations compared with the formal schools.

Research on the extent of private tutoring in Switzerland is rare. Analysing TIMSS (Trends in International Mathematics and Science Study) data from 1995 Baker and LeTendre (2005) revealed a weekly participation rate of 25% for 8th graders. A study for the canton of Tessin using PISA 2003 data reported a participation rate of 15% for 9th graders (Mariotta, 2006).

2.2.3 Data

This study uses data from the Programme for International Student Assessment (PISA) 2009 conducted by the Organization for Economic Cooperation and Development (OECD). Every three years since 2000, PISA has measured the performance of 15-year-old students at the end of compulsory schooling. Performance in mathematics, science and reading are investigated; PISA in 2009 focused on reading (OECD, 2011a).

The PISA survey follows a two-stage sampling process. First, schools are sampled, and then, students are sampled in the participating schools. In a simple random sample of schools, every

school has the same selection probability, and within the selected schools, the student selection probability will vary according to the school size because, in reality, schools differ in size. Therefore, in a small school, the student selection probability will be larger than in a large school. To avoid these unequal selection probabilities for pupils, the schools' probability to be selected are weighted with their size (OECD, 2009).

The PISA 2009 data collection for Switzerland includes a large nationally representative sample of 15-year old students and a supplementary study of grade-9 students from a selection of cantons. These surveys include a national option on the demand of private tutoring. These questions provide information about the frequencies, motives, and other relevant variables related to private tutoring demand in 8/9th grade among 9th graders. The analysis in this chapter makes use of the national 9th grade survey with 13,472[3] students.

The nationwide representative PISA 2009 dataset indicates a participation rate in private tutoring of 30% for Switzerland (Hof & Wolter, 2012) for 9th grade students, with approximately 40% of the students receiving private tutoring attending private tutoring classes on a regular basis in mathematics or reading. Girls and students with more educated parents are significantly more often sent to private tutoring lessons.

The main outcome variable is students' academic achievement. Academic achievement is measured with the PISA 2009 scores of Swiss 9th graders in mathematics and reading. To tighten the nonparametric bounds, an instrument variable approach is applied in which parents' schooling serves as a monotone instrument variable. This will be explained in more detail in the next section. Table 2-1 shows the descriptive statistics for students without any private tutoring and with private tutoring in reading and mathematics.

[3] Because of item non-response, 2383 observations were deleted.

Table 2-1: Descriptive Statistics of the Sample

	No private tutoring			Private tutoring in reading			Private tutoring in mathematics		
Variable	*N*	Mean	SD	*N*	Mean	SD	*N*	Mean	SD
Female	13,472	0.51	0.5	1216	0.49	0.5	2757	0.6	0.49
Age	13,472	15.76	0.63	1216	15.88	0.66	2757	15.76	0.63
Tertiary education	13,472	0.55	0.5	1216	0.59	0.49	2757	0.6	0.49
Secondary education	13,472	0.1	0.3	1216	0.11	0.31	2757	0.1	0.31
Vocational education	13,472	0.31	0.46	1216	0.25	0.43	2757	0.27	0.44
No postobligatory	13,472	0.03	0.18	1216	0.05	0.22	2757	0.03	0.18
First gen	13,472	0.09	0.29	1216	0.16	0.37	2757	0.1	0.3
Second gen	13,472	0.3	0.46	1216	0.35	0.48	2757	0.35	0.48
Native	13,472	0.61	0.49	1216	0.49	0.5	2757	0.55	0.5
Foreign lang	13,472	0.15	0.36	1216	0.25	0.43	2757	0.17	0.37
ISEI	13,472	51.02	15.83	1216	50.52	15.92	2757	52.19	15.3
Siblings	13,472	0.89	0.32	1216	0.87	0.34	2757	0.85	0.36
Single	13,472	0.14	0.35	1216	0.14	0.35	2757	0.14	0.35
Mixed	13,472	0	0.06	1216	0.01	0.09	2757	0	0.07
Nuclear	13,472	0.86	0.35	1216	0.85	0.36	2757	0.85	0.35
Latin Swiss	13,472	0.28	0.45	1216	0.29	0.46	2757	0.33	0.47
Upper level track	13,472	0.35	0.48	1216	0.24	0.43	2757	0.35	0.48
Intermediate track	13,472	0.39	0.49	1216	0.39	0.49	2757	0.4	0.49
Basic level track	13,472	0.25	0.44	1216	0.34	0.47	2757	0.23	0.42
PISA read	13,472	510.8	85.01	1216	473.92	89.8			
PISA math	13,472	545.25	92.4				2757	519.45	88.36

Notes: Female is a dummy variable with female = 1 and male = 0. Tertiary, secondary, vocational, and no postobligatory education refer to the mother's or father's highest education level. First gen, second gen, and native refer to the immigrant status. Foreign lang refers to the language spoken at home, which is one if the language spoken at home is not the test language. ISEI refers to the *International Socio-Economic Index of Occupational Status* with values from 16 (lowest status) to 90. Siblings is a dummy variable with one if the student has at least one sibling. Single, mixed, and nuclear refer to single-parent, mixed, or nuclear household, respectively. Latin Swiss is a dummy variable for the French- and Italian-speaking region in Switzerland. Upper level, intermediate, and basic-level track refer to the track the student is enrolled in. PISA read and PISA math refer to the PISA competences in reading and mathematics. The data stem from the Swiss PISA 2009 cohort.

2.3 Partial Identification Strategy

I consider the problem of learning the effect of private tutoring on students' academic achievement (in mathematics and reading). The analysis aims to identify the average treatment effect (ATE) of going to private tutoring classes on students' achievement, that is,

$$ATE_{r,m}(1,0) = E[y(1)|x] - E[y(0)|x], \quad [1]$$

where y is student's academic achievement in PISA, and y(1)[4] denotes a student's outcome if attending private tutoring classes and y(0) if not. For each student, there are two potential outcomes, y(1) and y(0). The ATE represents the causal effect of tutoring on achievement and is calculated by the mean outcome if all students receive private tutoring (y(1)) versus the mean outcome if all students do not attend private tutoring (y(0)). See Equation [1].

Under the assumption of exogenous treatment selection (ETS), the ATE is point estimated. ETS assumes that E[y(1)|z=0] = E[y(1)|z=1] and E[y(0)|z=0] = E[y(0)|z=1], and therefore (Beresteanu & Manski, 2000) the ATE = E[y(1)|z=0] - E[y(0)|z=1] = E[y|z=1] - E[y|z=0]. In particular, z=1 indicates that the pupil truly received the treatment, and z=0 otherwise.

For each student, we do not observe one of the two potential outcomes z=0 (e.g., what a student's academic achievement would have been if he had not attended private tutoring); therefore, this approach leads to biased results because students who receive private tutoring may differ in various unobserved variables from those who do not. This is referred to as the selection problem.

Instead of imposing assumptions that lead to a point estimate, this analysis applies the nonparametric bounds method (Manski, 1990, 2007; Manski & Pepper, 2009) and imposes as few assumptions as possible to calculate a lower and an upper bound of the private tutoring effect. The true causal effect of the treatment lies somewhere between the lower and upper bounds. These bounds lead to partial conclusions.[5]

For these bounds, I define the outcome (y) as the PISA test score of a student in mathematics or reading and t as the treatment indicator. $z \varepsilon T$ denotes, as well, the treatment received by person. z=1 denotes that a student participated in private tutoring in the 8th and/or 9th grade in mathematics or reading, and z= 0 otherwise. The response function y(.) : T -> Y maps the treatments $t \varepsilon T$ into outcomes $y(t) \varepsilon Y$. y(t)(t=z) is the realised outcome, and $y(t)(t \neq z)$ is the counterfactual. The outcome space Y has in general bounds $-\infty < K_0 < K_1 < +\infty$ and when specified,

[4] To make the notation more compact, I leave the conditioning on covariates (x) and the notation for mathematics (m) and reading (r) implicit in the following.

[5] It is important to notice that these bounds are not confidence intervals. They express the ambiguity created by the selection problem (Manski & Pepper, 2011).

the greatest lower bound $K_0 \equiv \inf Y$ and the least upper bound $K_1 \equiv \sup Y$. Using the Law of Iterated Expectations and following Manski and Pepper (Manski, 2007; Manski & Pepper, 2011), I decompose

$$E[y(1)] = E[y(1)|z=1]\, P(z=1) + E[y(1)|z=0]\, P(z=0), \qquad [2]$$

where P(z=1) or P(z=0) are the probabilities of receiving or not receiving the treatment.

2.3.1 Worst-Case Bounds for Average Treatment Effects

Manski (1990) shows that it is possible to identify bounds by adding very weak assumptions. I am, however, not able to identify the unobservable counterfactual (latent outcome) E[y(1)|z=0] or E[y(0)|z=1] from my data without imposing very strong and probably implausible assumptions. Therefore, this analysis replaces the unobserved by its bounds, and these are, for each treatment t, the worst-case (WC) bounds (no-assumptions bounds following Manski (1990)) with the very weak assumptions of a bounded output y(t) and stable unit treatment value. This yields the following sharp bounds for y(t) in the binary treatment case of private tutoring:

$$E[y(1)|z=1]\, P(z=1) + K_0\, P(z=0) \leq E[y(1)] \leq E[y(1)|z=1]\, P(z=1) + K_1\, P(z=0) \qquad [3]$$

$$E[y(0)|z=0]\, P(z=0) + K_0\, P(z=1) \leq E[y(0)] \leq E[y(0)|z=0]\, P(z=0) + K_1\, P(z=1).$$

The resulting bound on the ATE[6] is

$$E[y(1)|z=1]\, P(z=1) + K_0\, P(z=0) - [E[y(0)|z=0]\, P(z=0) + K_1\, P(z=1)] \qquad [4]$$

$$\leq E[y(1)] - E[y(0)] \leq$$

$$E[y(1)|z=1]\, P(z=1) + K_1\, P(z=0) - [E[y(0)|z=0]\, P(z=0) + K_0\, P(z=1)].$$

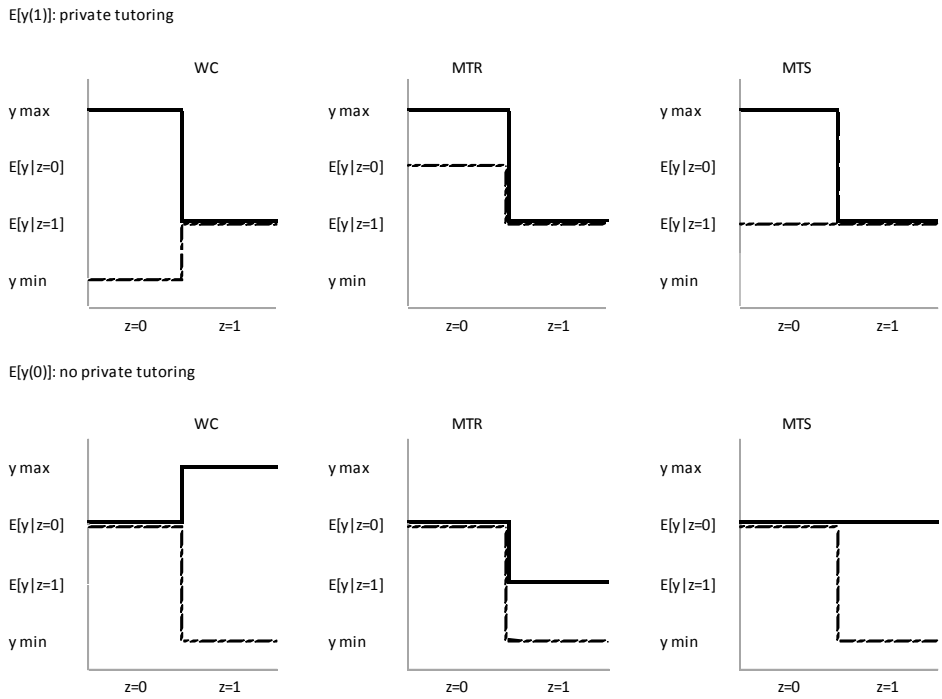
The two illustrations on the left in Figure 2-1 show the upper and lower bounds for E[y(1)] and E[y(0)] without assumptions. These worst-case bounds[7] are often too wide to be useful. To obtain narrower bounds, a few assumptions can be invoked. The analysis will subsequently add the MTS, MIV, and the MTR assumptions.

---

[6] The ATE (E[y(1)] - E[y(0)]) is calculated as follows: The lower bound on E[y(1)] minus the upper bound on E[y(0)] is the lower bound of the average treatment effect. The upper bound on E[y(1)] minus the lower bound on E[y(0)] is the upper bound of the ATE.

[7] Manski bounds are sharp bounds, i.e., nothing else can be learned in face of the censored data (see the proof in Heckman & Leamer, 2007; Heckman & Vytlacil, 2000; Manski, 2007).

Figure 2-1: How MTS and MTR Tighten the Bounds in the Binary Case



Source: Figure based on De Haan (2011).

2.3.2 Monotone Treatment Selection (MTS)

The first assumption introduced in the analysis is MTS, which supposes that sorting into treatments is not exogenous but monotone in the sense that the counterfactual outcome is smaller for those students who participated in private tutoring (z=1) than for those who did not participate (z=0). In other words, students who participated in private tutoring have a higher probability (because of observed and unobserved characteristics) of being a bad achiever than those who did not participate in private tutoring would have had if they had participated in private tutoring. Therefore, I assume a negative self-selection with $E[y(1)|z=1] \leq E[y(1)|z=0]$ and $E[y(0)|z=1] \leq E[y(0)|z=0]$. This MTS assumption implies that if all students received private tutoring, students actually receiving tutoring would, on average, still perform worse than students actually without private tutoring. However, this assumption may be problematic[8] because it implies that students who do not select into private tutoring can benefit from tutoring; private tutoring, though, may be ineffective for already high-performing students or may even have a detrimental effect. To account for this possibility, I analyse the bounds for high and low performing students' (high and low PISA scores) separately.[9]

The two illustrations on the right in Figure 2-1 show how the MTS assumption can tighten the bounds. One can observe the mean achievement for students that did not attend private tutoring

[8] Furthermore, it could be that ability and taste for private tutoring are be positively associated. Therefore, more able students would want to go to further lessons after school.

[9] High-scoring students are students with a PISA competence level 6, and low-performing students are those with a PISA competence level equal or below 1.

lessons. Under an MTS assumption, this achievement will not be lower than the mean achievement for students actually receiving private tutoring. Hence, the mean realised students' achievement for students with private tutoring is the lower bound, indicating that students with treatment could not have performed any better in the control state than those observed in the control group. MTS yields a lower bound for the counterfactual E[y(1)|z=0], which is E[y(1)|z=1], because for each z<t, it must be true that E[y(t)] is at most as large as E[y|z=t] and an upper bound for E[y(0)|z=1], which is E[y(0)|z=0]. In the binary case, the bounds under MTS are as follows:

$$E[y|z=1] \leq E[y(1)] \leq E[y(1)|z=1]\, P(z=1) + K_1\, P(z=0), \qquad [5]$$

$$E[y(0)|z=0]\, P(z=0) + K_0\, P(z=1) \leq E[y(0)] \leq E[y|z=0]. \qquad [6]$$

2.3.3 Monotone Instrument Variable (MIV)

A second assumption to tighten the bounds is the presence of an instrument variable (IV). This analysis will use the parents' education v as a MIV. With this additional variable v, it is possible to create sub-samples for each value of v and then to obtain bounds on the mean potential outcomes within each of these sub-samples (Manski & Pepper, 2000).

This approach applies the traditional IV[10] but loosens the assumptions with mean monotonicity[11] (Manski & Pepper, 2000):

$$u_1 \leq u \leq u_2 \;\rightarrow\; E[y(t)|v=u_1] \leq E[y(t)|v=u] \leq E[y(t)|v=u_2]. \qquad [7]$$

In contrast to an IV assumption with mean independence, the MIV assumption allows a weakly monotone positive relationship between v and the mean potential outcome (Manski & Pepper, 2000). By using the parents' education as an MIV, I assume that the mean schooling function of the pupil is monotonically increasing in the parents' education.[12] This innocuous MIV assumption allows for a direct impact of the parents education on students' academic achievement as long as the effect is not negative. The choice of the instrument is based on research on intergenerational mobility that indicates that educational achievement is positively correlated with the parents' education (Björklund & Salvanes, 2011; Black & Devereux, 2011).

---

[10] For example, Ono (2007) uses tutoring during secondary education as in IV to measure the effect of tutoring in tertiary education.

[11] The identifying power of an MIV is examined in Manski and Pepper (2000).

[12] The MIV used is discrete and takes four possible values: no post-obligatory education, vocational education, secondary academic education, and tertiary education.

The MIV bounds are (similar for E[y(0)][13]) as follows:

$$\sum_{u\varepsilon V} P(v=u)\ \{\sup_{u1\leq u}\ [E(y|\ v=u_1,\ z=1)\ P(z=1|\ v=u_1) + K_0\ P(z=0|\ v=u_1)]\}$$

$$\leq E[y(1)] \leq \qquad [8]$$

$$\sum_{u\varepsilon V} P(v=u)\ \{\inf_{u2\geq u}\ [E(y|\ v=u_2,\ z=1)\ P(z=1|\ v=u_2) + K_1\ P(z=0|\ v=u_2)]\}.$$

From Equations [7] and [8], it follows that for the sub-sample v = u, there is a new lower bound, which is the largest lower bound over all sup-samples v ≤ u. The new upper bound is the smallest upper bound over all sub-samples v ≥ u. To calculate these bounds, the analysis divides the sample into four groups of parents' education and uses the average estimates of MTS or MTS-MTR bounds to obtain the MTS-MIV or MTS-MTR-MIV bounds.

2.3.4 Monotone Treatment Response (MTR)

The third assumption employed is the MTR (Manski, 1997). MTR states, ceteris paribus, that the outcome is a weakly increasing function of the treatment, such that $\delta \geq 0$ for every student. The assumption implies that there exist no negative impacts of private tutoring on students' academic performance.

However, potential negative effects of private tutoring on students' academic outcome could arise if private tutoring lowers students' self-esteem/motivation or crowds out students' self-learning time or reduces students' attention in class. Lee (2013) shows that private tutoring positively affects low-achieving students in terms of their attention to school lessons and has no effect for middle- and upper-achieving students. To control for a possible crowding-out effect on students' self-learning time, I make use of the PISA 2006 questions about tutoring out-of-school and self-learning time.[14] Comparing the self-learning time for students with and without tutoring (see Appendix) reveals a significant positive effect of tutoring on students' self-learning time in reading and mathematics. These results are robust for all levels (zero up to six and more hours per week) of self-learning time. However, neither a possible crowding out effect or a negative effect on students' motivation can be completely excluded. Thus, MTR is a controversial assumption.

---

[13] For proof, see Manski and Pepper (2000).

[14] Self-learning time was not questioned in PISA 2009. For this present research, the questions in the international PISA student questionnaire concerning tutoring are not detailed enough to distinguish between private (and privately-paid) tutoring and other out-of-school time lessons. Comparing participation rates in tutoring (international question) and private tutoring (national option on privately-paid tutoring) shows an overestimation in the international question. The international question leads to participation rates of 40% (OECD, 2011b) in tutoring compared to 30% participation rate in privately paid tutoring.

Furthermore, it may be that the quantity of private tutoring makes a difference, i.e., a certain threshold may exist beyond which private tutoring may be ineffective or even detrimental. To account for the possibility of an inverse U-shaped function for the relation between the amount of private tutoring and student academic outcome, I split the sample according to the amount of private tutoring and checked whether tutoring affects individuals with low and high values in tutoring differentially. With the data at hand, I can divide the intensity of private tutoring in private tutoring on a regular basis, e.g., tutoring over several weeks or months (high value of private tutoring) and private tutoring on an irregular basis, e.g., tutoring during some lessons (low value of private tutoring).[15]

However, MTR allows estimating whether there exists a positive effect of private tutoring or whether there is no effect at all. MTR assumes that treatments are ordered, and y(.) is monotone in the treatment, and therefore, observations of the realised outcome y can be informative about the counterfactual outcomes y(t), $t \neq z$ (Manski, 2007). MTR for E[y(1)] is specified as follows, when private tutoring is assumed to weakly increase students' performance:

$$E[y(1)|z=0] \geq E[y(0)|z=0],$$

$$E[y(1)|z=1] \geq E[y(0)|z=1]. \quad [9]$$

The two illustrations in the middle of Figure 2-1 show how the MTR assumption can be used to tighten the bounds around the two potential outcomes. The data provides information on the mean outcome of students without private tutoring. Under the MTR assumption, for students without private tutoring, their observed mean outcome will not be lower than to what their mean outcome would have been if they had attended private tutoring classes. Therefore, the observed mean outcome for these students without private tutoring E[y|z=0] can be used to tighten the lower bound for students with z=0. For the students with private tutoring, under MTR assumption, the potential outcome will not be higher than the mean outcome we observe. E[y|z=1] can therefore be used as an upper bound for the students with z=1.

In the case of a binary treatment, the bounds under MTR can be expressed as follows:

$$E[y] \leq E[y(1)] \leq E[y|z=1]\, P(z=1) + K_1\, P(z=0), \quad [10]$$

$$E[y|z=0]\, P(z=0) + K_0\, P(z=1) \leq E[y(0)] \leq E[y]. \quad [11]$$

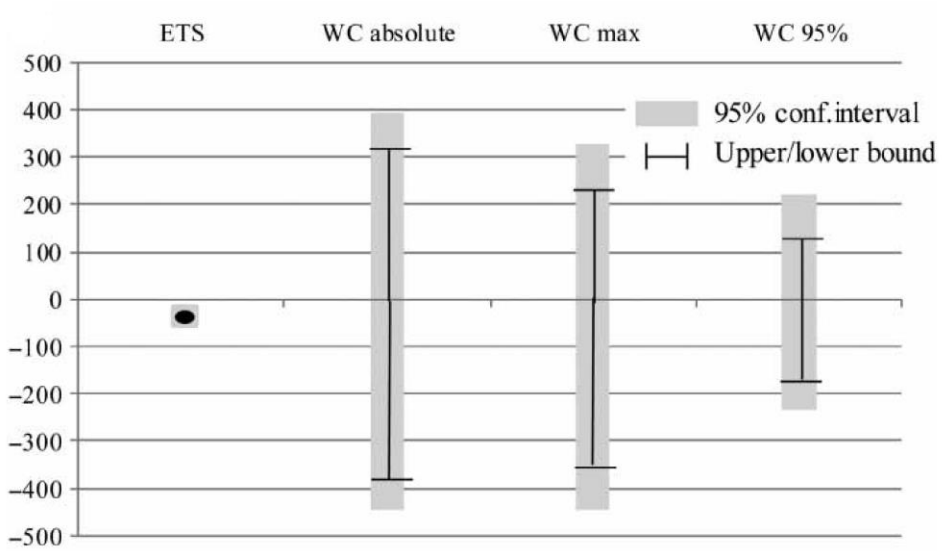
If I impose MTR as well as MTS, the lower bound for E[y(1)] is the higher lower bound of MTR and MTS. The upper bound on E[y(0)] is the lower bound of MTR and MTS.

[15] The data at hand do not allow an analysis on the number of weekly hours spent on private tutoring.

2.4 Results

Under ETS, a negative impact of private tutoring on students' academic achievement (Figure 2-2 and Figure 2-3) is measured. With the data at hand, there might be a self-selection of bad performing students into private tutoring explaining the negative relationship between private tutoring on student performance. Figure 2-2 and Figure 2-3 show the worst-case nonparametric bounds on students' academic achievement in reading (Figure 2-2) and mathematics (Figure 2-3) as a function of private tutoring. PISA scores of students can never be lower than 0 points and never higher than 1000 points.[16] The achieved PISA scores 2009 for reading and mathematics lie between the interval 120 and 860. Thus, the absolute worst-case bounds indicate the ATE for reading must be in the interval [-394, 346]. For the ATE for mathematics, the absolute worst-case bounds lies in the interval [-422, 318]. Using the actual maximum and minimum points (WC max) in PISA 2009 for reading (124, 771) and mathematics (125, 856), the bounds shrink a little bit, and the ATE for reading test must be in the interval [-382, 265] and for mathematics in [-417, 312]. In the Swiss PISA 2009, 95% of the students scored in reading in the interval [347, 675] and in mathematics in the interval [369, 721]. Thus, using these 95% minimums and maximums to calculate the upper and lower worst-case bounds, the ATE for reading is in the interval [-171, 157] and for mathematics must be in the interval [-184, 168]. Applying the very weak assumption that the students will score somewhere in between where 95% of all students allows to reduce the interval for the ATE significantly. However, without additional assumptions about the selection, I cannot eliminate the possibility that private tutoring has a positive or negative effect on students' academic achievement.

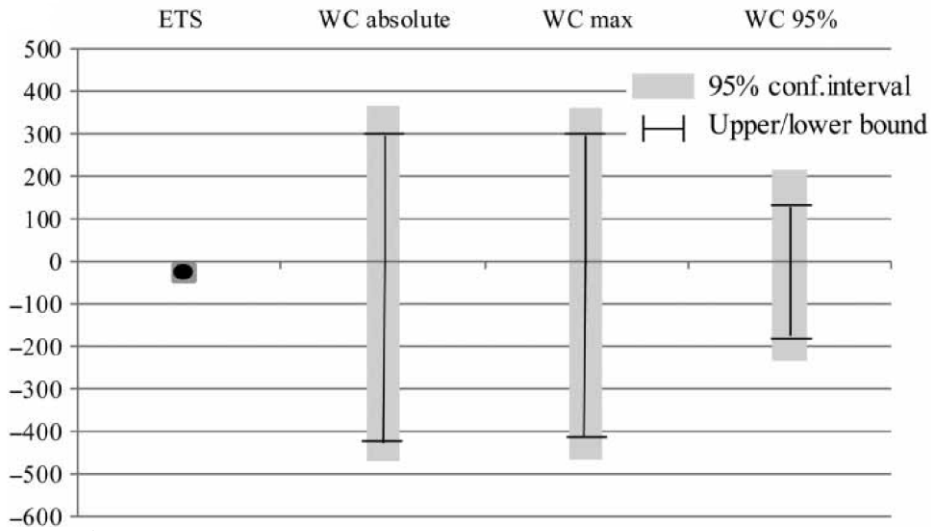
Figure 2-2: Exogenous Treatment Selection and Worst-Case Bounds on the ATE in Reading



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

[16] Manski and Pepper (2011) applied the method of restricting the minima and maxima.

Figure 2-3: Exogenous Treatment Selection and Worst-Case Bounds on the ATE in Mathematics

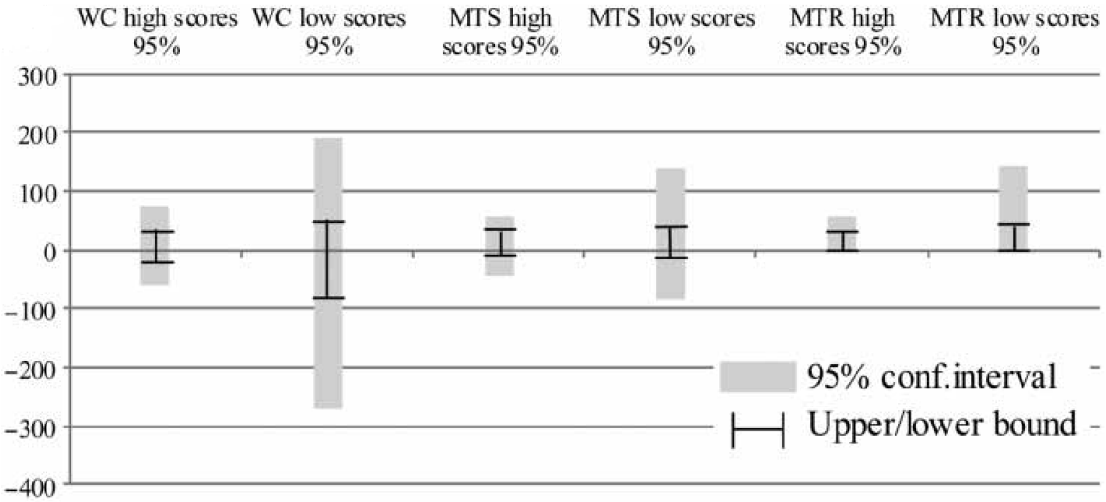


Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

Adding the MTS assumption significantly increases the lower bound (Figure 2-6 and Figure 2-7). Under the MTS assumption, the lower bound for the 95% distribution, for example, is measured to be -40 for reading, which is notably improved compared with the worst-case lower bound of -170. However, the applied MTS assumption may be controversial because it implies that students who do not select into private tutoring can benefit from tutoring. To account for the possibility that private tutoring may be ineffective for already high-performing students or may even have a detrimental effect, Figure 2-4 and Figure 2-5 show the bounds for the two subsamples of high- and low-performing students.[17] MTS results indicate a positive effect of private tutoring in mathematics for low-performing students but a lower variance for high-performing students. With this evidence, however, I cannot reject the hypotheses that high-performing students will not profit from private tutoring.

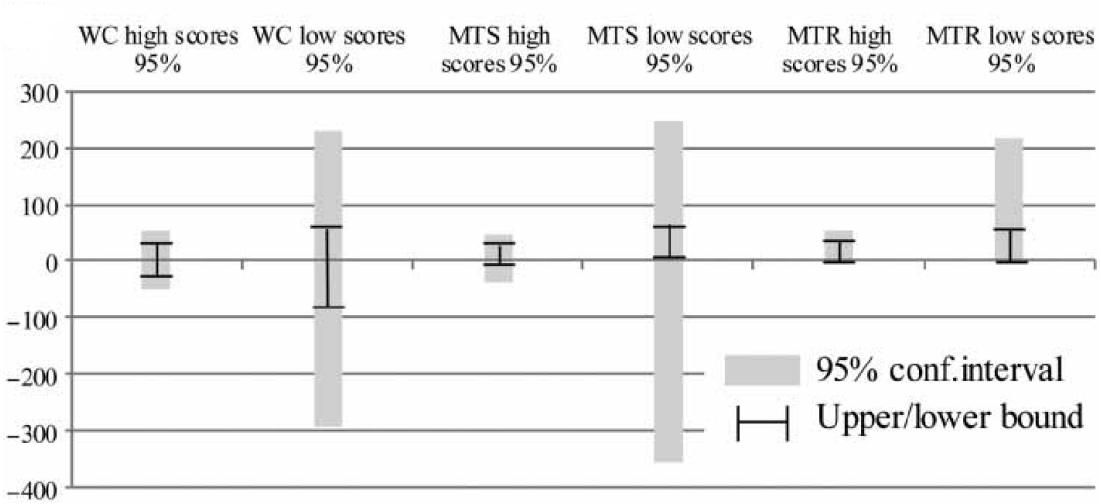
[17] Because of the small numbers for some MIV values, calculation of MIV-bounds for high and low performing students is not possible.

Figure 2-4: Bounds on the ATE in Language: High and Low PISA Scores



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

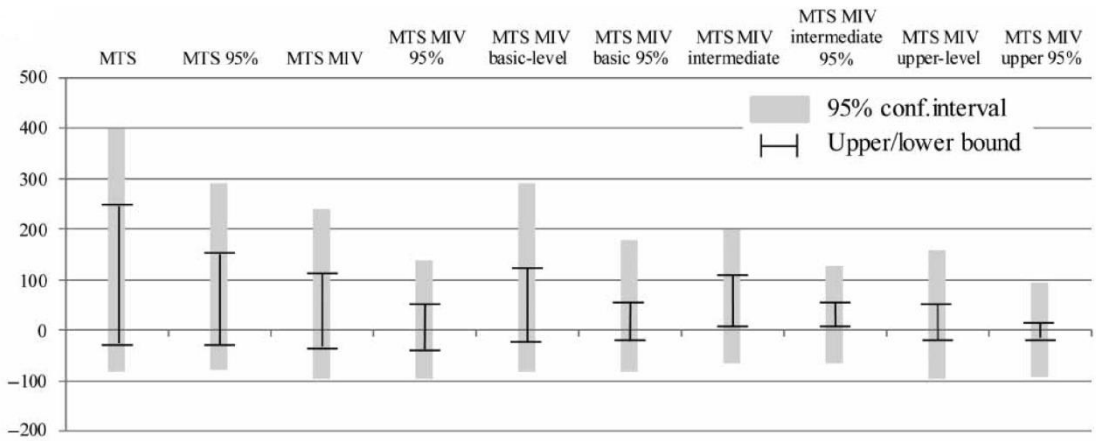
Figure 2-5: Bounds on the ATE in Mathematics: High and Low PISA Scores



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

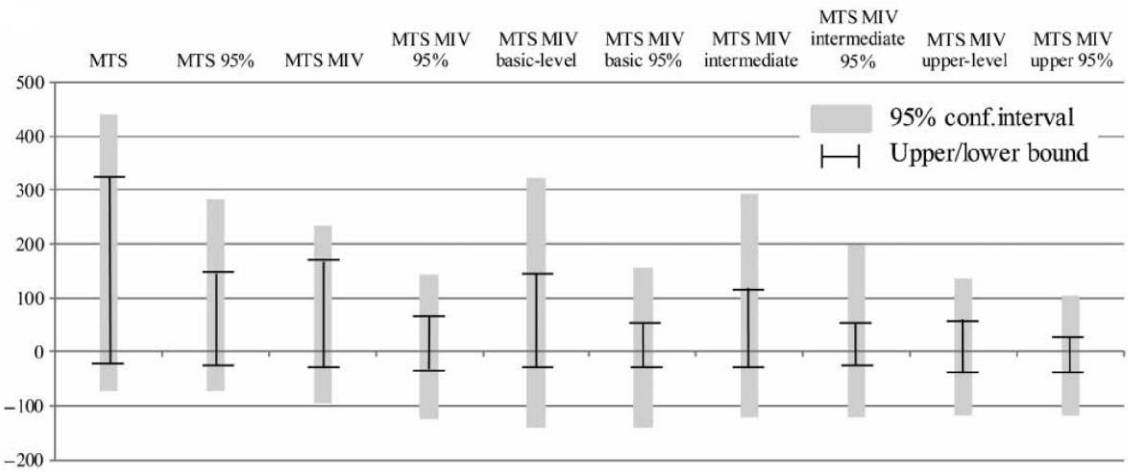
Combining the MTS and MIV assumptions (Figure 2-6 and Figure 2-7) does not further reduce the lower bound but significantly reduces the upper bound, for example, in reading to 60 compared with the worst-case upper bound of 150. While adding this MIV assumption substantially reduces the ambiguity created by the selection problem, there still remains uncertainty about the ATE. Calculating the MTS-MIV bounds in reading for the different school tracks, the bounds narrow to [2, 60] for the intermediate track; i.e., the impact of private tutoring appears to be at least slightly beneficial in reading. While this bound is positive, the confidence interval includes zero; therefore, I cannot reject the hypothesis that private tutoring is ineffective for all students. The narrowest bounds are found for students in the upper-level school track. However, all MTS-MIV bounds exclude the ETS point estimate.

Figure 2-6: Bounds on the ATE in Reading: MTS, Joint MTS and MIV Assumptions



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

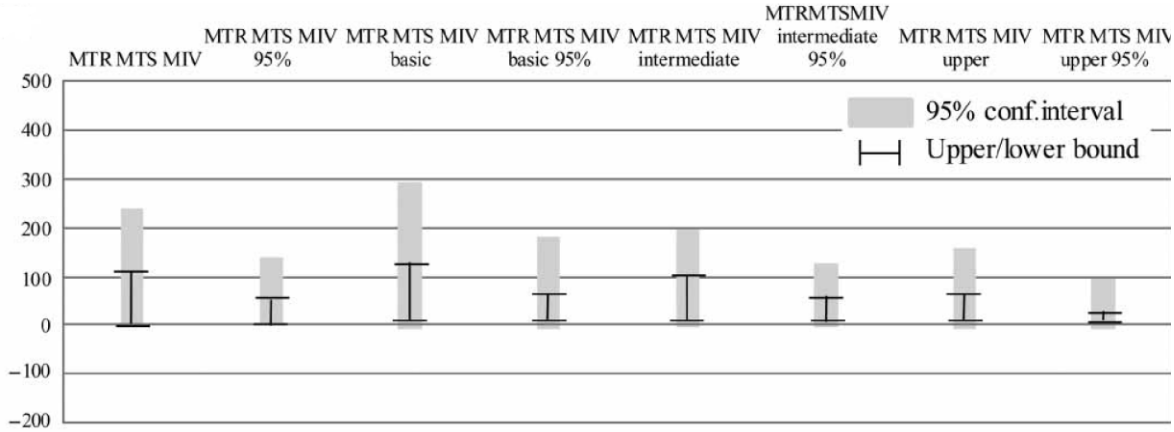
Figure 2-7: Bounds on the ATE in Mathematics: MTS, Joint MTS and MIV Assumptions



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

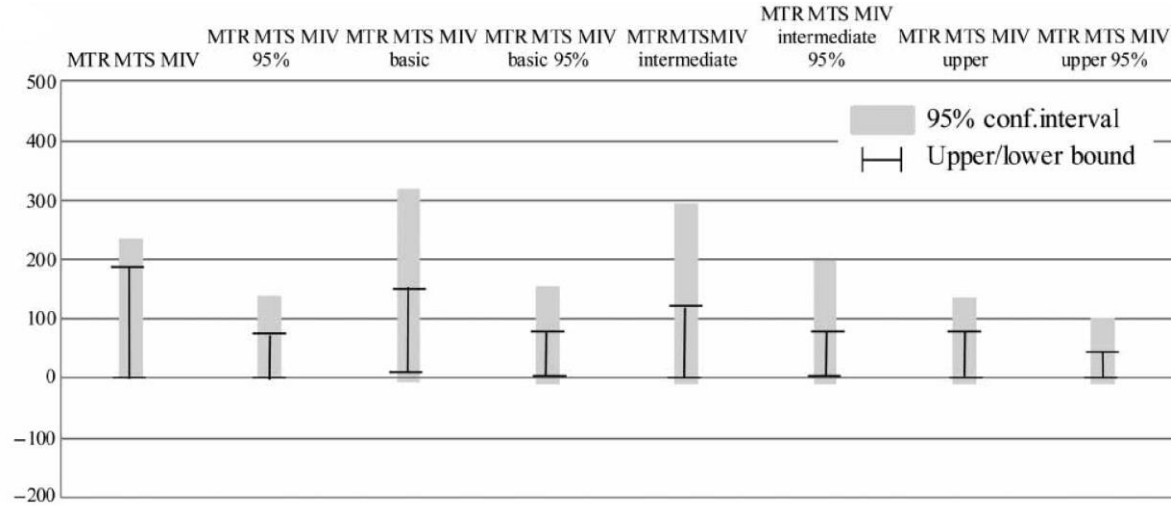
Imposing all three assumptions (MTR, MTS, MIV) jointly leads to the bounds in Figure 2-8 and Figure 2-9. The combined assumptions increase the lower bound significantly in reading and mathematics. Adding this additional MTR assumption implies that the ATE must be nonnegative. Therefore, private tutoring cannot increase the probability of a low academic outcome. This MTR is a very controversial assumption because private tutoring may crowd out self-learning time or negatively affect students' self-esteem and motivation.

Figure 2-8: Bounds on the ATE in Reading: Joint MTS, MTR and MIV Assumptions



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

Figure 2-9: Bounds on the ATE in Mathematics: Joint MTS, MTR and MIV Assumptions

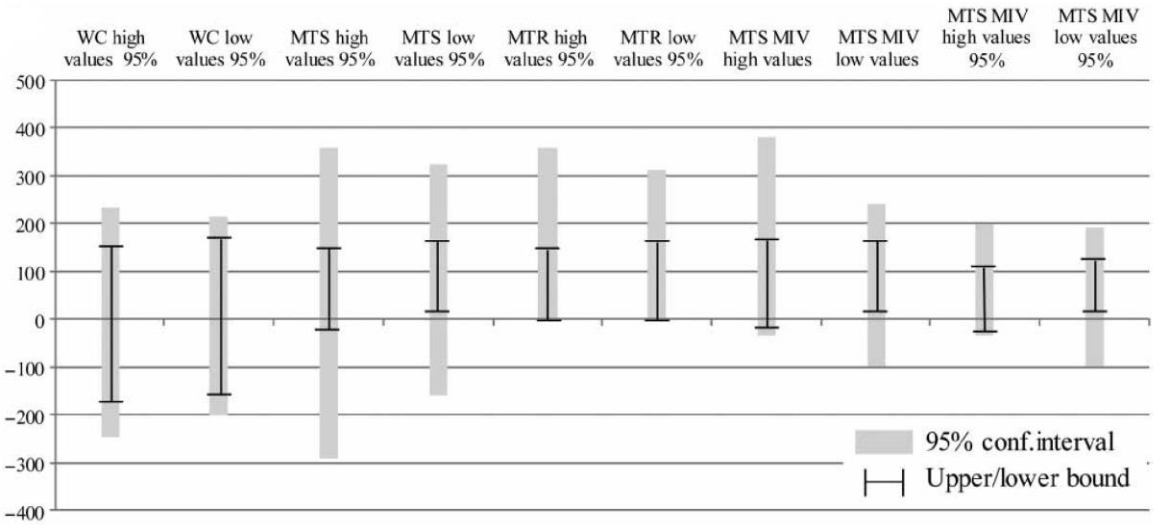


Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples.

However, a certain threshold may exist beyond which private tutoring may have no effect or may even be detrimental. Analysing the impact of high values of private tutoring compared to low values of private tutoring, Figure 2-10 and Figure 2-11 account for this possibility. These bounds indicate that in regard to private tutoring, the intensity of the treatment may matter and that private tutoring on a regular basis may be inefficient or even detrimental on academic outcomes. Although MTS and MTS-MIV bounds suggest a positive impact of low values of self-initiated private tutoring on student academic outcome in reading and mathematics, the 95% confidence intervals for the estimates still fall in a negative range; thus, the hypotheses that private tutoring for low values of tutoring is ineffective cannot be rejected.

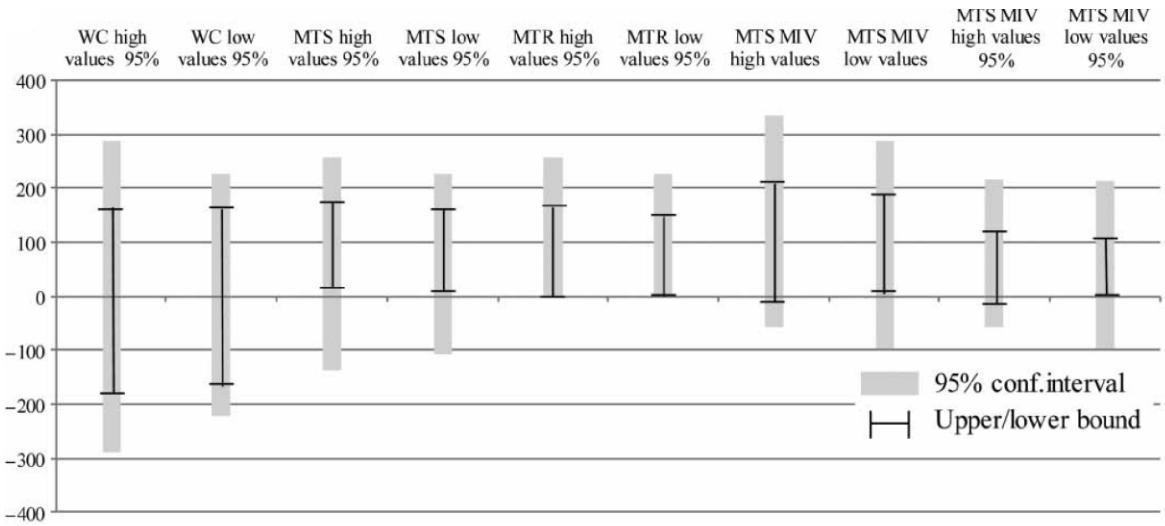
The calculated bounds demonstrate that additional assumptions can have substantial identifying power compared to the worst-case bounds, as the lower and upper bounds shrink. While these findings indicate that private tutoring improves students' academic achievement in reading on the intermediate level school track and for low values of private tutoring in reading and mathematics, these results have to be interpreted carefully. For all assumptions and subsamples, the 95% confidence interval includes zero; thus, I cannot reject the hypothesis that private tutoring is ineffective. Although the imposed assumptions are relatively weak, there is still a large ambiguity concerning the impact of private tutoring on students' academic achievement in reading and mathematics.

Figure 2-10: Bounds on the ATE in Language: High and Low Values of Private Tutoring



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples. High values of private tutoring refer to private tutoring on a regular basis.

Figure 2-11: Bounds on the ATE in Mathematics: High and Low Values of Private Tutoring



Note: Confidence intervals are estimated by using the variation around lower and upper bound with 300 pseudosamples. High values of private tutoring refer to private tutoring on a regular basis.

2.5 Conclusion

Given the large proportion of students employing private tutoring and the monthly spending of households for private tutoring, private tutoring has become a significant part of the educational system that cannot be ignored by educational researchers and policymakers. Thus, better understanding of the causal effect of self-initiated private tutoring in academic subjects is crucial for policy decisions. However, it is challenging to accurately assess the effectiveness of private tutoring because students who choose to participate may be quite different from the students who do not.

Regressing students' academic achievement on private tutoring generally gives large negative estimates. As there is a high probability of a special selection into private tutoring, these estimates are not all informative about the causal effect of private tutoring on students' academic outcome. Therefore, different identification strategies have been used in the empirical literature to estimate the true causal effect of private tutoring. While research on remedial private tutoring programmes indicates positive effects on outcome, empirical evidence on self-initiated private tutoring indicates mixed effects for point estimates on the effect of private tutoring on academic outcome.

A potential cause of diverging results is that the findings indicate different LATE and not ATE, which are equal only if the effect of the tutoring is linear and homogenous. Another potential cause for the diverging results might be that the assumptions made do not hold, thus leading to an invalid estimate. For example, even though the PISA results could be constructed as part of a longitudinal study, comparing the results of two standardised tests over time would lead to biased or invalid results because of the unlikely assumption of comparable time-trends.

Therefore, the present study contributes to the literature by applying an alternative method to overcome the selection bias and to identify the effect of self-initiated private tutoring on student's outcome. This article uses a nonparametric bounds method to analyse the causal effect of private tutoring by relying on a set of relatively weak nonparametric assumptions. The step-by-step approach applied in this chapter allows the reader to identify which assumptions tighten the bounds in which direction. Moreover, the analysis drops the probably unrealistic assumption of a linear and homogenous effect of private tutoring on students' academic achievement, e.g., it allows for the effect of private tutoring may vary with the quality of the private tutoring. The applied method obtains bounds around the average treatment effect even when the treatment effect differs between schools or students.

For my preferred MTS-MIV models, the results imply that private tutoring leads to increased academic achievement in reading for students on the intermediate track. In particular, estimates reveal that private tutoring increases outcome by at least 5.8% of a standard deviation. Although

these results suggest that private tutoring leads to notable improvements in students' academic achievement in reading, I cannot reject the hypothesis that private tutoring may be ineffective for the students' outcome.

In addition to these results, the findings suggest that the quantity of private tutoring makes a difference and that there may be a threshold beyond which private tutoring is ineffective or even detrimental. The more the better may not be true because bounds on ATE of self-initiated private tutoring on academic outcome show a positive lower bound only for students with private tutoring on an irregular basis for reading and mathematics.

To summarise briefly, self-initiated private tutoring was found to have mixed effects on the two analysed subjects and heterogeneous effects depending on track, competence level and quantity of private tutoring. However, the identified bounds are still quite large and include zero. All 95% confidence intervals of the estimates still fall in the negative range; thus, the hypothesis that self-initiated private tutoring is ineffective cannot be rejected. The reason for the latter might be the different types of private tutoring, indicating that there is a heterogeneous and probably not linear effect of private tutoring on students' achievement in reading and mathematics. Different types of tutors (e.g., retired teacher, students, older pupils) who may vary in the quality of private tutoring, the type of settings (e.g., one-to-one, two-to-one) or the frequencies for students with already high values of private tutoring (once a week or twice a week) may have a different impact on students; therefore, more research is needed to be able to rate the different forms of private tutoring and to further tighten the bounds for different sub-samples.

2.6 Appendix

Table 2-2: Self-Learning Time and Tutoring in Reading and Mathematics, PISA 2006

	No tutoring reading		Tutoring in reading		No tutoring in math		Tutoring in math	
Self-learning time per week	*N*	%	*N*	%	*N*	%	*N*	%
No time	1621	12.55	302	5.76	1082	9.95	380	5.21
Less than two hours	9195	71.16	3200	61.11	6881	63.28	4147	56.93
Two to four hours	1840	14.24	1322	25.25	2511	23.09	2113	29.01
four to six hours	223	1.72	309	5.91	327	3.01	466	6.4
six or more hours	43	0.33	103	1.97	73	0.67	178	2.44
Total	12,922		5237		10,875		7284	

Table 2-3: Self-Learning Time in Reading Per Week, PISA 2006

	Self-learning		>4 hours self-learning	
	(1)	(2)	(3)	(4)
Tutoring in reading	0.427***	0.539***	0.629***	0.552***
	(8.72)	(9.60)	(10.55)	(9.11)
Female		0.369***		0.119**
		(8.02)		(1.96)
Age		0.0566		0.143***
		(1.57)		(2.69)
Immigrant		0.218***		−0.212***
		(3.66)		(−3.03)
Parent education		−0.00566		0.0217**
		(−0.80)		(2.03)
Highest parental ISEI		−0.00258*		−0.00192
		(−1.81)		(−1.11)
Canton		−0.00290		0.0177***
		(−0.98)		(4.31)
PISA read		0.00249***		−0.00145***
		(8.46)		(−3.68)
Observations	18,159	18,159	18,159	18,159

Note: Marginal effects; *t* statistics in parentheses; (d) for discrete change of dummy variable from 0 to 1.
$^{*}p < .10$.
$^{**}p < .05$.
$^{***}p < .01$.

Table 2-4: Self-Learning Time in Mathematics Per Week, PISA 2006

	Self-learning		>4 hours self-learning	
	(1)	(2)	(3)	(4)
Tutoring in math	0.340***	0.394***	0.439***	0.378***
	(7.89)	(8.14)	(9.72)	(8.31)
Female		0.457***		−0.00356
		(11.09)		(−0.07)
Age		0.0461		0.0181
		(1.22)		(0.43)
Immigrant		0.174***		−0.113
		(2.97)		(−1.61)
Parent education		−0.0113		0.0147*
		(−1.30)		(1.65)
Highest parental I∼I		0.0000134		−0.000218
		(0.01)		(−0.13)
Canton		−0.00666**		0.00887***
		(−2.10)		(2.58)
PISA math		0.00115***		−0.00157***
		(3.48)		(−4.25)
Observations	18,159	18,159	18,159	18,159

Note: Marginal effects; *t* statistics in parentheses; (d) for discrete change of dummy variable from 0 to 1.
$^{*}p < .10$.
$^{**}p < .05$.
$^{***}p < .01$.

Chapter 3

Does Truancy Cause Bad PISA Results? What Can Be Learned About Its Effect in the Presence of Measurement Error

3.1 Introduction

Truancy is often thought of as a problematic student behavior related to low academic achievement. Although only a limited number of studies have investigated the causal relationship between intentionally skipping school lessons and academic achievement, the majority of research publications dealing with the association of truancy and achievement conclude that either low achievement precedes truancy (Bosworth, 1994) or that truancy is followed by low academic performance (Duarte, Escario, & Molina, 2011; Malcolm, Thorpe, & Lowden, 1996; Shoenfelt & Huddleston, 2006). However, a small number of studies argue that the association of truancy and high academic achievement is similarly plausible in cases where students decide to skip lessons in subjects they are very good at and tend to feel bored with (Renzulli & Park, 2000). In sum, there is agreement on the association of truant behavior and academic achievement. But the nature of this association and especially its causality has only rarely been addressed in scientific studies. Of the still very few studies that deal with causal effects of students' academic performance on truancy, there is consistent evidence that low achievement at school affects secondary students' attendance behavior and may result in truancy (Bosworth, 1994). With regard to these findings, Shute and Cooper (2015) argued that truancy can be seen as a rational student behavior depending on their academic achievement at school – in which case the assumption that high-achieving students have different reasons for being truant than low-achieving students is crucial. Accordingly, the need to know more about the causal connection between truant student behavior and academic achievement is a highly relevant research topic for educational policy makers that needs to be investigated in more detail. Adding to this, truant students are probably not a very homogeneous group of students. Since there are several plausible interconnections of academic achievement and truant behavior, one can assume that there are differential effects linking both elements together that can be found in different subgroups of students. This chapter aims to analyze data from the *Programme for International Student Assessment* (PISA) 2012 with regard to possible causal patterns connecting

student truancy in mathematics with mathematical proficiency. In order to do so, a number of challenges have to be met.

First, the fact that the group of truant students is likely to be a self-selective one needs to be taken into account when studying truancy. Second, the actual quality of data obtained through student self-reports is somewhat unclear. There is reason to assume that truancy is sometimes misreported in terms of exaggerated frequencies or denied behavior, meaning that students' responses to questionnaire items may involve both false positive and false negative classifications of students as truants or non-truants (e.g. Siziya, Muula, & Rudatsikira, 2007). For this reason, we propose considering a way of analyzing student self-reported data on truancy which explicitly tackles the problem of self-selection and measurement error. In this study, we apply a nonparametric bounds method (Kreider et al., 2012; Manski, 1990, 1997; Manski & Pepper, 2000) in order to calculate the upper and lower bounds of the treatment effect, making as few assumptions as possible. This partial identification approach relies on observed sample averages and differences in these averages between the treated and untreated groups, therefore producing a range of the average treatment effect under different assumptions instead of a point estimate. We can successively layer stronger assumptions and can therefore show how these assumptions influence the causal effect. We start by making no assumptions, and then we successively impose weak nonparametric assumptions to tighten the bounds step by step. Furthermore, we allow for different misclassification of truancy in order to account for measurement error in students' participation status.

This chapter is guided by the idea that many large-scale student assessments, such as PISA, measure students' proficiency levels and use them as indicators of the quality of educational systems. A specific focus is set on high-performing students (achieving a mathematics competence in PISA 2012 of at least level 4 out of 6). The main research question is whether self-reported truant behavior has an impact on the probability of reaching at least competence level 4 in the PISA mathematics test. This approach follows the assumption that many proficiencies measured in large-scale student assessments are at least mostly acquired at school. This holds especially true for mathematics, which is hardly learned in out-of-school settings. We pick up on this and take a closer look at students who decide not to attend every mathematics class that is on their schedule. One likely consequence of being truant in mathematics is a low test score on the PISA mathematics test. To date, very little is known about the treatment effect that adolescent truancy may have with regard to academic achievement.

3.2 Truancy as a Research Topic

When, as in this study, truancy is measured using self-reported student data, it is best defined as "absences which pupils themselves indicated would be unacceptable to teachers" (Wilson,

Malcolm, Edward, & Davidson, 2008, p. 3). The most central aspect of this definition is that truancy implies an active decision of students to skip a lesson or a school day, knowing that they should have attended school.

3.2.1 Specific Challenges of Studying Truancy

Aside from the issues of selection into the group of truant students (see the following section) and of potential misreporting of truancy in student self-reports (see "Classification Error Assumptions"), studies dealing with truancy generally face a number of challenges that may impact the quality of the results. For example, in countries like Germany, access to school records or files is denied due to privacy regulations. Hence, researchers cannot use school records as a source of data. School files are often thought to be somewhat objective in a way that they are comprehensive and complete. However, from a scientific point of view, using school files as a data source to investigate the prevalence of truancy would not be very valid in Germany, as each absence has to be legitimated by a written parental excuse, several days of absence have to be justified by a doctor's note, and irregularities have to be investigated. Thus, schools in general will leave no missed lesson unexcused by either a written excuse by parents or a doctor's note, so that truancy will disappear from school records as soon as there is a formal justification. The fact that schools have to keep their records straight makes these records a presumably invalid source of data for research on truancy. Adding to this, any excused absence can disguise truancy, as only the students themselves know when they have been truant.

3.2.2 The Complex Relationship between Truancy and Academic Achievement

Although there seems to be a vast consensus that there is an association between truancy and academic achievement (Bosworth, 1994; Duarte et al., 2011; Malcolm et al., 1996; Shoenfelt & Huddleston, 2006; Vaughn, Maynard, Salas-Wright, Perron, & Abdon, 2013), two aspects are still quite unclear: the direction of this relationship and the role of confounding variables. In other words, it is plausible that the performance of students being truant may decrease due to the missed lessons, but it is also reasonable to assume that especially those students whose performance at school is already rather weak are truant. Furthermore, there is evidence that truancy may be related to high academic achievement as well, if students decide to skip classes due to boredom or a feeling of being underchallenged in certain subjects (Renzulli & Park, 2000; Sälzer, Trautwein, Lüdtke, & Stamm, 2012). With regard to possibly confounding variables, student performance often cannot be fully separated from immigrant or socioeconomic status (SES) or the school track, so the predictive value of academic achievement as a predictor of truancy is highly dependent on the model and technique used for data analysis. Furthermore, there are very few studies that have investigated causal effects of truancy on academic achievement.

With regard to the direction of the relationship between truancy and achievement, two studies in particular examined this association and controlled for a differentiated background model. In a recent longitudinal study, Veenstra, Lindenberg, Tinga, and Ormel (2010) had a sample of 2,230 Dutch students, five percent of whom were persistent truants. The authors were able to show that most of the persistent truants had low academic achievement. Henry (2007) looked at the characteristics of eighth and 10th graders who were truant and showed that high academic achievement was related to low truancy rates. Student SES and their immigrant status as potential confounding variables have mostly been found to be negatively related to truancy (Considine & Zappalà, 2002; Dunkake, 2006; Rothman, 2001, 2004). That is, immigrant students and those from lower SES families tend to be truant more often than non-immigrant students and those from higher SES families. However, immigration background and student SES are often confounded and therefore not clearly separable; moreover, results have not been fully consistent across studies. One plausible reason may be the different possibilities of operationalizing SES (Ehmke & Siegle, 2005; Mallinson, 2007; Warren, Sheridan, & Hauser, 1998).

When studying the relationship between student achievement and truancy, one needs to take into account that at least part of this relationship may be attributed to students being sorted into different school tracks according to their prior academic achievement. Most industrialized countries worldwide use some form of achievement-based grouping of students (also known as “streaming” or “tracking”). In-depth discussions of potential effects of tracking inequality can be found elsewhere (Weissbrodt, 2007). Many critics of between-school tracking have argued that being assigned to a low-achieving group has negative effects on student motivation. Although these effects are not necessarily automatic (e.g., Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006), it seems quite plausible that truancy rates are higher in lower-track schools. Reid (1999) repeatedly stated that the degree of truancy changes over time, and one cannot assume a linear relationship between age or grade level and truancy. Moreover, Bongers, Koot, van der Ende, and Verhulst (2004) analyzed data from a longitudinal multiple birth study of children aged 4–18 years (N = 2,076) and found truancy increasing with grade level, using multilevel growth curve analyses. In this study, we make a subgroup analysis for the school tracks in order to take into account that truancy may have a different impact on students’ outcome in different tracks. Hence, being sorted at around age 10 is a central aspect of one’s school biography (Sälzer et al., 2012).

Studies focusing on the causal effect of truancy on academic achievement have mostly found that attending lessons – that is, more instructional time – is positively related to student achievement (Bellei, 2009; Cortes & Goodman, 2014; Lavy, 2015; Rivkin & Schiman, 2015). However, this is a district-specific variable, and in this research we study a student-specific variable: truancy. We

anticipate that truancy will therefore have a different impact on different students.[1] Some truant students falling behind the pace of instruction will partially hinder their ability to learn new things. Others will use the hours absent from school productively and gain a great deal by studying at home or spending time with subject-matter content. In this regard, we assume that the causal effect of truancy may be heterogeneous and even stronger than that of instructional time.

To date, few studies deal with the causal effects of truancy on student performance. Truancy has mostly been found to cause low academic performance or vice versa (Duarte et al., 2011; Malcolm et al., 1996; Shoenfelt & Huddleston, 2006). Yet in contrast, there is also scarce evidence that high-achieving, underchallenged students skip classes they feel are not much use to them and that returns to attendance in class are likely to vary across the distribution of students (Arulampalam, Naylor, & Smith, 2012; Dobkin, Gil, & Marion, 2010). More concretely, these studies indicate that a negative effect of truancy on academic achievement can only be found for high-achieving students. Conclusively, there is argument for both a negative and positive impact of truant behavior on academic performance. While truant behavior on the one hand means ignoring school lessons that provide an opportunity to learn (McDonnell, 1995), it can also be the result of a rational choice of students (Shute & Cooper, 2015) who excel at school and are not in need of attending every lesson in order to perform well.

3.3 The Present Study

This chapter investigates the causal impact of truancy on students' academic performance. Using a grade-based national oversample of PISA 2012 in Germany, the main objective is to estimate the average treatment effect (ATE) of being truant on high achievement in PISA, taking into account both self-selection into treatment and misreporting of treatment status. We investigate bounds around the ATE calculated by the mean outcome that would result if all students had skipped mathematics classes in the current school year versus the mean outcome if none of the students had skipped mathematics classes. We assume that the group of students who indicated intentionally skipped mathematics classes in the current school year have a lower probability of being a high-achieving student in mathematics (measured by means of the PISA 2012 competence score in mathematics) than students who regularly attend mathematics lessons (hypothesis 1). Further, we anticipate to find heterogeneous treatment effects for students in different school tracks (hypothesis 2). Dealing with the possibility of measurement error, we report the results first in the absence of misclassification and second under the assumption of different levels of misclassification. We

[1] Goodman (2014) has shown in a natural experiment on the impact of instructional time on achievement that the coordination of students is the central challenge; therefore, a disruption such as absence that affects different students at different times (in the case of snowfall, unintended, but in the case of truancy, intended) seems to be the important factor for student math achievement.

suppose that under the assumption of measurement error, the calculated bounds will widen and include zero – and thus allow for an increasing probability of truant students being high-achievers (hypothesis 3).

3.4 Data Source and Sample[2]

Sample. The data analyzed in this chapter come from a grade-based national oversample of PISA 2012 in Germany (Prenzel, 2013). The students analyzed attended one of two complete ninth-grade classrooms per participating school ($N_{(class)}$ = 220, $N_{(student)}$ = 9998). Class size was 23.9 students on average (*SD* = 4.25). 49% of the sample were male, students were on average 15.3 years old (*SD* = .75).

A characteristic feature of the German educational system is between-school tracking. In Germany, students are tracked earlier than their counterparts in almost all other countries of the Organisation for Economic Co-operation and Development (OECD) except Austria (Sälzer et al., 2012). On average, students at age 10 complete primary schooling and start secondary schooling at one of several school types. Based on the sampling procedures in PISA 2012 (OECD, 2014), five school types can be reported within the German PISA sample: Hauptschule, offering secondary level I education (five years); Cooperative Secondary School, offering secondary levels I and II education (five to nine years); Realschule, offering secondary level I education (six years); Comprehensive School, offering secondary levels I and II education (five to nine years); and Gymnasium, an academic track offering secondary levels I and II education (eight or nine years). The schools included in this chapter are classified as one of the five secondary school types for which we have representative data in Germany. Vocational schools and special needs schools are not included in this sample due to their very small proportion of the population. Participating schools were located in all 16 federal states of Germany.

Schools were sampled as a stratified random sample according to the international sampling definitions (Heine, Sälzer, Bochert, Sibberns, & Mang, 2013). First, a so-called sampling frame consisting of a complete list of all schools possibly attended by 15-year-old students (the PISA target-population) was established. In order to obtain a sample representative for the population of 15-year-old students attending a school in their respective country of test, each country is divided into several areas, so-called explicit strata. A stratum is a partition of the population which is

[2] We make use of the PISA mathematics proficiency levels. The range of difficulty of the tasks is represented by six levels of mathematical proficiency. The levels range from the lowest, Level 1, to the highest, Level 6. Descriptions of each of these levels have been generated, based on the framework-related cognitive demands imposed by tasks that are located within each level, to describe the kinds of knowledge and skills needed to successfully complete those tasks, which can then be used as characterizations of the substantive meaning of each level. We distinguish between high proficiency students (level 4 or higher) and low proficiency students.

defined according to national specifications that may have an impact on the representativity of the results (OECD, 2014). In Germany, the federal structure of the educational system made it necessary to define a stratum for each of the 16 federal states. Conceptually, only Gymnasium schools have a common history and curriculum within Germany, whereas the other four school types vary in terms of their history and curricula across the 16 federal states. According to this multi-tier school system, the explicit strata representing federal states were subdivided into implicit strata so that in each federal state, all prevalent school types were present in the sample (OECD, 2014) and Gymnasium schools were separated from non-Gymnasium schools.

Based on this stratification, a random sample of schools was drawn within each stratum (federal state). After the sample representing the target population of 15-year-old students was drawn, the German national oversample of two complete grade-9 classrooms followed. In each school participating in PISA 2012, two classrooms of the so-called modal grade (i.e. the grade-level which is attended by the majority of 15-year-old students) were randomly selected for participation in the PISA test.

Test of students' proficiency in mathematics. Mathematical literacy is one of three cognitive domains captured in the PISA study. In the most recent cycle of PISA, PISA 2012, mathematics was the major domain for the second time after PISA 2003 (OECD, 2013b). This means that about half of the test units were classified as mathematical tasks, whereas the two minor domains of reading and scientific literacy contributed about 25% of the test units, respectively. PISA focuses on an aspect of mathematical proficiency which is especially relevant for adolescents approaching the end of compulsory schooling. According to the PISA 2012 Assessment Framework, mathematical literacy is defined as follows: "Mathematical literacy is an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens" (OECD, 2013b, p. 25). Building on this definition, mathematical literacy in PISA is structured according to content areas, processes and contexts in which mathematical skills are applied. While four different content areas are distinguished (change and relationship, space and shape, quantity as well as uncertainty and data), the processes are defined so that they are operationalizing the concept of mathematical literacy in the form of tasks that measure student's mathematical proficiency. All the processes are applied in tasks referring to the four content areas: formulating mathematical situations, applying mathemical concepts, facts, procedures and conclusions as well as interpreting, applying and evaluating mathematical results. Both the content areas and the mathematical processes are situated in one of four contexts in each mathematics task in the PISA test: personal, professional, scientific or societal.

The PISA tasks are grouped in so-called clusters, combining approximately 30 minutes of test units per domain. Each cluster contains tasks only from one domain and each student is assigned four clusters for the PISA test, followed by a student questionnaire. A typical PISA task consists of one stimulus text establishing the context of the task and several test items which have to be independent from each other. Item formats include simple and complex multiple choice items as well as open-ended questions.

Results reported in PISA generally feature country mean scores on the proficiency scales in mathematics, reading and science as well as a descriptive illustration of proficiency or competence levels (OECD, 2014). PISA proficiency levels are defined as a range of values on the PISA proficiency scales, representing tasks which students are typically able to solve. In PISA 2012 mathematics, six proficiency levels are distinguished, where level 1 is the lowest and level 6 is the highest and most demanding. Students at a certain level are assumed to be able to solve tasks at lower levels, meaning that the proficiency levels are ordered hierarchically.

Since the aim of PISA is comparing student performance and its context internationally, the concept of validity in PISA genuinely refers to the question of cross-cultural comparability. Cross-cultural construct equivalence is tested using structural equation modeling and Multiple Group Confirmatory Factor Analysis (MGCFA, cf. OECD, 2013b) in order to describe the validity of cross-country comparisons. Both item and method bias are analyzed in order to correct for a lack of comparability if necessary (OECD, 2013b).

3.5 Methodology: Nonparametric Bounds

Assessing the impact of being truant on students' academic performance faces fundamental identification problems. Being truant may be endogenous and correlated with at least some unobservable personal and family characteristics. One difficulty is that the outcomes a person would produce or experience under all treatments cannot be observed. At most, one can observe the outcome that a person experiences under the treatment actually received. We therefore only observe the PISA competences of truant and non-truant students, but cannot observe their *potential* outcome, assuming that they have been truant or not. This is referred to as the so-called selection problem.

Second, truancy is a juvenile behavior that is known to be socially undesirable. Accordingly, we have to assume that study results that refer to student self-reports will include at least some rates of measurement error. On the one hand, students' reports on their individual truancy are in part exaggerated if the image of truants is positive; but, on the other hand – what we will argue is the case – truancy is a topic that involves the fear of being uncovered or caught and then punished (e.g.

Corville-Smith, Ryan, & Adams, 1998) and therefore students might understate the true amount of truancy. Thus, the credibility of empirical analysis depends on the strength of the underlying assumptions, taking into account that there is self-selection into truancy and potential misclassification.

The partial identification approach developed in this chapter (see, for example, De Haan, 2015; Hof, 2014; Kreider et al., 2012; Manski, 1990; Millimet & Jayjit, 2015) allows for the evaluation of bounds on the average treatment effect of truancy under different assumptions. This approach permits us to successively layer stronger identification assumptions and therefore elucidate how assumptions shape inferences about the causal effect of truancy.

By combining the distribution of a random sample with prior information, we intend to identify the ATE. The analysis aims to identify the ATE of being truant in mathematics classes on the probability of being a high achieving student in mathematics; that is,

$$\text{ATE}\,(1,0) \;=\; \text{P}[\text{y}(1) = 1|\text{x}] - \text{P}[\text{y}(0) = 1|\text{x}],$$

where y(1) denotes the outcome with treatment, y(0) is the outcome without treatment, and P[.] denotes the probability of the argument being true. The ATE is the causal effect of being truant on the probability of being a high achieving student and is calculated by the mean assuming *all* students were truant minus the mean if *no* student was truant. Conditioning on covariates X helps in this nonparametric approach only to define subpopulations (Kreider et al., 2012), which is dropped for simplicity in the following derivations. Denoting treatment status by the indicator z* (where z = 1 denotes that a student had been truant in the last school year, and z = 0 otherwise), the observed outcome for a particular student is given by $y = z^{*}y(1) + (1 - z^{*})y(0)$. Using the Law of Total Probability (following Almada, McCarthy, & Tchernis, 2015; Millimet & Jayjit, 2015), we decompose

$$\text{P}[\text{y}(1) = 1] \;=\; \text{P}[\text{y}(1) = 1|z^{*} = 1]\,\text{P}(z^{*} = 1) \;+\; \text{P}[\text{y}(1) = 1|z^{*} = 0]\,\text{P}(z^{*} = 0),$$

$$\text{P}[\text{y}(0) = 1] = \text{P}[\text{y}(0) = 1|z^{*} = 1]\text{P}(z^{*} = 1) + \text{P}[\text{y}(0) = 1|z^{*} = 0]\text{P}(z^{*} = 0).$$

We face two problems that must be addressed: First, the sampling process alone cannot identify the counterfactual probabilities P[y(1) = 1|z* = 0] and P[y(0) = 1|z* = 1] (selection problem) and second, the true treatment status may not be observed for all students (measurement error). Instead of observing z* we observe z, and when $t \neq t^{*}$ for some students, we cannot observe the true treatment status. We will not know if a student's reported truancy status is their actual truancy status. We therefore denote a* as an indicator for whether reported treatment is accurate or not, where a* = 1 if z* = z and zero otherwise. Previous research has shown that P[y(1) = 1] may be decomposed as follows:

$$P[y(1) = 1] = [P(y = 1, z = 1) - \theta_1^+ + \theta_1^-] + P[y = 1|z^* = 0]\,[P(z = 0) + (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)],$$

where $\theta_j^+ \equiv P(y = 1, z = 1, a^* = 0)$ and $\theta_j^- \equiv P(y = 1, z = 0, a^* = 0)$ for students' realized outcome j = 1,0 represent the unconditional misreporting probabilities (Almada et al., (2015); Kreider et al., (2012); McCarthy, Millimet, & Roy, (2015); see also the following section "Classification Error Assumptions").

The analysis starts with investigating the effect of truancy without imposing assumptions (worst-case [WC] bounds). Then the analysis imposes weak nonparametric assumptions to tighten the bounds; it utilizes monotone treatment selection (MTS) assumption which states that being truant is weakly monotonically related to poor academic outcome. We consider bounds under the assumption of a monotone instrument variable (MIV), which implies that the probability of a high academic outcome is weakly monotonically increasing with an observed covariate. This chapter uses students' ESCS,[3] an indicator for student economic, social, and cultural status as MIV, and the underlying MIV assumption is that students with a high ESCS value are less likely to be truant than students with a low ESCS value.

3.5.1 Classification Error Assumptions

It seems plausible that there is a correlation between the students' probabilities of misreporting and observed covariates. Therefore, measured point estimates are subject to some degree of bias and inconsistency. We thus adapt the explained bounds method to provide a complete picture of the rank of possible effects of truancy including the possibility of measurement error.

Thus, when considering measurement error, we discuss two cases, based on Gundersen and Kreider (2008), McCarthy et al. (2015), and Tourangeau, Rips, and Rasinski (2000).[4] In the first case, no structure is imposed on the pattern of reporting errors. We refer to this as arbitrary errors. In the second case, misreporting has a structure (see Table 3-1): 1) False negatives, which result from the students' failure to report truancy when in fact they were truant. 2) False positives, which result when students report having played truant when in fact they were not truant. However, after

[3] This index reflects the economic, social, and cultural resources of parents and is an indicator for the social background of the student.

[4] Tourangeau et al. (2000) found that respondents' concern about reporting socially undesirably behavior (such as truancy) is determined both by situational aspects and by the respondents' assumptions about the confidentiality of their data. The authors argue that sensitive behaviors are misreported systematically and they distinguish between two types of misreporting: overreporting and underreporting. While studies have shown that rates of sensitive behavior captured through standardized questionnaires in a self-administered setting like a classroom are higher than in a face-to-face interview situation at home (Oberwittler & Naplava, 2002), external validation strategies frequently fail due to large discrepancies between information from different sources (De Los Reyes & Kazdin, 2004; Eisner & Ribeaud, 2007; Reyes & Kazdin, 2006).

checking the results for arbitrary errors, we elected to assume the second case, with no false positive errors. Considering that Oberwittler and Naplava (2002) found that in self-administered settings the reported rates of socially undesirable behavior are higher than in interview data, we still have to take into account that the PISA student questionnaire was administered in a classroom setting and the students' fear of being detected as a truant through their responses cannot be ruled out. Hence, we assume that students will somewhat underrate their individual truancy when asked about it because they fear the potential consequences in their school. This means it is implyed that in this chapter students' responses can be treated as including no false positives and thus no students having reported being truant when in fact they were not.

Table 3-1: Classification of Measurement Error

	Student played truant	Student did not play truant
Student reports having played truant	true positive	false positive
Student reports not having played truant	false negative	true negative

Consequently, as explained, we have a selection problem, as we cannot identify the counterfactual probabilities directly and the true treatment cannot be observed for all individuals. The unconditional misreporting probabilities are as follows:

$\theta_1^+ \equiv P(y = 1, z = 1, a^* = 0)$ -> fraction of observations which are false positive with y = 1

$\theta_0^+ \equiv P(y = 0, z = 1, a^* = 0)$ -> fraction of observations which are false positive with y = 0

$\theta_1^- \equiv P(y = 1, z = 0, a^* = 0)$ -> fraction of observations which are false negative with y = 1

$\theta_0^- \equiv P(y = 0, z = 0, a^* = 0)$ -> fraction of observations which are false negative with y = 0

Additional details on the derivation of bounds under alternative misreporting assumptions are available in Almada et al. (2015), Kreider et al. (2012), Manski and Pepper (2000), and McCarthy et al. (2015).

Although estimated misreporting rates have not yet been quantified for studies measuring truancy through self-reports, misreporting has to be seen as prevalent when asking students about their truancy behavior (Tourangeau et al., 2000). The presence of misreporting widens the ATE bounds relative to the assumption of no misreporting; in addition, the extent depends on the assumptions of the level of misreporting in the data. Formally, we follow Gunderson and Kreider (2008) and McCarthy et al. (2015) with the following assumptions:

Upper bound error rate: $P(a^* = 0) \leq Q$.

No false positive: if z = 1, then $a^* = z = 1$,

with Q as the upper bound on the degree of misclassification. For arbitrary measurement errors, only assumption 1) is imposed, for the case of no false positive classification, assumptions 1) and 2) are imposed. The value of Q varies from 0 (no measurement error) to 10%. We assume misreporting rates of 1%, 2%, 5%, and 10%.[5]

Assumption one implies

$$0 \leq \theta_1^- \leq min\{Q, P(y=1, z=0)\} \equiv \theta_1^{UB-}$$

$$0 \leq \theta_0^- \leq \min\{Q, P(y=0, z=0)\} \equiv \theta_0^{UB-}$$

$$0 \leq \theta_1^+ \leq \min\{Q, P(y=1, z=1)\} \equiv \theta_1^{UB+}$$

$$0 \leq \theta_0^+ \leq min\{Q, P(y=0, z=1)\} \equiv \theta_0^{UB+}$$

And $\theta_1^- + \theta_1^+ + \theta_0^- + \theta_0^+ \leq Q$

And assumption two implies $\theta_1^+ = \theta_0^+ = 0$.

3.5.2 Exogenous Selection

The assumption of exogenous selection implies

$$\mathrm{P}[\mathrm{y}(1) = 1] = [\mathrm{P}(\mathrm{y}(1) = 1 | z^* = 1)] = [\mathrm{P}(\mathrm{y}(1) = 1 | z^* = 0)],$$

$$\mathrm{P}[\mathrm{y}(0) = 1] = [\mathrm{P}(\mathrm{y}(0) = 1 | z^* = 1)] = [\mathrm{P}(\mathrm{y}(0) = 1 | z^* = 0)].$$

And the ATE is given by

$$ATE = [\mathrm{P}(\mathrm{y}(1) = 1] - [\mathrm{P}(\mathrm{y}(0) = 1] = \mathrm{P}[\mathrm{y} = 1 \ | z^* = 1)] - \mathrm{P}[\mathrm{y} = 1] \ | z^* = 0)],$$

which is point identified in the absence of measurement error.

Allowing for measurement error z* is unobserved:

$$P[y(1) = 1] = \frac{P(y=1, z=1) + \theta_1^- - \theta_1^+}{P(y=1) + (\theta_1^- + \theta_0^-) - (\theta_1^+ + \theta_0^+)}$$

[5] To our knowledge, no studies have evaluated the misreporting rate for truant behavior. However, misreporting in survey reports for socially undesirable behavior is quite common (Kuntsche & Labhart, 2012; Monk, Heim, Qureshi, & Price, 2015; Tourangeau & Yan, 2007), and thus the applied misreporting rates in this study can be seen as a lower bound.

The ATE bounds are given by

$$UB_{ATE} = UB_{P[y(1)=1]} - LB_{P[y(0)=1]},$$

$$LB_{ATE} = LB_{P[y(1)=1]} - UB_{P[y(0)=1]},$$

with UB and LB denoting the upper and lower bounds, respectively.

With arbitrary errors, the bounds are as follows (following Kreider & Pepper, 2007; McCarthy et al., 2015; Millimet & Jayjit, 2015):

$$UB_{ATE} = sup_{c\epsilon(0,\min[Q,P(y=1,z=0)]])}\left[\frac{P[y=1,z=1]+c}{P(z=1)+2c-Q} - \frac{P[y=1,z=1]-c}{P(z=0)-2c+Q}\right],$$

$$LB_{ATE} = inf_{b\epsilon(0,\min[Q,P(y=1,z=4)]])}\left[\frac{P[y=1,z=1]-b}{P(z=1)-2b+Q} - \frac{P[y=1,z=1]+b}{P(z=0)+2b-Q}\right].$$

And with the assumption of no false positives, the bounds around the ATE are

$$\frac{P[y=1,z=1]}{P(z=1)+\theta_0^{UB-}} \leq P[y(1)=1] \leq \frac{P[y=1,z=1]+\theta_1^{UB-}}{P(z=1)+\theta_1^{UB-}},$$

$$\frac{P[y=1,z=0]-\theta_1^{UB-}}{P(z=0)-\theta_1^{UB-}} \leq P[y(0)=1] \leq \frac{P[y=1,z=0]}{P(z=0)-\theta_0^{UB-}}.$$

3.5.3 Worst-Case Bounds

We are not able to identify the unobservable counterfactual E[y(1)|z* = 0] or E[y(0)|z* = 1] from the data without imposing very strong and probably implausible assumptions. Therefore, this analysis replaces the unobserved by its bounds, and these are for each treatment the worst-case bounds (no-assumptions bounds). With a binary outcome (being a high-achieving student), the counterfactual expected values necessarily fall in the [0, 1] intervals, formally P[y(1) = 1| z* = 0], P[y(0) = 1|z* = 1] ε[0, 1]. The bounds without measurement error are as follows:

$$\mathrm{P}[y=1,z^*=1] \leq \mathrm{P}[y(1)=1] \leq \mathrm{P}[y=1,z^*=1] + \mathrm{P}(z^*=0)],$$

$$\mathrm{P}[y=1,z^*=0] \leq \mathrm{P}[y(0)=1] \leq \mathrm{P}[y=1,z^*=0] + \mathrm{P}(z^*=1)].$$

The bounds are sharp bounds; the width always equals unity and includes zero (Manski, 1990). So even when the direction of the effect is not known, these bounds provide some potentially useful information as the extreme values are excluded.

$$P[y=1,z=1]-\theta_1^+ + \theta_1^- \leq P\,[y(1)=1] \leq P[y=1,z=1]P(z=0)+\theta_0^+ - \theta_0^-$$

$$P[y=1,z=0]+\theta_1^+ - \theta_1^- \leq P\,[y(0)=1] \leq P[y=1,z=0]P(z=1)-\theta_0^+ + \theta_0^-$$

With arbitrary errors, the bounds on ATE[6] are

$$UB_{ATE} = P[y=1,z=1]+P(z=0)+\min\{Q,\theta_0^{UB+} + \theta_1^{UB-}\} - P[y=1,z=0)\,,$$

$$LB_{ATE} = P[y=1,z=1]-P(z=1)-\min\{Q,\theta_1^{UB+} + \theta_0^{UB-}\} - P[y=1,z=0).$$

With the assumption of no false positive, we have tighter bounds

$$UB_{ATE} = P[y=1,z=1]+P(z=0)+\theta_1^{UB-} - P[y=1,z=0),$$

$$LB_{ATE} = P[y=1,z=1]-P(z=1)-\theta_0^{UB-} - P[y=1,z=0).$$

3.5.4 Monotone Treatment Selection

Applying the first assumption, we assume that the counterfactual outcome is smaller for those students who were truant (z* = 1) than for those who were not truant (z* = 0). We therefore assume a negative self-selection into truancy; in other words, those who were truant are more likely to be low-achievers with: P[y(1) = 1|z* = 1] ≤ P[y(1) = 1|z* = 0] and P[y(0) = 1|z* = 1] ≤ P[y(0) = 1|z* = 0]. Therefore, MTS assumes that high-achieving students are less likely to be truant.

To tighten the bounds, we make use of the following (exemplarily for the lower bound): We can observe the mean achievement of students who are not truant. Under MTS assumption, this mean achievement will never be lower than the mean achievement for students actually being truant. Therefore, the mean realized outcome for truant students is the lower bound. The MTS bounds are the following:

$$P[y(1)=1]\epsilon\left[\frac{P(y=1,z=1)-\theta_1^+ + \theta_1^-}{P(z=1)-(\theta_1^+ + \theta_0^+)+(\theta_1^- + \theta_0^-)},\ P(z=0)+P[y=1,z=1]+\theta_0^+ - \theta_0^-\right]$$

$$\equiv LB_1\,, UB_1,$$

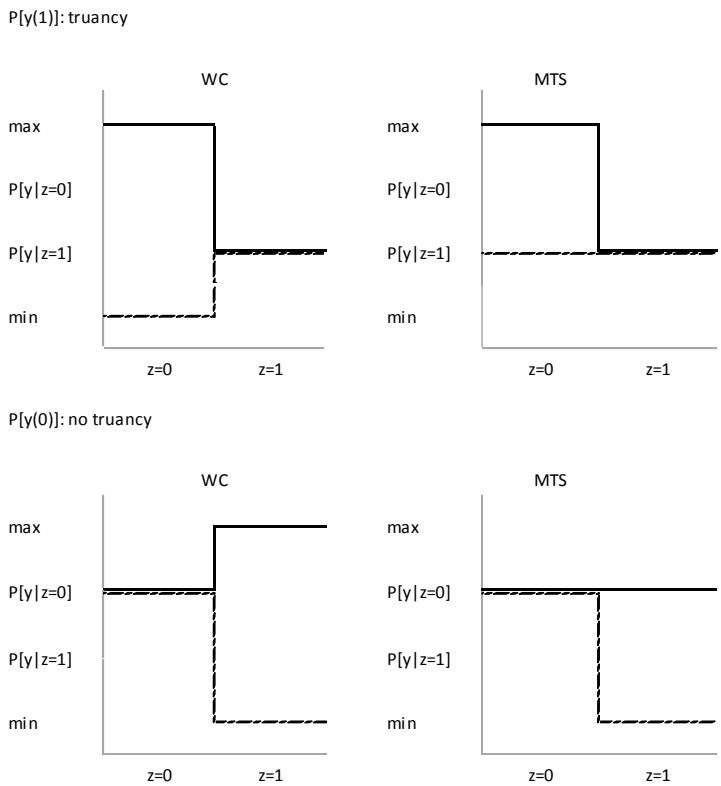
$$P[y(0)=1]\epsilon\left[P[y=1,z=0]+\theta_1^+ - \theta_1^-\ ,\frac{P(y=1,z=0)+\theta_1^+ - \theta_1^-}{P(z=0)+(\theta_1^+ + \theta_0^+)-(\theta_1^- + \theta_0^-)}\right] \equiv LB_0\,, UB_0.$$

The two illustrations on the right in Figure 3-1 show how the MTS assumption tightens the bounds. For example, the mean potential outcome of the treatment P[y|z = 1] is only observed for truant

[6] The ATE (P[y(1)] - P[y(0)]) is calculated as follows: The lower bound on P[y(1)] minus the upper bound on P[y(0)] is the lower bound of the ATE. The upper bound on P[y(1)] minus the lower bound on P[y(0)] is the upper bound of the ATE. $ATE \in [LB_1 - UB_0\,, UB_1 - LB_0]$

students. The mean potential outcome of truant behavior is not observed for students who were not truant and can be between y min and y max. Under the assumption of MTS, the mean outcome for non-truant students will not be lower than the mean outcome observed for students who were actually truant. P[y|z = 1] can therefore be used as the lower bound.

Figure 3-1: How the MTS Assumption Works



Source: Figure based on De Haan (2011) and Hof (2014)

3.5.5 Monotone Instrument Variable (MIV)

In contrast to an IV assumption with mean independence, the MIV assumption allows a weakly monotone positive relationship between v and the mean potential outcome. This innocuous MIV assumption allows for a direct impact of the socioeconomic background on students' academic achievement as long as the effect is not negative.

Let v denote the monotone instrument, and P[y(1) = 1] and P[y(0) = 1] are non-decreasing in v. $u_1 < u < u_2$, the MIV assumption implies:

$$P[y(1) = 1|v = u_1] \leq P[y(1) = 1|v = u] \leq P[y(1) = 1|v = u_{2]},$$

$$P[y(0) = 1|v = u_1] \leq P[y(0) = 1|v = u] \leq P[y(0) = 1|v = u_{2].}$$

Following Kreider et al. (2012) we combine the MIV assumption with the MTS assumption, as the MIV alone has no identification power. The bounds are given by

$$P[y(1) = 1]\epsilon \left[\frac{P[y = 1, z^* = 1]}{P(z^* = 1)}, P(z^* = 0) + P[y = 1, z^* = 1]\right],$$

$$P[y(0) = 1]\epsilon \left[P[y = 1, z^* = 0], \frac{P[y=1,z^*=0]}{P(z^*=0)}\right].$$

To calculate the bounds, the sample is split into four ESCS cells. Weighted averages of the estimates of the UB and LB across the four cells yield joint MTS-MIV bounds (see also McCarthy et al., 2015; Millimet & Jayjit, 2015). We use Kreider and Pepper's (2007) nonparametric finite sample bias-corrected MIV estimator (for discussion of this estimator see McCarthy et al., 2015).

3.6 Results

The results section is organized along the three hypotheses we formulated above. Hypotheses 1 and 2 are presented in two accordingly named sections; hypothesis 3, referring to the effect of measurement error, relates to both sections and thus has no explicit extra section. The partial identification approach allows us to evaluate bounds on the ATE of being truant under different assumptions about the selection and measurement error problems. We therefore present the results a) under different assumptions concerning the selection problem and b) with and without classification error (arbitrary errors and no false positive).

Hypothesis 1. We first present results of the overall effect of truancy on the proficiency level in PISA 2012 mathematics. This analysis tests hypothesis 1, truancy in mathematics causing a low proficiency level on the PISA 2012 mathematics test. Table 3-2 presents the full-sample non-parametric bounds on the estimated ATE of truancy on students' academic performance. The range in the brackets reflect the ATE bounds under the different assumptions (Exogenous Selection, No Monotonicity Assumption [Worst-Case Selection], MTS, and MIV with MTS). In each section of the table, the results for different misreporting assumptions are shown. The arbitrary misreporting column allows for misreporting in both directions, bounded with the assumed error rates. The non-false positive column assumes that no student reported having been truant when in fact he or she has not. For example, assuming exogenous selection and no misreporting (for illustration, see Figure 3-2), the estimated ATE is -0.144. This implies that truant students are 14.4% more likely to be low-achieving students compared with students who never were truant. The impact of misreporting is profound. If as little as two percent of the sample misreport their truancy behavior, the direction of the ATE even under exogenous selection is no longer clearly negative (see also Figure 3-2). That said, it is obvious that the association between truant behavior and student academic achievement is not robust to even small amounts of misreporting. Further, the assumption

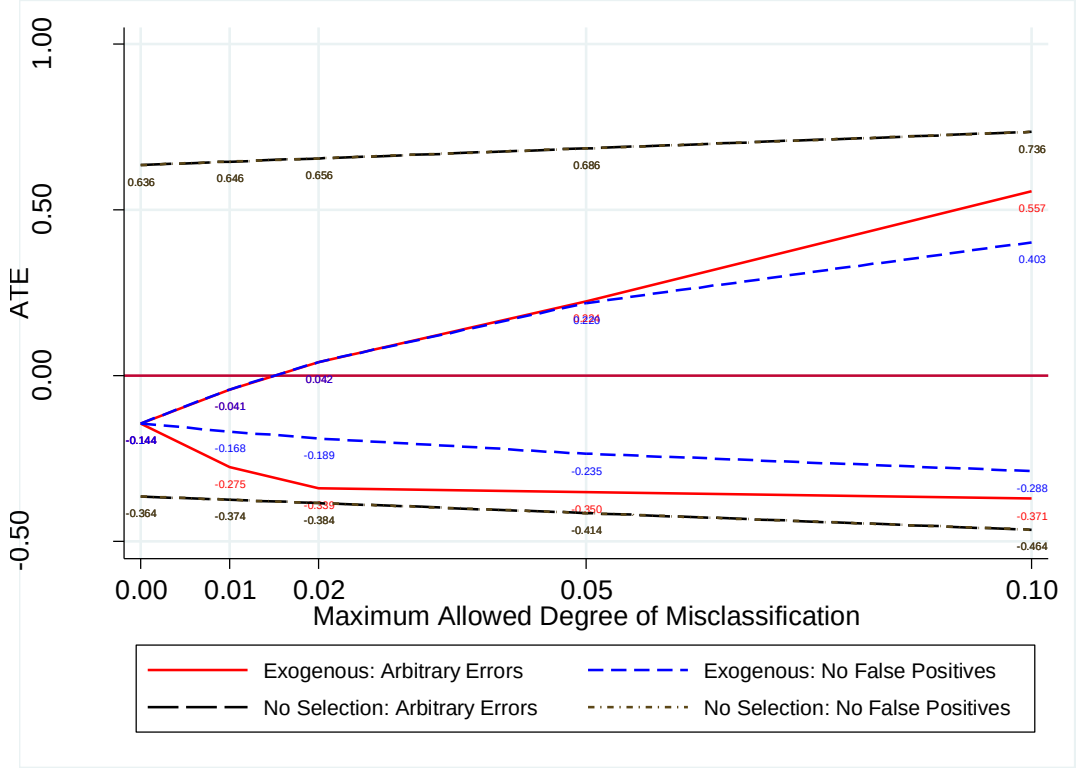
of no false positive provides some identifying information relative to the assumption of arbitrary errors.

Table 3-2: Bounds on the ATE of Truancy on High Academic Outcome[7]

	Arbitrary errors		No false positives	
Error rate	Bounds	95% CI	Bounds	95% CI
Exogenous selection model				
0	[-0.114, -0.144]	[-0.166, -0.121]	[-0.144, -0.144]	[-0.166, -0.121]
.01	[-0.275, -0.041]	[-0.300, -0.018]	[-0.168, -0.041]	[-0.190, -0.018]
.02	[-0.339, 0.042]	[-0.350, 0.066]	[-0.189, 0.042]	[-0.210, 0.066]
.05	[-0.350, 0.224]	[-0.361, 0.283]	[-0.235, 0.220]	[-0.254, 0.245]
.1	[-0.371, 0.557]	[-0.383, 0.583]	[-0.288, 0.403]	[-0.305, 0.428]
No monotonicity assumptions (worst-case selection)				
0	[-0-364, 0.636]	[-0.375, 0.650]	[-0.364, 0.636]	[-0.375, 0.650]
.01	[-0.374, 0.646]	[-0.385, 0.660]	[-0.374, 0.646]	[-0.385, 0.660]
.02	[-0.384, 0.656]	[-0.395, 0.670]	[-0.384, 0.656]	[-0.395, 0.670]
.05	[-0.414, 0.686]	[-0.425, 0.700]	[-0.414, 0.686]	[-0.425, 0.700]
.1	[-0.464, 0.736]	[-0.475, 0.750]	[-0.464, 0.736]	[-0.475, 0.750]
MTS assumption				
0	[-0.144, 0.636]	[-0.166, 0.650]	[-0.144, 0.636]	[-0.166, 0.650]
.01	[-0.275, 0.646]	[-0.300, 0.660]	[-0.168, 0.646]	[-0.190, 0.660]
.02	[-0.339, 0.656]	[-0.350, 0.670]	[-0.189, 0.656]	[-0.210, 0.670]
.05	[-0.350, 0.686]	[-0.361, 0.700]	[-0.235, 0.686]	[-0.254, 0.700]
.1	[-0.371, 0.736]	[-0.383, 0.750]	[-0.288, 0.736]	[-0.305, 0.750]
MIV and MTS assumptions				
0	[-0.117, 0.513]	[-0.158, 0.520]	[-0.117, 0.513]	[-0.158, 0.520]
.01	[-0.229, 0.530]	[-0.280, 0.537]	[-0.146, 0.521]	[-0.181, 0.529]
.02	[-0.333, 0.547]	[-0.335, 0.554]	[-0.170, 0.530]	[-0.200, 0.537]
.05	[-0.337, 0.595]	[-0.343, 0.600]	[-0.220, 0.556]	[-0.241, 0.563]
.1	[-0.357, 0.643]	[-0.363, 0.649]	[-0.274, 0.599]	[-0.287, 0.606]

[7] For the estimations, we used the tebounds command for stata (McCarthy, Millimet, & Roy, 2015).

Figure 3-2: Exogenous and No Selection (Worst-Case) Assumption



Without imposing any assumptions concerning the selection process, the bounds are of width one and necessarily include zero (WC-Bounds). Nevertheless, these bounds are informative to exclude possible values of the ATE (see Table 3-2). Allowing for measurement error, the bounds become wider; however, it is unexpected to see that the assumption of no false positive has *no* identification power over the range of assumed error values.

Self-selection into the truancy condition is a plausible explanation for the negative correlation between truant behavior and academic achievement. Adding this MTS assumption tightens the lower bounds and MIV further reduces the upper bounds. MIV in combination with MTS and probability of misreporting of one percent and no false positive, the estimates ATE bounds from -0.146 to 0.521. This corresponds to an interval of between a 14.6% decrease and a 52.1% increase in the probability of being a high-performing student when a student had been truant. Even in the absence of measurement error, the bounds include zero in all cases. Further, even small rates of misreporting significantly widen the bounds. Results for students who were frequently truant are comparable (Table 3-3). However, the lower bound is a little lower, which might be an indication for a larger negative effect of being truant frequently on student achievement.

Thus, the bounds still fail to identify the direction of the ATE. This leads to the conclusion that the treatment effect is very heterogeneous within the sample investigated and we therefore perform the analysis for students in different school tracks.

Figure 3-3: MTS and MTS-MIV Assumption (Arbitrary Errors)

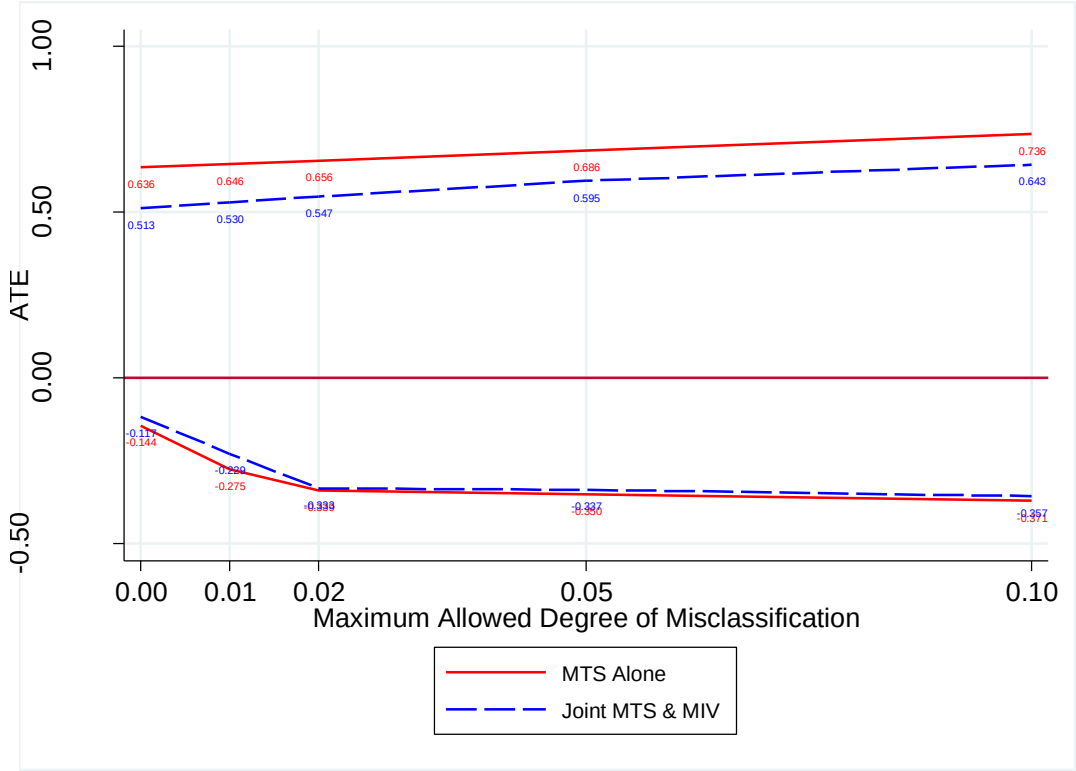


Figure 3-4: MTS and MTS-MIV Assumption (No False Positive Errors)

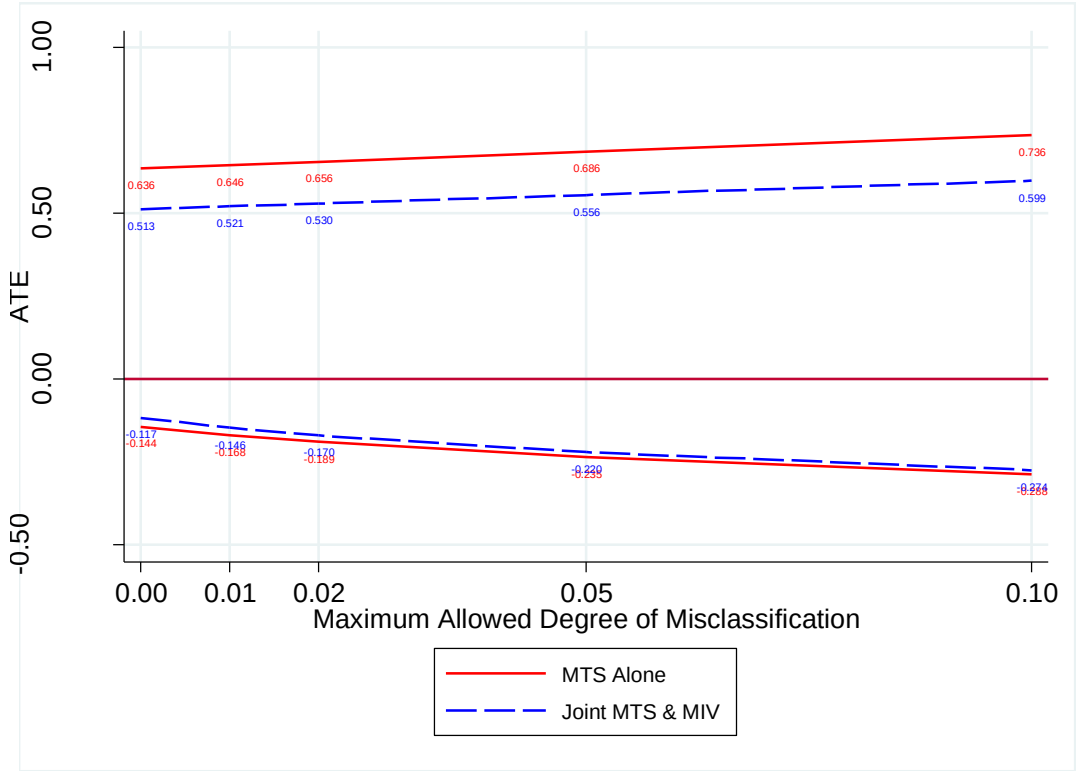


Table 3-3: Bounds on the ATE of Frequent Truancy on High Academic Outcome

	Arbitrary errors		No false positives	
Error rate	Bounds	95% CI	Bounds	95% CI
Exogenous selection model				
0	[-0.157, -0.157]	[-0.198, -0.102]	[-0.157, -0.157]	[-0.198, -0.102]
.01	[-0.326, 0.065]	[-0.336, 0.114]	[-0.203, 0.065]	[-0.236, 0.114]
.02	[-0.329, 0.213]	[-0.339, 0.329]	[-0.231, 0.200]	[-0.261, 0.243]
.05	[-0.340, 0.583]	[-0.350, 0.608]	[-0.278, 0.411]	[-0.301, 0.444]
.1	[-0.359, 0.691]	[-0.370, 0.709]	[-0.321, 0.570]	[-0.340, 0.597]
No monotonicity assumptions (worst-case selection)				
0	[-0.335, 0.665]	[-0.346, 0.679]	[-0.335, 0.665]	[-0.346, 0.679]
.01	[-0.345, 0.675]	[-0.356, 0.689]	[-0.345, 0.675]	[-0.356, 0.689]
.02	[-0.355, 0.685]	[-0.336, 0.699]	[-0.355, 0.685]	[-0.366, 0.699]
.05	[-0.385, 0.715]	[-0.396, 0.729]	[-0.385, 0.715]	[-0.396, 0.729]
.1	[-0.435, 0.765]	[-0.446, 0.779]	[-0.435, 0.765]	[-0.446, 0.779]
MTS assumption				
0	[-0.157, 0.665]	[-0.198, 0.679]	[-0.157, 0.665]	[-0.198, 0.679]
.01	[-0.326, 0.675]	[-0.336, 0.689]	[-0.203, 0.675]	[-0.236, 0.689]
.02	[-0.329, 0.685]	[-0.339, 0.699]	[-0.231, 0.685]	[-0.261, 0.699]
.05	[-0.340, 0.715]	[-0.350, 0.729]	[-0.278, 0.715]	[-0.301, 0.729]
.1	[-0.359, 0.765]	[-0.370, 0.779]	[-0.321, 0.765]	[-0.340, 0.779]
MIV and MTS assumptions				
0	[-0.138, 0.535]	[-0.157, 0.538]	[-0.138, 0.535]	[-0.157, 0.538]
.01	[-0.316, 0.552]	[-0.321, 0.555]	[-0.193, 0.543]	[-0.208, 0.547]
.02	[-0.319, 0.567]	[-0.323, 0.571]	[-0.224, 0.552]	[-0.237, 0.555]
.05	[-0.329, 0.595]	[-0.333, 0.601]	[-0.272, 0.578]	[-0.282, 0.581]
.1	[-0.348, 0.637]	[-0.352, 0.644]	[-0.313, 0.621]	[-0.322, 0.624]

Hypothesis 2. Since the group of truant students is quite heterogeneous and the effects of truant behavior have been found to be accordingly diverse, we investigated whether there are different treatment effects for two subgroups of the sample. In order to do so, bounds can be obtained by conditioning on covariates. To explore the heterogeneity in the ATE across different subgroups, we condition on school track. The results show the potential impact of misreporting, even at very low levels. Given that the sample was stratified according to different school tracks in Germany, we take into account that only the Gymnasium is prevalent in all 16 federal states. All other school types vary in terms of the numbers of study programs, denominations, and availability across federal states. In general, the Gymnasium schools represent the academic track qualifying for

tertiary education, whereas the non-Gymnasium schools offer different tracks both within and between schools and can be finished with diplomas qualifying for either vocational training, secondary level II education, or higher education. Thus, we distinguish Gymnasium schools from non-Gymnasium schools in this analyses, taking into account that differential effects of truancy on academic achievement have been found before that seem to be valid for the upper part of the performance distribution only (Arulampalam et al., 2012). Gymnasium schools across all federal states share a common history and similar curricula, while the non-Gymnasium school types are rather diverse with regard to their development and curricula.

Table 3-4: Bounds on the ATE of Truancy on High Academic Outcome: Gymnasium

	Arbitrary errors		No false positives	
Error rate	Bounds	95% CI	Bounds	95% CI
Exogenous selection model				
0	[-0.153, -0.153]	[-0.254, -0.102]	[-0.153, -0.153]	[-0.254, -0.102]
.01	[-0.273, -0.039]	[-0.398, 0.032]	[-0.235, -0.070]	[-0.326, -0.021]
.02	[-0.459, 0.137]	[-0.628, 0.214]	[-0.296, -0.007]	[-0.380, 0.038]
.05	[-0.662, 0.288]	[-0.691, 0.305]	[-0.419, 0.114]	[-0.488, 0.151]
.1	[-0.699, 0.356]	[-0.729, 0.368]	[-0.540, 0.224]	[-0.596, 0.253]
No monotonicity assumptions (worst-case selection)				
0	[-0.629, 0.371]	[-0.657, 0.382]	[-0.629, 0.371]	[-0.657, 0.382]
.01	[-0.639, 0.381]	[-0.667, 0.392]	[-0.639, 0.381]	[-0.667, 0.392]
.02	[-0.649, 0.391]	[-0.677, 0.402]	[-0.649, 0.391]	[-0.677, 0.402]
.05	[-0.679, 0.421]	[-0.707, 0.432]	[-0.679, 0.421]	[-0.707, 0.432]
.1	[-0.729, 0.471]	[-0.757, 0.482]	[-0.729, 0.471]	[-0.757, 0.482]
MTS assumption				
0	[-0.153, 0.371]	[-0.254, 0.382]	[-0.153, 0.371]	[-0.254, 0.382]
.01	[-0.273, 0.381]	[-0.398, 0.392]	[-0.235, 0.381]	[-0.326, 0.392]
.02	[-0.459, 0.391]	[-0.682, 0.402]	[-0.296, 0.391]	[-0.380, 0.402]
.05	[-0.662, 0.421]	[-0.691, 0.432]	[-0.419, 0.421]	[-0.488, 0.432]
.1	[-0.699, 0.471]	[-0.729, 0.482]	[-0.540, 0.471]	[-0.596, 0.482]
MIV and MTS assumptions				
0	[-0.153, 0.292]	[-0.199, 0.303]	[-0.153, 0.292]	[-0.199, 0.303]
.01	[-0.224, 0.310]	[-0.271, 0.321]	[-0.221, 0.301]	[-0.248, 0.312]
.02	[-0.348, 0.328]	[-0.398, 0.338]	[-0.275, 0.310]	[-0.292, 0.321]
.05	[-0.644, 0.357]	[-0.653, 0.370]	[-0.391, 0.337]	[-0.406, 0.348]
.1	[-0.678, 0.402]	[-0.682, 0.415]	[-0.511, 0.382]	[-0.523, 0.393]

Table 3-5: Bounds on the ATE of Truancy on High Academic Outcome: Non-Gymnasium

	Arbitrary errors		No false positives	
Error rate	Bounds	95% CI	Bounds	95% CI
Exogenous selection model				
0	[-0.061, -0.061]	[-0.080, -0.047]	[-0.061, -0.061]	[-0.080, -0.047]
.01	[-0.135, 0.042]	[-0.142, 0.051]	[-0.069, 0.042]	[-0.087, 0.051]
.02	[-0.137, 0.127]	[-0.143, 0.134]	[-0.076, 0.127]	[-0.093, 0.134]
.05	[-0.141, 0.322]	[-0.148, 0.330]	[-0.092, 0.322]	[-0.108, 0.330]
.1	[-0.150, 0.590]	[-0.157, 0.604]	[-0.112, 0.536]	[-0.126, 0.547]
No monotonicity assumptions (worst-case selection)				
0	[-0.201, 0.799]	[-0.213, 0.808]	[-0.201, 0.799]	[-0.213, 0.808]
.01	[-0.211, 0.809]	[-0.223, 0.818]	[-0.211, 0.809]	[-0.223, 0.818]
.02	[-0.221, 0.819]	[-0.233, 0.828]	[-0.221, 0.819]	[-0.233, 0.828]
.05	[-0.251, 0.849]	[-0.263, 0.858]	[-0.251, 0.849]	[-0.263, 0.858]
.1	[-0.301, 0.899]	[-0.313, 0.908]	[-0.301, 0.899]	[-0.313, 0.908]
MTS assumption				
0	[-0.061, 0.799]	[-0.080, 0.808]	[-0.061, 0.799]	[-0.080, 0.808]
.01	[-0.135, 0.809]	[-0.142, 0.818]	[-0.069, 0.809]	[-0.087, 0.818]
.02	[-0.137, 0.819]	[-0.143, 0.828]	[-0.076, 0.819]	[-0.093, 0.828]
.05	[-0.141, 0.849]	[-0.148, 0.858]	[-0.092, 0.849]	[-0.108, 0.858]
.1	[-0.150, 0.899]	[-0.157, 0.908]	[-0.112, 0.899]	[-0.126, 0.908]
MIV and MTS assumptions				
0	[-0.056, 0.652]	[-0.080, 0.664]	[-0.056, 0.652]	[-0.080, 0.664]
.01	[-0.135, 0.668]	[-0.142, 0.681]	[-0.065, 0.660]	[-0.087, 0.673]
.02	[-0.134, 0.685]	[-0.139, 0.698]	[-0.073, 0.668]	[-0.093, 0.681]
.05	[-0.138, 0.735]	[-0.144, 0.749]	[-0.090, 0.693]	[-0.108, 0.706]
.1	[-0.147, 0.790]	[-0.152, 0.806]	[-0.111, 0.731]	[-0.125, 0.745]

Note. CI = confidence interval; 95% confidence intervals obtained using 50 bootstrap repetitions.

To sum up, the results in this chapter ought to serve a note of caution to future evaluation of the effect of truant behavior on student performance. First, when ignoring the non-random selection or misclassification, we obtained a negative association between truant behavior and the probability of being a high performing student. Second, if only 2% of the students declare their truancy behavior incorrectly, the ATE cannot be assigned, even under the assumption of exogenous selection. Third, for the full sample, the bounds around the ATE that account for non-random selection always include zero, even if one assumes no misclassification. Fourth, the results depicted in Table 3-4 and Table 3-5 suggest that skipping mathematics classes seems to have a less negative effect on student

achievement on the PISA mathematics test in non-Gymnasium schools than in Gymnasiums. This may indicate that mathematics lessons in Gymnasium schools are used more efficiently by the students who attend them since their achievement is considerably higher than for students who skip classes on purpose. Another consideration is that being truant at Gymnasium schools is a more rational choice than in other school types, which may be because students are sorted into secondary school types in Germany as early as at age 10, after four years of schooling and according to prior achievement. Students attending Gymnasium schools, which in PISA 2012 were 36% of the cohort of 15-year-old students (Sälzer et al., 2012), are the highest achieving students within the tested cohort.

3.7 Discussion

In this study, we explored the causal impact of truancy on students' academic achievement while accounting for both non-random selection into the truancy group and measurement error. We addressed these identification problems with a partial identification framework. Instead of obtaining point estimates, this nonparametric method of analysis provides bounds around the ATE and is thus more credible. Moreover, the analysis drops the rather unrealistic assumption of a linear and homogeneous effect of intentionally skipping mathematics classes on students' performance on the PISA mathematics test. Generally, the bounds rely on observed sample averages and differences in averages between the treated and untreated groups (i.e., truant and non-truant students). We impose several weak assumptions concerning the nature of selection and the measurement error.

In this chapter, we included more than 8,000 students from the grade-based PISA 2012 national oversample in Germany. We tested three hypotheses. First, we investigated whether being truant in mathematics classes caused a low proficiency level on the PISA 2012 mathematics test. This hypothesis could be confirmed. However, regardless of measurement error, the bounds include zero and thus, being truant may also have a positive impact on the possibility of being a high performing student. This is often the case when bounding the ATE. Even though the imposed assumptions are relatively weak and plausible, there is still much ambiguity concerning the impact of truancy on students' academic achievement in mathematics. In this regard, our findings are meaningful since they indicate that being truant is not always a harmful decision for the student that results in low academic achievement. Instead, the relationship between truancy and academic achievement is rather complex – not linear and certainly not deterministic. Second, we tested whether there are different treatment effects for subgroups of students, namely, students in Gymnasium schools and other school types. Our results confirm this hypothesis and suggest that the decision to skip mathematics classes in Gymnasium schools and other school types might be taken due to different reasons and especially with differential effects. One plausible explanation for these differences

between Gymnasium and non-Gymnasium schools is that students in Gymnasium schools select the lessons to skip more carefully and with regard to which contents they anticipate missing. Missing a lesson that repeats content already mastered causes less harm than missing an introductory lesson presenting a new field or topic. Third, we assumed that under the consideration of measurement error, the bounds would widen and include zero, allowing for an increasing probability of truant students being high achievers. This hypothesis could be confirmed as well and is probably the most relevant of our findings since it suggests that truancy can also be associated with high academic achievement even under exogenous selection. Accounting for misreporting, the estimated effects of truancy on academic performance are largely inconclusive. This confirms the sensitivity to misclassification of analyses involving self-reported data on truancy.

Our study is supported by a number of strengths, such as a large national sample and mandatory participation of the students. It is, however, at the same time limited by several issues that should be addressed in further studies. First, our dataset is limited to only one country and it cannot be determined to what extent our results are generalizable to other countries and cultures. Especially with regard to the stratified school system, results may be different in other countries. Second, we focused our analyses on only one out of three domains tested in PISA. Taking into account scientific and reading literacy, the findings could either be scaffolded or qualified. Despite these limitations, the findings of our study contribute to the methodological development of truancy as a research topic in education. The identified bounds are still quite large and include zero, which allow for a range of conclusions. One aspect might be that students were truant for different reasons, indicating that there is a heterogeneous and probably non-linear effect of truancy behavior on students' achievement in mathematics. Having said this, we have to assume that as staying away from school without a valid reason tends to be more common in the final years of secondary education (Wagner, Spiel, & Tranker, 2003), the effect of truancy on behavior might be even more heterogeneous for later stages of school.

In sum, this study showed a possibility of conducting causal analyses with data from PISA, which per se are cross-sectional and therefore rather limited with regard to causal interpretations and conclusions. Besides the association of truancy and academic achievement in mathematics and differential treatment effects for subgroups, we showed that students' self-reported data on truant behavior need to be analyzed with great care. Their sensitivity to measurement error is high, and in order to avoid false conclusions models for analyzing such data need to be carefully chosen and based on strong theoretical assumptions.

Chapter 4

The Impact of Work-Based Education on Personality Skills

This chapter has been published in another version with Thomas Bolli as co-author under the Title "The Impact of Apprenticeship Training on Personality Traits: An Instrumental Variable Approach", (2014), KOF Working Paper No. 350

4.1 Introduction

"The most promising adolescent programs integrate aspects of work into traditional education. […] In earlier times, adolescents took apprenticeships and jobs where they were supervised and mentored by adults. Mentoring involved teaching valuable character skills – showing up for work, cooperating with others, and persevering on tasks" (Heckman & Kautz, 2013, p. 35).

The relationship between personality skills[1] and success in life has been widely demonstrated, as such skills have been found to be strong predictors of academic performance and life outcomes (see, e.g., Almlund, Duckworth, Heckman, & Kautz, 2011; Borghans, Duckworth, Heckman, & Weel, 2008; Boyce, Wood, & Powdthavee, 2013; Brunello & Schlotter, 2011; Falch, Nyhus, & Strøm, 2014; Fletcher, 2013; Heckman & Kautz, 2012; Lindqvist & Vestman, 2011). Adolescence is shown to be a time during which personality skills are still fluid compared to adulthood. Though personality skills may change as a result of educational experience, there is surprisingly little evidence on the effect of education on personality skills (Büttner, Thiel, & Thomsen, 2011; Dahmann & Anger, 2014; Hanushek, Welch, Machin, & Woessmann, 2011; Heckman, Stixrud, & Urzua, 2006; Meghir, Palme, & Simeonova, 2013), and none of the studies focuses on work-based education. Previous evidence has shown that work experience has effects on a wide variety of personality skills (for an overview see Roberts, Caspi, & Moffitt, 2003). Hence, breaking down the rigid separation between school and work, work-based education may affect personality skills differently than full-time school-based

[1] Other terms used for similar concepts in the literature include soft skills, character skills, psychological skills, personality traits, character, personality factors or socio-emotional skills (Borghans, Duckworth, Heckman, & Weel, 2008; Heckman & Kautz, 2013; Heckman, Pinto, & Savelyev, 2013).

education. Therefore, this chapter aims to provide first evidence on the causal effect of work-based upper secondary education on personality skills.[2]

We exploit a dataset that follows the participants of the 2000 Swiss Program for International Student Assessment (PISA) examination at grade 9 up to the year 2010. Students in the treatment group of work-based education receive on-the-job training at the training firm during three to four days a week in combination with one to two days of classroom learning in a vocational school per week (half of all students in Switzerland enroll in these kinds of VET programs with apprenticeships (SKBF, 2011). We compare these students with students in the control group of fully school-based education.

Personality skills of students might affect the selection of the educational track. Hence, to address these concerns regarding endogeneity due to selection and unobserved heterogeneity, we apply three different strategies. First, we make use of the panel structure of our data set to analyze changes over time. Second, we apply an Instrumental-Variable Approach that exploits regional differences in the relevance of general secondary education across Switzerland to account for potential selection in personality skill growth. The regional differences in the shares of general secondary education, which varies between 10 and 32%, are based on historical decisions made by the government and remained relatively stable over the last 20 to 30 years and we therefore argue that these historical differences produce exogenous variation. While this addresses reverse causality, potential endogeneity due to unobserved heterogeneity across regions correlated to both personality skills and general secondary education share, such as cultural variation, remains a problem. We address this issue in four ways. First, we control for the lagged dependent variable accounts for unobserved heterogeneity in the level of the dependent variable. Secondly, we compare regions within relatively homogenous areas which limits potential endogeneity problems. Thirdly, we exploit the small variation in the shares of general secondary education across time. Fourthly, we apply a second instrument based on the relevance of work-based education in the students' country of origin, allowing testing instrument validity formally.

Findings show that work-based education decreases emotion-centered coping, i.e. increases emotional stability. It potentially increases contact-centered coping, indicating an improvement in interpersonal relationship and potentially reduces intrinsic work motivation. No effect is found for task-centered coping. The effect sizes are economically significant. The results suggest that the impact on emotion-centered coping represents a permanent shift.

The remainder of this chapter is organized as follows. Section 4.2 reviews the existing evidence on the effects of education on personality skills and discusses how work-based education may affect

[2] This paper forms interpretable aggregates of facets of personality skills through factor analysis. This method summarizes the covariablity among different personality measures using low-dimensional latent variables. The latent factors variables are the factors.

personality skills. Section 4.3 reveals the data, and section 4.4 presents the estimation strategy. Section 4.5 reports our results of the impact of work-based education on personality skills, and section 4.6 presents this chapter's conclusions.

4.2 Literature and Theoretical Framework

Recent literature finds that non-cognitive skills, such as personality skills have a significant impact on a wide range of outcomes (Almlund et al., 2011; Borghans et al., 2008; Brunello & Schlotter, 2011; Fletcher, 2013; Gensowski, 2014; Heckman & Kautz, 2012; Lindqvist & Vestman, 2011). For example, recent evidence shows that 30 to 40% of the explained variance in achievement test scores across student is due to personality skills and not IQ (Heckman, Pinto, & Savelyev, 2013).

Substantial evidence exists that these personality skills are not permanently entrenched at birth (Boyce et al., 2013; Hanushek et al., 2011; Heckman et al., 2013). While the literature claims that genetic factors are responsible for the stability of personality skills, environmental factors are responsible for changes in personality skills (Blonigen, Hicks, Krueger, Patrick, & Iacono, 2006; Borghans et al., 2008; Lykken, Bouchard, McGue, & Tellegen, 1993). Late adolescence and early adulthood seem to be critical and sensitive periods, i.e., a time when personality skills are still very fluid compared to adulthood (Cobb-Clark & Schurer, 2012; Dahl, 2004; Roberts & Mroczek, 2008; Roberts, Walton, & Viechtbauer, 2006). As the predominant environment during adolescence and early adulthood is the educational environment, it may influence personality skills. Therefore, it is important to understand how personality skills can change, in particular, to what extent education influences the development of personality skills.

4.2.1 The Impact of Education on Personality Skills

Only a few empirical studies have examined the causal relationship between education and personality skills. Heckman, Stixrud and Urzua (2006) formulate a theoretical model for the effect of school years on cognitive skills and personality skills. Importantly, the model reveals the possibility of reverse causality, i.e. selection of students into education according to cognitive and non-cognitive skills. They find evidence that the number of years of schooling affects personality skills. Concretely, an additional year of either high school or college increases self-esteem, while the locus of control is primarily affected by high school, but not college attendance.[3] Büttner et al. (2011), in contrast, using a natural experiment in Germany induced by an educational policy reform, where the last year of higher secondary schooling was abolished, find no effect of learning intensity on personality skills. However, Dahmann and Anger (2014), analyzing the same educational reform for the whole country,

[3] Self-esteem refers to an individual's subjective sense of his own worth (De Wals & Meszaros, 2011). Locus of control refers to an individual's belief about whether the determinants of one's life are determined internally or externally (Rotter, 1966).

show a decreasing impact on emotional stability with substantial heterogeneity in the effects. Meghir et al. (2013) analyze an increase of schooling years in Sweden, suggesting that non-cognitive skills are improved, though only for students with high socio-economic background. In addition, Lüdtke et al. (2011) present evidence for Germany that a life experience, i.e. failing an important exam is associated with a change in personality skills, in this case an increase in neuroticism. Jackson (2011) analyzes the impact of educational experience on personality skills and suggests that educational contexts are important for the development of personality skills. In this study experiences outside the classroom were also related to changes in personality skills, e.g. spending time working for pay was associated with increases in extraversion, but not with changes in any other personality skills. However, this study does not identify causal effects.

4.2.2 The Impact of Interventions before or during School on Personality Skills

Some studies have analyzed the impact of different interventions[4] before or during school on personality skills. Studies based on the randomized Perry Preschool and STAR projects find that home visits, better peers and smaller classes[5] positively impact personality skills (Dee & West, 2011; Heckman et al., 2013). Heckman et al. (2013) analyze the channels through which these persistent changes in personality skills may occur: The reduction in externalizing behavior, i.e. aggressive, antisocial and rule-breaking behaviors, is especially strong. Thus, factors other than cognitive skills, such as personality skills, are potentially influenced by experiences within the educational system (J. J. Jackson, 2011). While these two projects are not designed to affect personality skills, there are programs that do. For example, a randomized 3-year socio-emotional learning program, the Promoting Alternative Thinking Strategies (PATHS) curriculum, is associated with an increase in authority acceptance, concentration and social competence (Bierman et al., 2010). Other interventions are more short-term and designed to isolate a particular effect. In a randomized experiment in Switzerland (Behncke, 2012), the treatment group received positive affirmation intervention before taking a math test. The test scores for the treatment group were significantly raised, which the author attributes to a change in non-cognitive abilities, such as an increase in student motivation and self-confidence and a decrease in test anxiety. Accordingly, the learning environment, e.g., teacher practices, seems to be crucial for the development of personality skills.[6]

---

[4] For an overview, see Almlund et al. (2011), Brunello and Schlotter (2011) or Heckman and Kautz (2013).

[5] For Sweden, Fredriksson et al. (2013) apply a regression discontinuity approach to show that a unit reduction in class size improves non-cognitive outcomes by 0.026 of a standard deviation.

[6] Others relate systemic features of school systems to personality traits (Falck & Woessmann, 2013). Luedemann (2011), for example, finds a small but significantly positive impact on students' personality traits results from the monitoring of teacher lessons by the principal or external inspectors according to assessments used to compare the school to district or national performance standards.

4.2.3 The Impact of Work on Personality Skills

The question whether work-based education causes a change in personality skills is addressed in this chapter. Good-quality workplace learning provides students with valuable labor market experience before graduation by enabling apprentices to develop technical skills and gain real world experiences (OECD, 2013a). Based on the neo-socioanalytic model (Roberts & Wood, 2006) change in personality skills is a result from the transaction between people and their environments, e.g. a person's participation in social norms and the social interactions. In general, these new structures prompt people to become more agreeable and conscientious and less neurotic (Roberts & Wood, 2006). Given the high proportion of time many individuals spend each day at the workplace, the workplace may be one of the domains within which personality changes. Several empirical studies have examined the relation between work experiences and personality, showing that work has effects on a wide variety of skills (for an overview see Roberts, Caspi, et al., 2003), consistent with the neo-socioanalytic model. For example for men using a wider variety of skills on the job is related to increases in emotional stability (Brousseau & Bruce, 1981). Women's participation in the paid labor force is associated with an increase in conscientiousness (Roberts, 1997). Moreover, for both sexes, occupational attainment and work satisfaction are associated with an increase in emotional stability and conscientiousness (Roberts, Caspi, et al., 2003; Roberts & Chapman, 2000). Lüdtke et al. (2011) show that positive experiences of beginning regular work were associated with increases in emotional stability. Further, applying a Diffs-in-Diffs Approach, they find that entering work or vocational education after general secondary education at age 19 (in this research we analyze the transition from compulsory education to secondary education at age 15) compared to starting college is associated with an increase in conscientiousness and a decrease in agreeableness (Lüdtke et al., 2011). However, a common trend is assumed.

4.2.4 Channels through which Work-Based Education Might Affect Personality Skills

Personality skills and coping strategies are highly associated (e.g. Amirkhan, Risinger, & Swickert, 1995; Leandro & Castillo, 2010). According to Endler and Parker (1990), it is possible to assign coping styles into tree main categories: 1) problem-centered (focused) coping with attempts to regulate the situation, 2) emotion-centered coping with attempts to regulate the emotion and 3) avoidance - centered coping which aims at avoiding the stressful situation (see also "Measures of Personality Skills")

As apprentices are exposed to different environmental factors than students in high school we expect to have different treatment effects for work-based and school-based secondary education on non-cognitive skills. We identify four channels, i.e. mechanisms that have the potential to be related to a different causal effect to change in non-cognitive skills for work-based education compared to full-time school-based secondary education. However, we cannot empirically distinguish between these channels.

An important step towards the theoretical framing of how non-cognitive skills are produced is formulated with the multistage production model of Cunha, Heckman and Schennach (2010) of the evolution of children's cognitive and non-cognitive skills as determined by parental investment at different stages of life. Further, recent evidence shows, that there is nearly no correlation between the effect of teachers on students' test scores and on students' non-cognitive outcomes (C. K. Jackson, 2012). Therefore teachers and schools face a trade-off between investments in cognitive skills and personality skills, which we label the *Trade-off channel*. Full-time schools measure student achievement by cognitive tests, as personality skills are difficult to measure. Moreover, general secondary education teachers are not allowed to rate or assess students' personality skills, while firms training apprentices have to follow a prescribed curriculum, including teaching of personality skills. Accordingly, full-time schools do not focus on the development of personality skills. Because apprentices, on the other hand, come into contact with clients, instructors are more inclined and have to invest resources in the development of personality skills.

Individuals are assumed to change personality skills as they learn social norms, most often on the basis of feedback from peers (Roberts, Caspi, et al., 2003; Turner, 2013). This *Feedback Channel* may affect apprentices different compared to students because apprentices are supervised and mentored by professionals in the training firm and most have contact with clients. Therefore, education in the workplace may involve the teaching of different personality skills (Heckman & Kautz, 2013; Lerman, 2013). For example, apprentices must report for work on time (punctuality), and they do not have the option of 'skipping' the first lesson. They also have to cooperate with others more intensely (team work (OECD, 2013a)) and not only with students of the same age but also with adults and professionals who are older and more experienced (OECD, 2013a). Therefore, apprentices face a much older and more experienced reference group. By serving as role models, these older group members may affect the personality skills of individuals. Furthermore, group members have the potential to sanction non-conforming social behaviors. Relatedly, as apprentices earn wages, the training firms also have the possibility to sanction non-conforming behaviors. Persevering on tasks (work discipline) and reliability represent examples of skills that apprentices must acquire to be successful in their workplace environment. Following a more disciplined schedule with structured expectations increases conscientiousness, which is shown for students entering vocational training or work after general secondary education compared to students entering college (Lüdtke et al., 2011).

Two other domains of work-related socialization are power, e.g. having the ability to get things done and the feeling that one is gaining financial security (Roberts, Caspi, et al., 2003). Through experiences of fulfilling task and obligations individuals develop responsibility (Roberts, Wood, & Smith, 2005). We use the term *Responsibility Channel* for this socialization process which may affect apprentices different compared to students because apprentices face more responsibility. First, they interact directly with clients. Second, they are responsible for valuable equipment, and third they serve as role models for the younger apprentices. Hence, during their education, they assume a supervisory and parental role for younger apprentices. Taking on a new role or obligations is described as the first step in the youth's cycle of developing responsibility (Salusky et al., 2014). Fourth, apprentices earn for the first time some money and therefore "feel that they have "made" the transition to adulthood or maturity" (Roberts, Caspi, et al., 2003, p. 584) which is associated with skills such as emotional stability, agreeableness, and conscientiousness (Roberts, Robins, Caspi, & Trzesniewski, 2003). Accordingly, acting responsibly and feeling responsible is important and may lead to increased self-confidence and reliable behavior. Hence, we hypothesize that these three channels increase task-centered and contact-centered coping while decreasing emotion-centered coping, i.e. increases emotional stability:

H1: Work-based education increases task-centered coping

H2: Work-based education increases contact-centered coping

H3: Work-based education decreases emotion-centered coping

And it can be assumed that educational climate, such as teacher behavior or peers have an important impact on students' motivation in the self-determination theory (Reeve, 2002). Students starting work-based education move earlier from the freedom of adolescence to the responsibilities of adulthood (process of social investment) than students moving to full-time school-based secondary education. This so-called social investment process (Lodi-Smith & Roberts, 2007)[7] in the workplace is associated with personality skill change (Hudson & Brent, 2012). We therefore label this channel the *Freedom Channel* that arises because students in school-based education profit from a higher degree of freedom and more leisure time. This includes both freedom regarding the way students learn and the amount of leisure time they have each week. Furthermore, school-based education offers more than twice as much vacation time for students compared to apprentices, which typically have 5 to 6 weeks of vacation per year. The lower degree of freedom might foster extrinsic over intrinsic motivation as suggested by the self-determination theory (Deci & Ryan, 1985; Ryan & Deci, 2000). Komarraju et al. (2009) support this hypothesis by showing that openness is related to intrinsic

[7] The social investment process is shown for the example of young adults who entered for the first time a long-term romantic relationship and experienced a simultaneous increase in emotional stability (Lehnart, Neyer, & Eccles, 2010).

motivation of college students. Komarraju et al. (2009) further show that extrinsic motivation is positively related to extraversion.

H4: Work-based education decreases intrinsic motivation

4.3 Data

In this research, we use the Transition to Education and Employment survey (TREE).[8] The TREE is a longitudinal follow-up panel study to the PISA 2000 that was conducted in Switzerland. The TREE survey is administered each year between 2001 and 2007 and in 2010. The sample is representative of both the country as a whole and its three main language regions (German, French, and Italian). This unique database combines the variables in the standard PISA survey, such as parental background, PISA test scores and living conditions with information on personality skills and employment/education status. The Appendix displays complementary Tables, i.e. variable definitions. Using a balanced panel we still face the problem of attrition in the TREE data set, however, we apply a robustness check including weights.

4.3.1 Swiss Education System

After completing the Swiss compulsory school (9th grade), adolescents can choose among several possibilities. Almost half of the students (43%) enter apprenticeship training programs. Apprenticeships are a core element of the vocational education and training (VET) in Switzerland and typically last three or four years. They combine on-the-job training at the training (host) firm with one to two days of classroom learning in a vocational school per week.[9] Roughly one third of the companies in Switzerland engage in apprenticeship training. Of the students finishing lower secondary education, 40% begin a school-based secondary education, 16% follow an alternative education path, one percent enter the workforce and one percent do nothing. However, these percentages differ substantially among the various Swiss cantons (member states), and these differences have been highly persistent for the last 20 years (SKBF, 2011).

---

[8] The Swiss youth panel study TREE (Transitions from Education to Employment; www.tree-ch.ch) has been ongoing since 2000 and is funded by the Swiss National Science Foundation, the University of Basel, the Swiss Federal Office of Statistics, the Federal Office of Professional Education and Technology, and the cantons of Berne, Geneva and Ticino. Distribution: Data service, FORS, Lausanne: http://www2.unil.ch/daris/spip.php?rubrique141&lang=en

[9] The legal basis for each VET program in Switzerland can be found in VET ordinances issued by the Federal Office for Professional Education and Training. Training plans form the basis for the vocational teaching concept used in the apprenticeship. They define not only technical but also social and personality skills as student must acquire. At vocational schools apprentices are educated by vocational teachers and at training companies they are trained by vocational trainers, both must meet certain standards.

4.3.2 Measures of Personality Skills

One of the most popular psychological concepts employed to measure personality are the Big Five personality dimensions (Costa & MacCrae, 1992; McCrae & Costa, 1987). Unfortunately, the Big Five inventory is not included in the TREE data. We therefore use a number of self-reported measures[10] of personality skills, as summarized in Table 4-5 in the Appendix. There are many ways to summarize the available psychological measures in TREE. The approach used in this chapter is exploratory factor analysis. Factor analysis is the standard approach for defining constructs in personality psychology (Borghans et al., 2008) and is for example applied by Heckman et al. (2013) when summarizing Perry Preschool psychological measures. Using the within-cluster correlations of the measures, we isolate a latent factor for each of the personality skills; thereby create a low dimensional and interpretable aggregate of the employed psychological measures. The rotated factor loadings of the principal component factor analysis displayed in Table 4-6 in the Appendix confirm that the employed proxies represent independent dimensions. The table further shows that the measures aggregated in each dimension are internally consistent, i.e. load into the same factor with sufficient strength. Table 4-8 in the Appendix shows summary statistics for the factor scores divided by the educational path.

4.4 Estimation Strategy

To assess the impact of work-based education on personality skills compared to full-time school-based education, we start by estimating an OLS equation of the following form:

$$P_{it} = \alpha + \alpha_t + \beta_1 A_i + \beta_2 B_{it} + \varepsilon_{it}, \qquad [1]$$

where A is a dummy variable indicating work-based education (apprenticeship) and P represents the personality skills of student i at time t. B is a set of control variables, e.g. gender, age, PISA reading scores, socioeconomic background of mother and father, family structure, urban living. Table 4-7 and Table 4-8 contain a description of all control variables included in the estimations as well as descriptive statistics. ε is a random error with mean 0, clustered at the individual level.

We construct our treatment group from individuals who start work-based education in 2001 and remain apprentices until 2003, and we compare the outcomes to a control group[11] that participates in full-time schooling. This control group starts school-based secondary education in 2001 and remains

[10] Measurement of latent factors with self-reports may be false when false responses are made because of impression management or due to self-deception (Paulhus, 1984; Paulhus & Reid, 1991).

[11] Using observations of individuals in the respective track in 2001, 2002, 2003 and 2004 yields qualitatively similar results.

enrolled in the program until 2003.[12] We restrict the sample to observations of individuals who have responded in all employed variables between 2001 and 2003, i.e. create a balanced sample for this period.

There might be selection into education track according to personality skills (Hanushek et al., 2011; Heckman et al., 2006). Hence, OLS estimates may suffer from an endogeneity bias due to selection. For example, low intrinsic work motivation might lead the student to choose a school-based education rather than work-based secondary education. However, both cognitive and personality skills contribute to education choices and performance (Cunha, Heckman, & Lochner, 2006; Heckman, Humphries, Veramendi, & Urzua, 2014; Heckman et al., 2006) and we therefore expect a negative selection into apprenticeships (PISA test scores in Table 4-8 support this hypothesis), hence OLS estimates will underestimate (overestimate for negative personality skills) the effect.

Because we can never observe the same student under different secondary education treatments, the credibility of an empirical analysis depends on the plausibility of the identification strategy. Our first approach to tackle selection is to exploit the longitudinal structure of the data base. Including the lagged dependent variable P_{it-1} on the right hand side accounts for selection in terms of the personality skill level (LDV):

$$P_{it} = \alpha + \alpha_t + \beta_1 A_i + \beta_2 B_{it} + \beta_3 P_{it-1} + \varepsilon_{it.} \qquad [2]$$

While this approach represents a first step towards a causal interpretation, we still assume a common trend, i.e. that personality skill trends would be the same for both groups of students in the absence of the treatment. Since we observe personality skills in 2001 for the first time, we cannot test this assumption. Furthermore, the first observation of personality skills takes place a few months after the secondary education has started. Hence, this approach disregards any effect that arises during the first months after the start of the education.

Due to these issues with the estimation of [2], we additionally report the results of an instrumental variable (IV) approach (Angrist, Imbens, & Rubin, 1996) in which case A_i in formula (1) represents predicted values based on the following first-stage model:[13]

[12] Bolli and Hof (2014) show that the results hold for both students in general and vocational school-based upper secondary education.

[13] Due to the binary character of the endogenous variable, we estimate the model with the treatreg command of Stata12 and use the ivreg2 command of Stata 12 as a robustness check and to provide statistical tests for weak instruments and overidentification.

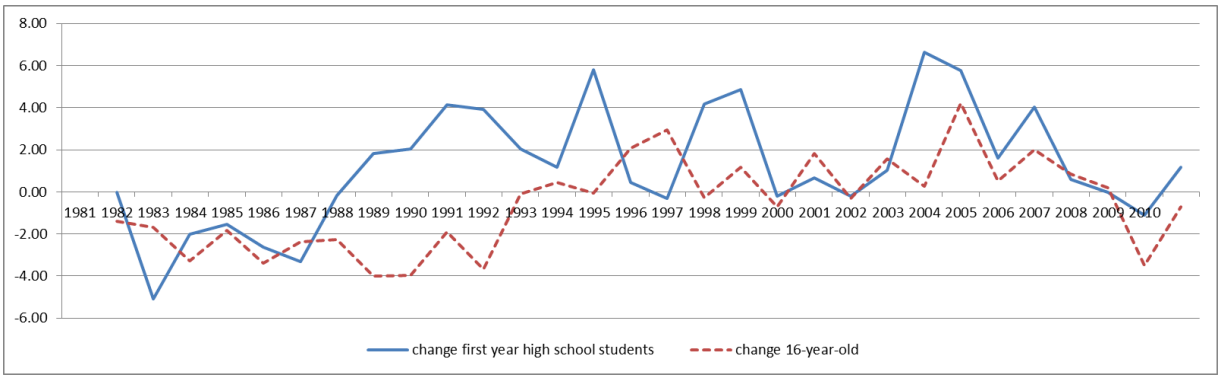
$$A_i = \delta + \delta_t + \delta_1 B_{it} + \theta_1 z_{1j} + u_{it}, \quad [3]$$

where z denotes our instrumental variables and u is a random error term. Subscript j refers to the level of the instrument.

We use two different instruments as described in Table 4-7 and summarized in Table 4-9 in the Appendix. Thereby, we provide evidence that the instrumental variable approach holds for two different types of arguably exogenous variation. Furthermore, using both instruments simultaneously allows us to conduct a Sargan test, thereby testing the validity of our instrument formally.

The first instrument refers to the share of general secondary education among cantons in Switzerland in 1998. As Table 4-9 shows, this share varies substantially across Switzerland. The cantonal differences in the share of general secondary education are based on historically set shares by the government and reflect the historical differences in the importance of work-based education in the region. The reasons for the historical differences in the cantonal share remain unexplained, but Table 4-9 shows that the differences remained stable over the last 20 to 30 years. Figure 4-1 displays that the national share of general education in Switzerland is independent of the cohort size of the 16 year old. We use this historical pattern as a natural experiment to estimate the causal effect of work-based education on students' personality skills in cross-cantonal student-level analyses.

Figure 4-1: Cohort Sizes of General Upper Secondary Education and 16-Year Old in Switzerland

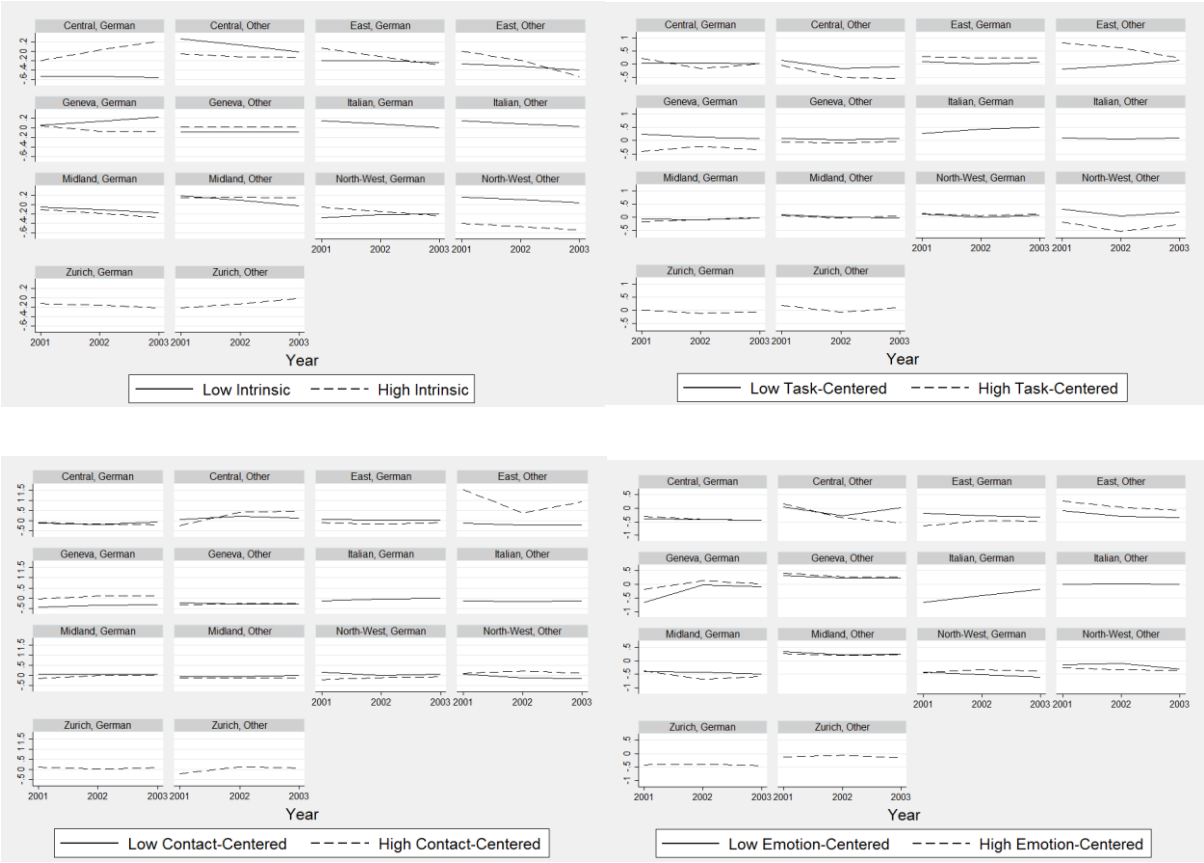


The internal validity of our IV approach relies on the assumption that the cantonal shares of general upper secondary education and personality skills are independent. This assumption, however, may be violated, e.g., because of culture. Beside of controlling for an extensive vector of control variables, we address this issue in four main ways. First, we include the lagged dependent variable on the right hand side, thereby removing any unobserved heterogeneity in the level of personality skills across cantons.

Secondly, we add dummy variables for five areas in combination with the student's language at home into our estimation, implying that we only exploit within-area variation for the identification of the effect and thereby homogenize the variation in terms of culture. In order to illustrate the relationship

between the within-area-culture development and personality skills, Figure 4-2 shows the development of personality skills in cantons with low and high share of school-based education for each of the area-culture groups. This descriptive evidence suggests that the relationship between both the level and the development of personality skills is unrelated to the within-area-culture distribution of the share of school-based education, thereby providing suggestive evidence that the instrument might be valid.

Figure 4-2: Within-Area-Culture Average of the Development of Personality Skills in Cantons with High and Low Share of School-Based Education



Thirdly, in order to address remaining concerns regarding the instrument validity, we report estimates that control for the share of school-based education in 1980, thereby instrumenting the probability of selecting an apprenticeship by the growth of school-based education between 1980 and 1998.

Fourthly, we report estimates that drop all control variables except for the lagged dependent variable and the area and culture dummies. The stability of estimation results suggest that the vector of control variables is orthogonal to the instrument, providing further credibility to the applied instrument.

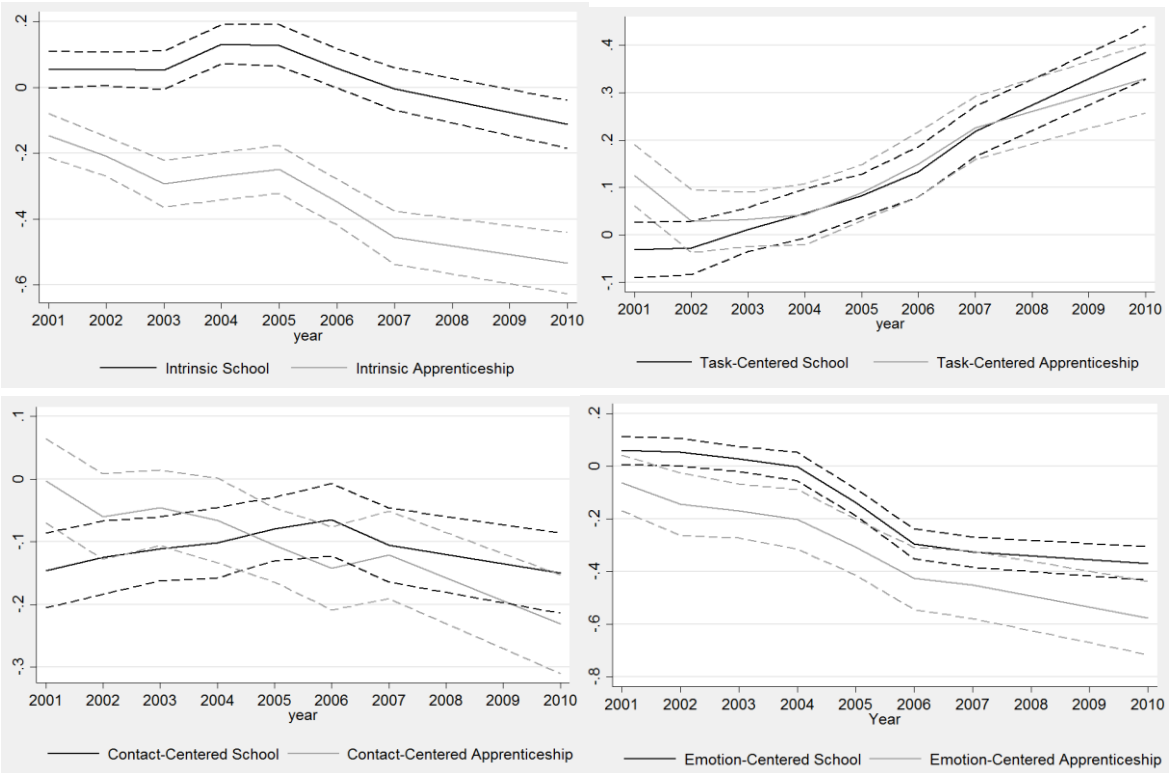
Another potential confounder of our estimation arises because personality skills affect occupational choice (Heckman et al., 2006). We address this issue in two ways. First, we include dummy variables for the first job (4-digit ISCO codes). Second, we report estimates that restrict the sample to apprentices in a commercial apprenticeship (KV).

The second instrument exploits the fact that foreigners that are less familiar with the Swiss education system are less likely to choose a work-based education. We approximate familiarity with the principle of work-based education by the share of work-based education in the country of origin, based on the OECD indicator "Students enrolled by type of institution" available at http://stats.oecd.org/.

4.5 Results

This section starts by presenting the results that address selection by exploiting the longitudinal dimension of the data set to analyze the impact of work-based education compared to school-based schooling on personality skills. Concretely, Figure 4-3 represents the approach graphically and Table 4-1 displays the corresponding estimation results.

Figure 4-3: Graphical Representation of Personality Skills Change by Education Track



Note: School refers to students starting school-based secondary education in 2001; Apprenticeship refers to students starting work-based education in 2001. Dashed lines show the confidence intervals.

Figure 4-3 and Table 4-1[14] suggest that work-based education increases intrinsic work motivation but decreases emotional-centered coping. The coefficient estimates for task-centered coping are negative, but insignificant; those fore contact-centered coping are positive, but also insignificant.

[14] Using the differences over time in the dependent variables yields qualitatively similar results.

Table 4-1: OLS Estimates Including Lagged Dependent Variable: Work-based Education vs. School-based Education

		Intrinsic Work Motivation	Task-Centered Coping	Contact-Centered Coping	Emotional Coping
OLS	Apprentice	-0.211***	-0.039	0.085*	-0.170***
		(0.046)	(0.046)	(0.046)	(0.043)
LDV	Apprentice	-0.109***	-0.049	0.001	-0.172***
		(0.034)	(0.039)	(0.038)	(0.036)

Notes: N=4442. The table displays OLS coefficients and standard errors clustered at the individual level in parentheses based on the TREE dataset for 2002 and 2003 pooled and separately. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The sample consists of students continuously enrolled in either a school-based or work-based educational track between 2001 and 2003. All estimates include year dummies and the control variables described in Table 4-7. The LDV estimates further include lagged dependent variable.

However, while these estimates control for selection in the level of the dependent variable, we are unable to show that control and treatment group have common trends. Hence, our preferred estimation strategy is based on an instrumental variable approach that exploits the regional variation in the share of general secondary education which was set by government historically. Table 4-2 displays the baseline results in addition to a number of robustness checks. Apprentice refers to the second-stage coefficients of the endogenized variable indicating work-based education and compares work-based educated students to students in school-based secondary education. All estimates (except NO CONTROLS) include covariates that may affect the choice for work-based education, i.e., socio-demographic and socio-economic background, age, gender, language, cantonal religion and particularly important student's competence level measured at the end of compulsory education in the standardized PISA competence measurement.[15] We report Kleibergen-Paap rk Wald F weak instrument statistics, which substantially exceed the critical value of about 16 in all regressions.

These baseline AREA+ CULTURE estimates, include dummy variables for 7 areas in Switzerland, thereby homogenizing the compared cantons and dummy variables for the native language to account for potential difference in the cultural heritage. The LDV estimator homogenizes cantons by including the level of the lagged dependent variable. The baseline estimation suggests a negative influence of work-based education on emotional-centered coping, while there is potentially an increasing effect on contact-centered coping. There is no effect on intrinsic work motivation and task-centered coping. Including LDV in the estimation the coefficient of emotional-centered coping declines from -0.6 to −0.4. Subsequently, the coefficients of the baseline IV estimates remain stable after weighting and, even more important; remains stable when we drop all control variables. Further, the effect also remains stable when we control for the cantonal share of general secondary education in the year 1980 (DELTASHARE). Thus, the applied IV estimates are robust.

[15] We use the PISA Reading Score as opposed to the PISA Math Score due to fewer observations. However, the qualitative results are the same.

However, we show several robustness checks. The KV and NOGA estimates test whether the effect is due to occupational choice by restricting the sample of work-based education to commercial apprenticeships and adding dummy variables for the first job, respectively. COUNTRY estimates exploit the fact that foreigners that are less familiar with the Swiss education system are less likely to choose a work-based education. Note that country of origin may affect skills and, hence, educational choice. Therefore it is particularly important to add the control vector for estimates using the share of work-based education in the country of origin as instruments as this share is negatively related to the choice of work-based education otherwise.[16] Finally, CANTON+COUNTRY estimates include both the cantonal share of general school students and the share of work-based education in the country of origin as instruments, allowing testing the validity of the instruments formally. The Hansen p-value of the overidentification test supports the exogeneity assumption of the instruments. Our robustness checks confirm the decreasing effect of work-based education on emotional-centered coping found in the baseline estimations, which is stable across methodologies and samples.

To aid in interpreting the magnitude of the estimated effects, remember that the dependent variables take values between -6 and 4, have a mean of 0 and a standard deviation of 1. Hence, a coefficient of 1 suggests that work-based education (a change from 0 to 1) results in a change in the order of one standard deviation. Therefore, the estimated effects are economically significant.

[16] The results shown in Table 4-2 exclude control variables for Language, Swiss and Time Swiss. Hence, identification rests largely on the difference in the share of work-based education between Switzerland and other countries.

Table 4-2: IV Estimates: Work-based Education vs. School-based Education

Specification	Variable	Intrinsic Work Motivation	Task-Centered Coping	Contact-Centered Coping	Emotional Coping
Baseline					
AREA+CULTURE	Apprentice	-0.624*	0.174	0.201*	-0.607***
		(0.335)	(0.136)	(0.103)	(0.179)
	Kleibergen	34.797	34.797	34.797	34.797
Baseline					
LDV	Apprentice	-0.046	0.014	0.216***	-0.468***
		(0.048)	(0.063)	(0.071)	(0.073)
	Kleibergen	21.468	22.529	21.948	20.616
Baseline					
LDV+AREA+CULTURE	Apprentice	-0.037	0.022	-0.020	-0.309***
		(0.069)	(0.097)	(0.085)	(0.088)
	Kleibergen	33.809	34.785	34.539	34.458
WEIGTHED					
LDV+AREA+CULTURE	Apprentice	-0.048	-0.118	-0.111	-0.329***
		(0.119)	(0.121)	(0.246)	(0.074)
	Kleibergen	15.325	16.026	15.545	15.729
NOCONTROL					
LDV+AREA+CULTURE	Apprentice	-0.132	0.136	-0.037	-0.374**
		(0.088)	(0.122)	(0.094)	(0.158)
	Kleibergen	23.657	24.653	24.909	23.730
DELTASHARE					
LDV+AREA+CULTURE	Apprentice	-0.061	0.091	0.044	-0.392***
		(0.073)	(0.074)	(0.140)	(0.110)
	Kleibergen	7.596	8.274	8.206	7.521
KV					
LDV+AREA+CULTURE	Apprentice	-0.173*	0.256***	-0.084	-0.168
		(0.102)	(0.086)	(0.115)	(0.170)
	Kleibergen	15.161	15.272	15.091	15.445
NOGA					
LDV+AREA+CULTURE	Apprentice	0.061	0.078	0.005	-0.290***
		(0.121)	(0.070)	(0.100)	(0.047)
	Kleibergen	27.025	27.778	27.214	27.509
INDSHARE					
LDV+AREA+CULTURE	Apprentice	-0.046	0.017	-0.088	-0.226***
		(0.061)	(0.123)	(0.101)	(0.087)
	Kleibergen	9.211	9.871	9.872	9.692

ORIGIN					
LDV	Apprentice	-0.019	0.179***	0.313**	-0.191***
		(0.046)	(0.031)	(0.138)	(0.035)
	Kleibergen	34.595	37.326	35.074	34.986
ORIGIN+CANTON					
LDV	Apprentice	-0.061	0.040	0.243***	-0.453***
		(0.043)	(0.049)	(0.077)	(0.071)
	Kleibergen	24.598	25.966	27.027	25.956
	Hansen p-value	0.663	0.114	0.490	0.131

Notes: N=4442. The table displays coefficients and standard errors clustered at the cantonal level in parentheses of an IV estimation with the binary endogenous variable Apprentice and the share of general school students in the canton in 1998 as instrument. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The sample based on the TREE dataset consists of students continuously enrolled in an educational track between 2001 and 2003. Kleibergen refers to the Kleibergen-Paap F statistic, which has a critical value of 16.38 for 10% maximal IV size. All estimates include year dummies and the control variables described in Table 4-7. AREA and CULTURE estimates include dummy variables for 7 large areas in Switzerland and for whether the native language is German, French or Other. LDV refers to estimates that include the lagged dependent variable. NOCONTROL omits observable characteristics from the estimation. DELTASHARE includes a control for the share of general school students in the canton in 1980. KV restricts the treatment group of apprentices to commercial apprenticeships (N=3146). NOGA estimates include dummy variables for the first industry (N=3096). INDSHARE estimates further control for the cantonal employment share in 2-digit industries. ORIGIN estimates use the share of work-based education in the country of origin as instrument and displays block-bootstrapped standard errors at the country of origin level. Kleibergen refers to the Kleibergen-Paap F statistic, which has a critical value of 16.38 for 10% maximal IV size. These estimates exclude Language, Swiss and Time Swiss[17] as control variables. CANTON+ORIGIN uses both instruments simultaneously and displays robust standard errors clustered on the level of the canton. Hence, the Kleibergen statistic has a critical value of 19.93. Hansen p-value refers to the p-value of a Hansen overidentification test.

4.5.1 Extensions

The following paragraphs extend the analysis of the effect of work-based education on the personality skills of adolescents in two directions. First, we compare the effect of work-based education on female and male students. Second, we use information in 2007 and 2010 to evaluate whether the estimated effects are merely transitory or whether work-based education shifts personality skills permanently.

4.5.2 Heterogeneity of the Treatment Effect

This paragraph analyzes whether the impact of work-based education differs between men and women. To this end, Table 4-3 displays the pooled OLS results including the lagged dependent variable, the baseline IV estimates with the cantonal shares of general upper secondary education as instrument and the corresponding IV estimates that control for the lagged dependent variable for the samples of men and women separately.

Regarding emotional stability, Table 4-3 suggests that emotional stability is increased for both women and men compared to school-based general education. However, compared to vocational school emotional stability of women is more affected than the emotional stability of men. The decreasing effect on openness is statistically significant on the one percent levels for males, but is not significant for females. Further the results indicate, that females profit from work-based education with an

[17] Including these control variables doesn't affect the estimates qualitatively but the Kleibergen statistic drops beneath 1, questioning the validity of the approach.

increase in conscientiousness. Agreeableness has similar effect sizes for women and men, though statistical significance is more stable for women. Splitting the sample in this way suggests that extraversion is affected for neither women nor men.

Analyzing the heterogeneity of the effects between females and males reveals that work-based education compared to school-based education decreases emotion-centered coping for females more than for males. The decreasing effect on intrinsic work motivation seems to be driven through males. The significance of the effect on contact-centered coping is more stable for females, though the effect size is similar.

Table 4-3: Heterogeneity of Treatment Effect between Men and Women: Work-based Education vs. School-based Education

		Intrinsic Work Motivation	Task-Centered Coping	Contact-Centered Coping	Emotional Coping
Men					
OLS LDV	Apprentice	-0.061	-0.081**	0.000	-0.128***
		(0.038)	(0.039)	(0.039)	(0.036)
IV LDV+AREA+CULTURE	Apprentice	-0.081	-0.077	-0.143	-0.329***
		(0.178)	(0.116)	(0.233)	(0.121)
	Kleibergen	11.991	12.040	12.119	11.688
Women					
OLS LDV	Apprentice	-0.067**	0.006	0.019	-0.085***
		(0.033)	(0.037)	(0.034)	(0.033)
IV LDV+AREA+CULTURE	Apprentice	0.028	0.177**	0.006	-0.273**
		(0.068)	(0.089)	(0.089)	(0.129)
	Kleibergen	43.427	42.495	45.051	43.760

Notes: N=1878 for men and 2564 for women, respectively. The OLS LDV estimates display coefficients and standard errors clustered at the individual level for estimates using OLS with lagged dependent variables. The IV estimates display coefficients and standard errors clustered at the cantonal level in parentheses of an IV estimation with the binary endogenous variable Apprentice and the share of general school in the canton as instrument. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The sample based on the TREE dataset consists of students continuously enrolled in an educational track between 2001 and 2003. Kleibergen refers to the Kleibergen-Paap F statistic, which has a critical value of 16.38 for 10% maximal IV size. All estimates include year dummies and the control variables described in Table 4-7 in addition to the lagged dependent variable, dummy variables for the area and for language.

4.5.3 Long Run Effects

However, in the long-term, the initial impact of education on personality skills may diminish or even disappear, e.g. because students might start to work after the school-based upper secondary education. Therefore, this section analyses whether the differences still exist in 2007 and 2010, i.e. about four to seven years after concluding secondary education. Table 4-4 shows the development of the dependent variables for an extended period of time. Our results indicate that the impact on the personality skill emotional-centered coping remains in the long run, while the long-run evidence further questions the statistically less robust increase of contact-centered coping and the decrease of intrinsic work

motivation. This finding is consistent with idea that personality skills are more malleable during adolescence than during early adulthood.

Table 4-4: Transcience vs. Persistence of the Effect: Estimates: Work-based Education vs. School-based Education

		Intrinsic Work Motivation	Task-Centered Coping	Contact-Centered Coping	Emotional Coping
OLS LDV 2007	Apprentice	-0.277***	-0.093*	-0.010	-0.167***
		(0.060)	(0.051)	(0.053)	(0.052)
IV 2007					
LDV+AREA+CULTURE	Apprentice	-0.407	-0.532**	0.442	-0.689***
		(0.257)	(0.234)	(0.380)	(0.226)
	Kleibergen	25.101	24.360	25.277	24.626
OLS LDV 2010	Apprentice	-0.322***	-0.170***	-0.113*	-0.180***
		(0.071)	(0.051)	(0.058)	(0.056)
IV 2010					
LDV+AREA+CULTURE	Apprentice	-0.183	0.069	0.613**	-0.788**
		(0.312)	(0.215)	(0.304)	(0.384)
	Kleibergen	29.084	28.040	29.612	28.879

Notes: N(2007)=1707 and N(2010)=1454. The table displays coefficients and standard errors clustered at the cantonal level in parentheses of an IV estimation with binary endogenous variable and the share of general high school in the canton as instrument. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively. The sample based on the TREE dataset consists of students continuously enrolled in an educational track between 2001 and 2003. The sample refers to the year 2007 and 2010, respectively. Kleibergen refers to the Kleibergen-Paap F statistic, which has a critical value of 16.38 for 10% maximal IV size. All estimates include year dummies and the control variables described in Table 4-7 in addition to the lagged dependent variable, dummy variables for the area and for language.

4.6 Conclusion

Recent evidence documents that personality skills predict a wide range of life outcomes including educational achievement and labor market outcomes. Hence, information about how education impacts personality skills is crucial. Following the hypotheses of Heckman and Kautz (2013) that work-based education may involve the teaching of valuable personality skills, we provide first evidence regarding the effect of work-based secondary education compared to school-based secondary education on personality skills.

We make use of a large representative PISA 2000 follow-up sample in Switzerland (TREE) and apply an IV approach to control for endogeneity in the growth of personality skills. Identification in our model results from the fact that the share of general secondary education between cantons in Switzerland varies substantially. These differences reflect historically shares set by government, which remain persistent over the last 20 to 30 years. However, since the regional differences in these shares could be correlated with other features of regions related to personality skills, we apply several

robustness checks. But finally some concerns about the exclusion assumptions remain and therefore more research is needed.

The evidence in this chapter indicates that education can change personality skills. Our estimates suggest that work-based education decreases emotional-centered coping statistically and economically significant. Intrinsic work motivation might be decreased as well and contact-centered coping might be increased, though these effects are less robust in terms of the econometric specification.

4.7 Appendix

This section displays the tables complementing the tables of the main analysis shown in the chapter, i.e. the factor analyses, variable definitions and summary statistics, conditional correlations, information about the instruments, complete estimation results including control variables and further robustness checks of the IV strategies.

Table 4-5: Dependent Variables

Variable	Questionnaire Items
Intrinsic Work Motivation	Thinking about the future, how important is it to have a job, where I have a lot of contact with other people (4-point Likert Scale).
	Thinking about the future, how important is it to have a job, where I can help other people (4-point Likert Scale).
	Thinking about the future, how important is it to have a job, which gives me the feeling of doing something sensible (4-point Likert Scale).
Task-centered coping	When I am stressed or find myself in a difficult situation, I focus on the problem and see how I can solve it (5-point Likert Scale).
	When I am stressed or find myself in a difficult situation, I think about the event and learn from my mistakes (5-point Likert Scale).
Contact-centered coping	When I am stressed or find myself in a difficult situation, I try to be with other people (5-point Likert Scale).
	When I am stressed or find myself in a difficult situation, I visit a friend (5-point Likert Scale).
Emotional Coping	When I am stressed or find myself in a difficult situation, I get angry (5-point Likert Scale).
	When I am stressed or find myself in a difficult situation, I feel anxious about not being able to cope (5-point Likert Scale).
	When I am stressed or find myself in a difficult situation, I blame myself for not knowing what to do (5-point Likert Scale).
	When I am stressed or find myself in a difficult situation, I wish I could change what has happened (5-point Likert Scale).

Table 4-6: Rotated Factor Loadings of Items for Defining Constructs of Personality Skills

Item	Emotional Coping	Intrinsic Work Motivation	Contact-Centered Coping	Task-Centered Coping
Intrinsic Work Motivation 1	0.0509	**0.7347**	0.3330	-0.0208
Intrinsic Work Motivation 2	-0.0661	**0.8250**	0.0800	-0.0154
Intrinsic Work Motivation 3	-0.0723	**0.6806**	-0.0983	0.1956
Task-centered coping 1	0.0905	-0.0130	-0.0147	**0.7878**
Task-centered coping 2	0.0247	0.0860	0.1144	**0.7956**
Contact-centered coping 1	0.0702	0.1140	**0.8278**	0.0825
Contact-centered coping 2	-0.0697	0.0823	**0.8417**	0.0080
Emotion-centered coping 1	**0.5390**	-0.0584	-0.0004	0.2835
Emotion-centered coping 2	**0.7594**	-0.0473	0.0018	0.0915
Emotion-centered coping 3	**0.7780**	-0.0091	-0.0049	0.0107
Emotion-centered coping 4	**0.6957**	-0.0270	0.0082	-0.0133

Table 4-7: Explanatory Variables

Variable Name	
Endogenous Variable	
Apprenticeship	Dummy variable that takes the value 1 if an individual is continuously enrolled in apprenticeship training (work-based secondary education) between 2001 and 2003, and 0 otherwise.
Control Group	
School-based Education	Dummy variable that takes the value 1 if an individual is continuously enrolled in full-time school based secondary education between 2001 and 2003, and 0 otherwise.
Control Variables	
PISA Read	PISA score in reading in the year 2000
Books	Variable taking values 1 to 7 for 0, 1-10, 11-50, 51-100, 101-250, 251-500, more than 500 books at home in 2000.
ISEI Father	Social status of father according to ISEI in 2000
Age	Age of the individual
Male	Dummy variable that takes the value 1 if the individual is male, and 0 otherwise.
Male*Age	Interaction term of Age and Male
Urban	Dummy variable that takes the value 1 if the individual lives in an urban area in 2000, and 0 otherwise.
Family Structure	Dummy variables that take the value 1 for nuclear, mixed and other family structures, and 0 otherwise. Single is the base category.
Education Mother	Dummy variables that take the value 1 if the mother has the highest education of ISCED2, ISCED3B/ISCED3C and ISCED3A, and 0 otherwise. Mother's education of ISCED5A/ISCED5B/ISCED6 represents the base category.
Live with Parent	Dummy variable that takes the value 1 if the individual lives with at least one parent, and 0 otherwise.
Language	Dummy variable that takes the value 1 if the individual speaks the PISA test language at home in 2000, and 0 otherwise.
Swiss Born	Dummy variable that takes the value 1 if the individual was born in Switzerland, and 0 otherwise.
Swiss Time	Number of years living in Switzerland
Catholic Share	Cantonal share of Catholic inhabitants
Instruments	
Canton 1998	Canton average of the share of general secondary education degrees in 1998
Canton 1980	Canton average of the share of general secondary education degrees in 1980
Country 1998	1998 share of work-based education[18] in the country (CH, DE/AT, FR/BE, IT, ES, PT, YU, TR, OTHER) the individual was born. Due to missing values, YU and OTHER are set to 0.

[18] Based on the OECD indicator "Students enrolled by type of institution" available at http://stats.oecd.org/.

Table 4-8: Summary Statistics of Dependent and Control Variables

	Apprenticeship (Work-Based Education)					School-Based Education				
Variable	Obs	Mean	Std. Dev	Min	Max	Obs	Mean	Std. Dev	Min	Max
Intrinsic 2001	921	-0.16	0.96	-4.31	1.89	1300	0.04	0.99	-4.35	1.72
Task-Centered 2001	921	0.08	0.93	-3.80	2.20	1300	0.03	0.98	-3.67	2.30
Contact-Centered 2001	921	-0.12	0.98	-2.73	2.95	1300	0.06	0.99	-2.86	2.68
Emotion-Centered 2001	921	-0.16	0.96	-4.31	1.89	1300	0.04	0.99	-4.35	1.72
Intrinsic 2002	921	-0.21	0.87	-4.28	1.89	1300	0.04	0.90	-3.82	1.79
Task-Centered 2002	921	-0.01	0.90	-3.80	2.16	1300	-0.01	1.00	-3.77	2.45
Contact-Centered 2002	921	-0.27	0.95	-2.79	2.83	1300	0.06	0.98	-2.54	2.63
Emotion-Centered 2002	921	-0.21	0.87	-4.28	1.89	1300	0.04	0.90	-3.82	1.79
Intrinsic 2003	921	-0.29	1.02	-4.48	1.78	1300	0.03	1.03	-3.66	1.57
Task-Centered 2003	921	0.02	0.81	-3.55	2.45	1300	0.06	0.82	-2.77	2.24
Contact-Centered 2003	921	-0.27	0.87	-2.73	2.50	1300	0.03	0.89	-2.47	2.65
Emotion-Centered 2003	921	-0.29	1.02	-4.48	1.78	1300	0.03	1.03	-3.66	1.57
PISA Read	1842	514.76	73.21	256.74	738.72	2600	573.75	68.33	323.89	804.7
Books	1842	4.50	1.42	1	7	2600	5.30	1.39	1	7
ISEI Father	1842	42.81	15.37	16	90	2600	53.23	18.09	16	90
Age	1842	18.37	0.81	17	22	2600	18.13	0.81	16	22
Male	1842	0.56	0.50	0	1	2600	0.33	0.47	0	1
Urban	1842	0.56	0.50	0	1	2600	0.72	0.45	0	1
Single Family	1842	0.08	0.27	0	1	2600	0.09	0.29	0	1
Nuclear Family	1842	0.85	0.36	0	1	2600	0.86	0.35	0	1
Mixed Family	1842	0.05	0.21	0	1	2600	0.03	0.16	0	1
Other Family	1842	0.02	0.15	0	1	2600	0.02	0.13	0	1
ISCED2	1842	0.24	0.43	0	1	2600	0.11	0.31	0	1
ISCED3B/3C	1842	0.59	0.49	0	1	2600	0.50	0.50	0	1
ISCED3A	1842	0.16	0.37	0	1	2600	0.38	0.49	0	1
ISCED5A/5B/6	1842	0.01	0.08	0	1	2600	0.01	0.10	0	1
Live Parent	1842	0.89	0.32	0	1	2600	0.91	0.29	0	1
Language	1842	0.11	0.32	0	1	2600	0.12	0.32	0	1
Swiss	1842	0.92	0.27	0	1	2600	0.91	0.29	0	1
Swiss Time	1842	14.94	2.20	1	17	2600	14.69	2.04	1	17
Catholic Share	1842	47.71	21.09	16	81.2	2600	50.88	22.12	16	81.2

Notes: Dependent variables are shown for each year separately. Control variables refer to the year 2002 and 2003.

Table 4-9: Summary Statistics of Instruments

Canton	N	Area	1980	1998
ZH	96	4	12.5	18.90
BE	229	2	7	13.30
LU	32	6	5.8	11.80
UR	0	6	8.6	11.50
SZ	21	6	5.9	11.90
OW	34	6	6.3	10.60
NW	14	6	5.6	17.50
GL	2	5	10.3	16.00
ZG	19	6	10.7	15.10
FR	170	2	10	20.50
SO	34	2	9	13.90
BS	25	3	18.2	21.10
BL	51	3	16.5	21.10
SH	27	5	6.5	18.80
AR	4	5	7.7	14.60
AI	0	5	6.3	12.70
SG	253	5	6.1	12.60
GR	26	5	7.9	12.50
AG	121	3	9.5	16.30
TG	33	5	6.1	10.50
TI	273	7	17	26.00
VD	136	1	12.5	20.90
VS	161	1	8.4	19.60
NE	135	2	13.5	24.00
GE	235	1	21.3	31.80
JU	90	2	9	25.40
Country	**N**			**1998**
CH	2031			0.58
DEAT	9			0.47
ES	5			0.03
FRBE	14			0.11
IT	10			0.00
PT	35			0.00
TR	7			0.00
YU	51			0.00
OTHER	59			0.00

Chapter 5

Teaching in Vocational Education as a Second Career

This chapter has been published in another version with Mirjam Strupler Leiser as co-author under the Title "Teaching in Vocational Education as a Second Career", in Empirical Research in Vocational Education and Training (2014). 6(8)

5.1 Introduction

The present study investigates the determinants governing change to a career in teaching. The question of what motivates individuals to forfeit their original occupation to become a teacher is important for policy making not only in times of teacher shortages but also in light of the quality of individuals that can be motivated to change their career mid-life to become a teacher.

More specifically, we use a particular feature of the Swiss vocational education and training (VET) system to study career change – that teachers of job-related subjects cannot choose teaching as their first career but can only become teachers after (a) acquiring the highest education qualification in their job category and (b) accumulating a certain number of years of job experience. Thus, all teachers of vocational subjects have changed careers. This special feature of the Swiss education system allows us to analyze a large number of career changers turning to teaching and to draw conclusions about how educational systems can attract high-quality career changers into teaching.

For our empirical analyses, we assess a unique data set of teachers who decided to change their careers to teaching. The data set is representative of the whole German- and French-speaking Switzerland.[1] For our comparative analyses, we use data on the workforce from the Swiss Labor Force Survey (SLFS). We use the former wage position compared with the average wage of similar individuals in the former occupation of the teachers as an indicator for their performance in their former occupation as well as for teaching quality.[2] With this approach, we follow Chingos and West (2012), who showed that better teachers leaving the teaching occupation earn more in the outside career and therefore provide evidence of a positive correlation between teaching quality and earnings outside teaching.

---

[1] German and French-speaking Switzerland accounts for more than 95% of the Swiss population.

[2] This implies the assumption that wage in the former occupation is positively correlated with productivity in the former occupation and commensurately also in teaching.

Our results show first that those who change careers to teaching (who are on average 40 years old) do not change careers because of a lack of (financial) success in their original career. Although the results are unsurprisingly heterogeneous, we can nevertheless explain at least part of this heterogeneity. Second, although the average teacher tends to rank among the higher earners in their original career, the majority of career changers expect to earn more as a teacher than in their original career. Again, we find a substantial heterogeneity: around one-third of career changers expect a wage reduction after changing.

The chapter is structured as follows. The next section contextualizes our research with a description of the recruiting procedures for Swiss vocational education and training teachers followed by a brief review of the relevant literature. The "Method and Data" section describes our data and outlines our research questions and estimation strategy. The "Results" section presents the empirical results and the final section concludes.

5.1.1 Recruitment of Vocational Teachers in Switzerland

The legal provisions governing the hiring of teachers for instruction in vocational subjects require candidates to meet two conditions. First, candidates must have the highest possible training qualification in the particular occupation, which in most cases is a professional education at tertiary level B (ISCED 5B)[3] or even an academic degree at tertiary level A (ISCED 5A).[4] Second, they must have very good subject knowledge, i.e., a minimum of six months of occupational experience, with several years being the norm. Most teachers have a long history of job experience because the highest possible training qualification generally consists of practically oriented and occupationally specific training at an advanced level (57% of teachers in our sample have tertiary level B professional education and training), all of which involves many years of job experience. Extensive job experience is also apparent in our sample, with the average age of teachers being over 40 (see Table 5-1). Meeting all of these requirements an individual can be recruited by a vocational school as a teacher. Therefore, all vocational teachers in Switzerland are career changers. Only after recruitment – and besides teaching – teacher training starts. Individuals begin teacher training for full-time or sideline positions. Full-time teachers can teach full-time or part-time at vocational schools. Qualification for sideline teaching, by contrast, consists of much fewer training hours and, therefore, the certificate limits these

[3] Tertiary level B professional education and training (PET) consists of the Federal PET Diploma Examination, the Advanced Federal PET Diploma (also referred to as Meisterprüfung) and PET Colleges (SKBF, 2011). Professional (occupational) college programs last two to three years full- or part-time, the duration of the Federal PET Diploma Examination and Advanced Federal PET Diploma is unspecified as attendance of the preparation courses is not compulsory. However, they require a certain number of years' work experience in the relevant occupation.

[4] However, the sample also contains some teachers whose highest previous educational qualification was only upper secondary education (8%). Either these teachers come from professions with no tertiary education programs, or the vocational school recruited the teachers despite their low academic qualifications because of a shortage of teachers in a particular field.

teachers to teaching part-time and not full-time. We use the term sideline teachers to distinguish the teachers from those who are trained for full-time teaching but who may choose to teach part-time. For sideline teachers, their work at vocational school is precisely that – on the sideline of their work in their original occupation. The teachers are recruited from occupations being taught to students at the vocational schools. Therefore, the original occupations of the teaching staff reflect the respective apprenticeship market. The strong rooting of apprenticeship in the manufacturing sector also explains why more (one-third) vocational teachers than the working force average worked in industry and manufacturing prior to changing careers into teaching. One crucial factor directly affecting the monetary appeal of teaching at vocational schools is the competitiveness of the original occupations and business sectors in terms of the prevailing pay and working conditions.

5.1.2 Literature Review

The existing literature on the choice of teaching as a (first) career investigates both factors governing candidate quality and aptitude and factors governing the quantity of teachers available in the workforce. Economic literature tends to focus on the relative wage as a factor influencing the quantitative and qualitative supply of teachers on the labor market. Non-monetary factors have also been studied but only to some extent because these factors are usually much more difficult to address empirically.

Most studies demonstrate a positive wage elasticity of labor supply (Chevalier, Dolton, & Mcintosh, 2007; Denzler & Wolter, 2009; Dolton, 1990; Dolton & Chung, 2004; Falch, 2010; Manski, 1987). These findings may be related to labor supply elasticity being influenced to some extent by how high the wage differential is in absolute terms. Labor supply elasticity appears very high in cases where teachers earn less than individuals in similar occupations, whereas elasticity is relatively low where teachers tend to earn more.

For the qualitative selection of the teaching occupation, the results of known empirical studies are less conclusive. Whereas US studies find ample evidence of negative selection in terms of cognitive criteria (see, e.g., Corman, 1993; Hanushek & Pace, 1995; Manski, 1987; Podgursky, Monroe, & Watson, 2004; Stinebrickner, 2001) the results are less conclusive for German-speaking countries (Denzler & Wolter, 2009).

What makes applying these data most difficult to the subject matter explored in this chapter is that all of these studies focus on why people select teaching as a first career after graduation, i.e., there is a deficiency of comparable empirical studies investigating why people select teaching as a second career.

Therefore, to construct the hypotheses on career change for teachers we also use career change literature to explore the predictive factors of career change. Standard search and matching models

(Burdett, 1978; Jovanovic, 1979; Mortensen, 1987; Neal, 1999) start by assuming that labor markets feature heterogeneous employers and employees as well as imperfect information. In these models, employee productivity is highest where there is the perfect match to the specific job. Because neither employer nor employee will know the optimal match in advance, employees will keep changing jobs until they achieve the perfect match. As a consequence of this search process, changing jobs correlates with increasing wages (Rubinstein & Weiss, 2006). However, individuals do not continuously change employers and careers because any change involves a loss of human capital and, therefore, of productivity and wage.

According to the standard human capital theory (Becker, 1962), assuming that wage corresponds to the workers' productivity, jobs are associated with the acquisition of employer-, occupation- and industry-specific human capital that may be forfeited with a change of employer or and even more a change of career. The better match in the new job would therefore have to raise the productive value of general human capital enough to compensate for the loss in employer- and job-specific human capital. However, increased employer- and job specific skills lower turnover intentions (or a change of career) as employer-specific skills are less valuable to other employers (Doeringer & Piore, 1971). Corresponding to the logic of human capital theory, a change is all of the more unlikely the longer the period of investment in employer- and job-specific human capital. One should therefore be able to observe a lower incidence of job changes (with or without career changes) as a function of seniority. Furthermore, in addition to the standard human capital theory, the strategy of backloading the compensation profile (i.e., paying the worker less than his marginal productivity when young and more when old to increase workers motivation over the whole working career and to alleviate monitoring problems) also explains a decreasing propensity in employer change with seniority (Daniel & Heywood, 2007; Heywood, Jirjahn, & Tsertsvardze, 2010; Lazear, 1979, 1981).

Refinements of the human capital theory (e.g., the skills weight theory, see Lazear, 2009) assume the existence of no general or specific human capital but only of different combinations of skills. These refinements suggest that, regardless of the existing duration of employment, mobility between employers, occupations and industries can still be high provided that the potential employment alternative requires a similar mix of skills (see, e.g., Geel, Mure, & Backes-Gellner, 2011). For our hypotheses, this additional factor is relevant because a vocational teacher's job not only calls for levels of expertise similar to those required in the former occupation but also requires above-average expertise levels (i.e., a long history of skill-building) in the original occupation. As we can assume that a large amount of the expertise accumulated in the former occupation can be transferred to the new one (teaching), it is likely that the probability of changing a career to teaching will not correlate negatively with seniority. In contrast, individuals who changed employers frequently also tended to be those who had already changed careers once or several times. The lack of consistency in their employment history makes it more difficult for these individuals to enter the vocational teaching

occupation because they do not fulfill the relevant requirements (see the section on the recruitment of vocational teachers in Switzerland).

Furthermore, empirical work on the role of job characteristics (also non-monetary characteristics) suggests that labor supply is affected by the specific characteristic of a job (e.g., Altonji & Paxson, 1986; Atrostic, 1982; Kunze & Suppa, 2013). More favorable working conditions affect job and life satisfaction (e.g., Cornelissen, 2009; Luechinger, Meier, & Stutzer, 2010) and job satisfaction or dissatisfaction may explain job changes (e.g., Clark, 2001; Cornelissen, 2009). Furthermore, job characteristics can also explain wage differentials (e.g., Wells, 2010). Therefore, in addition to earnings prospects, job quality in teaching may motivate individuals to change to teaching.

As for forecasting based on search and matching models, individuals will want to change to the teaching occupation only if they expect it to be a better match. However, the extent to which a higher wage is expected to be part of that better match remains unclear. The reason is first, we do not know the relative position of different jobs regarding the job characteristics and hence non-monetary benefits. Second, the assumption that higher productivity will translate to a higher wage (wage reflects productivity) does not automatically apply in the public sector, where schooling takes place. One feature of vocational schools is that all teachers receive the same wage, depending on age, canton, experience and training, independent of their original occupation. Wages are set by cantonal laws, and schools therefore have no room for maneuver in wage setting. The financial attractiveness of teaching therefore depends essentially on the wage level of the individual's original occupation. Therefore, we expect that schools can choose among several candidates for teaching positions from occupations with relatively low wage levels, whereas schools will face difficulties finding suitable candidates from occupations with high compensation levels.

Furthermore, we hypothesize that teachers change to teaching when their cumulated future compensation bundle (monetary and non-monetary benefits) is superior in the teaching occupation compared with their former occupation.

5.2 Methods and Data

5.2.1 Data

Because they are career changers, teachers in vocational education need to complete a pedagogical education in their first years of teaching in addition to teaching. We conducted the survey among all teacher trainees for vocational education at the Swiss Federal Institute for Vocational Education and Training (SFIVET). At the time of this study (spring semester 2010), three tertiary institutions were offering vocational teacher training in Switzerland but SFIVET, which provided the sample used for this chapter, had a market share of over 80% of vocational teacher trainees.

Because we conducted the survey during classes, we achieved a response rate of 100%. The teachers completed the survey using either a computer-assisted questionnaire or, if no computer was available, an identical paper and pencil version. We tested the questionnaire in an extensive pre-test on teachers who had previously trained at SFIVET.

The information elicited in the survey comprises personal details, training, job experience, wage and wage expectations of the 483 respondents in German- and French-speaking Switzerland. About one half of the 483 teachers (230) were pursuing a degree qualifying them to be full-time vocational teachers. The other half of the respondents (253) was working toward a certificate qualifying them to be sideline vocational teachers. Despite the 100% response rate, some data were missing on account of item non-response. We excluded 93 observations (19%) from the analysis because of missing wage information or other important data, leaving a final dataset of 390 vocational teachers. Item non-response analyses show that the exclusion of the 93 observations should not influence or bias our results.

To compare those individuals who have chosen[5] to change career with similar individuals who have chosen to remain in their initial occupation, we sourced a comparison group data from the Swiss Labor Force Survey (SLFS). Because the vocational teachers in our sample opted to change to teaching at different points in time, we used three different cohorts (2004, 2006 and 2008) of the SLFS. To obtain the comparison group for the teachers, we excluded all individuals who would not have been able to become a teacher, i.e., the unemployed, pensioners, students, individuals without compulsory-post-schooling qualifications and individuals who were under 20 years old prior to our analyses. Details of each of the variables for the vocational teacher and SLFS subjects appear in Table 5-1. Table 5-5 in the Appendix provides a description of the variables.

[5] Whether movers from other occupations had quit or had been laid off may have a substantial impact on selection into teaching jobs. However, only six percent of the career changers were involuntarily unemployed within three years before changing to teaching and the expected wage gains/losses from career change do not differ significantly between the two groups.

Table 5-1: Descriptive Statistics of the Sample

Attributes	All teachers		Teachers without self-employed		Sideline teachers		Primary occupation of teachers		Swiss Labor Force Survey (Total)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female (dummy)	0.33	0.47	0.33	0.47	0.40	0.49	0.25	0.34	0.46	0.50
Age	40.51	7.28	40.1	7.25	40.67	7.57	40.34	6.98	44.38	10.19
Gross yearly wage in CHF	90,741	31,146	30,836	25,546	91,825	35,028	89,611	26,552	82,538	29,537
Education (share)										
Upper secondary education	0.08	0.27	0.08	0.27	0.09	0.28	0.07	0.26	0.63	0.48
ISCED 5b	0.57	0.50	0.54	0.50	0.62	0.49	0.51	0.50	0.12	0.32
ISCED 5a+6	0.35	0.48	0.38	0.49	0.30	0.46	0.41	0.49	0.26	0.44
Occupational field (share) in the non-teaching occupation										
Agriculture, forestry and livestock breeding	0.03	0.16	0.02	0.15	0.03	0.16	0.03	0.16	0.02	0.15
Industry and manufacturing	0.28	0.45	0.28	0.45	0.23	0.24	0.33	0.47	0.12	0.32
Engineering and informatics	0.17	0.38	0.17	0.37	0.16	0.36	0.19	0.39	0.12	0.32
Construction sector and mining	0.08	0.27	0.07	0.26	0.10	0.30	0.05	0.22	0.06	0.23
Trade and transport	0.05	0.21	0.05	0.21	0.03	0.17	0.06	0.24	0.15	0.36
Hospitality industry and service sector	0.09	0.29	0.08	0.27	0.12	0.33	0.06	0.23	0.07	0.26
Management, administration, banking, insurance and law	0.08	0.27	0.09	0.28	0.06	0.24	0.10	0.31	0.22	0.41
Health services, culture, science	0.23	0.42	0.25	0.42	0.28	0.45	0.18	0.38	0.25	0.43
Children (yes/no, dummy)	0.63	0.48	0.64	0.48	0.59	0.49	0.66	0.47	0.59	0.49
Leadership position (yes/no, dummy)	0.66	0.47	0.65	0.48	0.72	0.45	0.60	0.49	0.39	0.49
Tenure	10.15	6.88	9.60	6.58	10.65	7.14	9.62	6.56	7.39	7.14
Self-employed (yes/no, dummy)	0.15	0.36			0.19	0.39	0.12	0.32	0.11	0.32
Sideline teacher (yes/no, dummy)	0.51	0.50	0.49	0.50	1.00		0.00			
Observations	390		330		199		191		36251	

5.2.2 Empirical Strategy

First, we compare how much the teachers earned relative to others with the same characteristics in their original occupation. This shows us whether those who choose teaching earn an average, above-average or below-average wage in their former occupation.

Although automatically inferring suitability for teaching from what a person earned in their original occupation is not possible, a somewhat direct relationship nonetheless exists between productivity as measured by earnings in the former occupation and suitability for vocational teaching. This relationship is probably even stronger for vocational teaching than for other teaching categories or occupations. The main task of vocational teachers is to teach young people occupation-specific knowledge relevant to the teacher's former occupation. Therefore, if we assume that individuals whose skills are above average in their original occupation are more productive and earn commensurately more, then a positive selection should also be beneficial to vocational education.

Our analysis is a form of inversion of that presented by Chingos and West (2012), who investigated, of the teachers who left teaching, whether those who earned more in the new occupation had also been the better teachers (i.e., obtained better student performance). Chingos and West (2012) identified a positive correlation, which they interpreted to indicate that the same skills that made this group of teachers more productive in the educational system also led to higher wages and therefore higher productivity in other occupations.

We identify a comparison group by matching each teacher with individuals from the SLFS who do not differ from our teachers in terms of key characteristics, such as gender, age, education and occupational field. Because we have a very large pool of non-teaching individuals from the SLFS, we are able to match each teacher with several non-teaching individuals with the same characteristics.[6] For each teacher, we estimate a comparable wage in the original occupation by averaging the salaries from SLFS individuals who work in the same occupational field, who hold the same degree of highest education and who are of the same gender and age. These average wages are then compared with the teachers' wages in their initial occupation (outside education).[7] Obviously, our analysis allows us to match only the observable characteristics. However, the individuals might differ in terms of ability, motivation or personality, and it is not within the scope of this study to control for such differences. Therefore, our results have to be interpreted as being the relative wage position of a teacher compared with the average wage of individuals with the same observable characteristics.

---

[6] Therefore, we use exact matching (Abadie, Herr, Imbens, & Drukker, 2004).

[7] The matching strategy is similar to an approach where residuals from a wage regression (for SLFS individuals and teachers) would be examined and actual wages and predicted wages would be compared with see whether teachers earned more or less than their peers (predicted wage) in their original occupation. The advantage of the matching strategy is that no assumptions about the functional form of the wage regression are imposed.

If (a) a correlation exists between the occupational abilities and wages in the economy, (b) the hiring authority (usually the head teacher) imputes a correlation between occupational ability and teaching ability, (c) the hiring authority is in a position to choose among different candidates and (d) teaching wages are on average competitive with wages outside education, then teachers who have earned more than comparable colleagues in their initial occupation should be a positive selection into teaching. It is clear that the fourth point, i.e., that teachers' wages are, on average, competitive with other wages, will not be equally met for all occupations and business sectors because wages vary substantially between sectors but are relatively uniform in teaching, depending on canton, individual's training, age and experience. We therefore form the hypothesis that individuals have less of an incentive to change to teaching from occupations and industries with average salaries as high as or higher than in teaching. Accordingly, teachers from these occupations and industries would on average be counted among the lower earners in their sector (otherwise teacher pay is not competitive), whereas exactly the opposite would apply in less well-paid occupations and industries. Obviously, there is the possibility that some high performing individuals from high paying industries with wages above the teaching salary change career to teaching and accept a wage cut, for example, because of better working conditions in teaching. However, we expect that, on average, teachers from high paying industries rank among the lower earners in their former occupation.

Second, we investigate the wage prospects of career changers who have opted to become teachers. Teachers can earn more or less as teachers than they would have earned had they decided to stay in their former occupation.

Those who opt to become teachers are unlikely to represent a random sample of all individuals who could theoretically become teachers. Thus, a simple comparison of teachers' wages with average alternative wages is not a useful way of learning whether the decision to change to teaching pays off financially. This study explores therefore the counterfactual situation to the decision to become a teacher by surveying teacher's expectations on both options. The questionnaire asked teachers to indicate their wage expectations for two scenarios: first, expected wages (five and 10 years after training) if they stay in teaching and second, expected wages (five and 10 years after training) if they had continued to work in their original occupation.

Whether these wage expectations are indeed accurate *ex post* is irrelevant to what is at stake here, i.e., the selection to the teaching occupation. What matters are the expectations of individuals who decided to enter teaching at the time they made those decisions (*ex ante*). In accordance with search and matching models, one expects the average teacher to expect a monetary benefit from the change. A conscious decision to accept a monetary disadvantage from the decision to enter teaching likely occurs only in cases in which the relative non-monetary benefits of teaching are high enough to more than compensate for the monetary disadvantages.

5.3 Results

The results show that teachers represent a positive selection on average, i.e., they earned significantly more in their former occupation than comparable individuals who did not change career to teaching (see Table 5-2). Teachers earned, on average, 5,497 Swiss Francs (CHF)[8] more per annum in their former occupation than comparable colleagues (this is some 6% more than the average salary in the control group). The wage advantage over matched individuals is also positive for formerly self-employed individuals (approximately 18% of our sample); however, this estimate is not very precise. This result is no surprise given the high earnings heterogeneity among the self-employed population. The result shows that the decision-making scenario for the formerly self-employed is much more difficult to model than that of former salaried employees. To take these differences into account, we conduct the following analyses separately for each of the two initial employment situations.

Table 5-2: Average Wage Difference (Teacher – Non-Teacher)

	Average teacher wage (in CHF)	Average wage difference (in CHF)	Std. Err.	Observations
Employed	90,836	5,497	1,329***	330
Self-employed	90,218	7,165	6,274	60

Note: Exact Matching for the variables gender, age, degree of highest education and occupational field. Average difference in yearly gross wage in the pre-teaching occupation. * p< 0.1, ** p< 0.05, *** p < 0.01

In a further analysis, we regressed the individual wage differential (the teacher's wage in the non-teaching occupation minus the average wage of a comparable colleague in the same occupation) against the various characteristics of the teachers. This analysis shows us which individuals earned more or less in their former occupation than comparison subjects. The results in model 3 (Table 5-3) show that individuals in senior positions and specific industries earned significantly more than comparison subjects, whereas others earned significantly less. Other characteristics, such as gender or qualifications, have no significant impact on wage differences for an average teacher.

In keeping with the hypotheses outlined earlier, the large effect of sizes for the occupational categories show a consistent picture. The higher the average wage level in an occupation, the more likely that teachers will constitute a negative selection, i.e., those who tend to earn less than comparison subjects, and vice versa. The average wage is about CHF 98,805 in the "engineering and informatics" category

[8] At the time of the study one Swiss Franc (CHF) was roughly equivalent to 1.45 Euro (EUR).

and CHF 89,022 in “management, administration, banking, insurance and judiciary”, and the coefficient is negative in both cases.

In contrast, the average wage level in the reference category “industry and manufacturing” is CHF 69,922, and all of the occupation categories that deviate positively from the reference category feature average wages in the under CHF 80,000 range. This finding demonstrates that average salaries in the former occupation are the main factor determining whether teachers are more likely to constitute a positive or negative selection from their occupational sector in terms of earnings.

Table 5-3: Wage Difference "Teacher - Non-Teachers" (Regression)

	(1)	(2)	(3)	(4)
Female	-7,273	-6,611	-4,029	-2,180
	(2,653)***	(2,741)**	(3,149)	(3,252)
Reference category: Tertiary level B professional education and training (ISCED 5b)				
Upper secondary		-725	4,269	1,742
education		(4,335)	(4,155)	(5,747)
ISCED 5a+6		-5,821	1,195	6,778
		(3,065)*	(3,323)	(3,998)*
Leadership position			7,170	6,019
			(2,731)***	(2,711)**
Reference category: Industry and manufacturing				
Agriculture, forestry and			1,935	3,939
livestock breeding			(11,434)	(12,415)
Engineering and informatics			-19,568	-19,077
			(4,002)***	(3,952)***
Construction sector and			6,582	7,815
mining			(5,265)	(5,343)
Trade and transport			-17,835	-15,231
			(6,849)***	(6,853)**
Hospitality industry and			1,989	2,648
service sector			(4,330)	(4,219)
Management, admin., banking,			-21,904	-22,541
insurance and law			(5,566)***	(5,779)***
Health services, culture,			-11,502	-10,050
science			(4,287)***	(4,331)**
Interaction: upper sec. educ.				1,186
x option part-time				(7,301)
Interaction: ISCED 5a+6				-14,224
x option part-time				(4,430)***
Interaction: ISCED 5b				-651
x option part-time				(3,582)
Constant	9,682	11,572	3,091	3,555
	(2,428)***	(2,414)***	(5,760)	(5,735)
Controlled for age	Yes	Yes	Yes	Yes
Controlled for region	No	No	Yes	Yes
Controlled for firm size	No	No	Yes	Yes
Controlled for tenure	No	No	Yes	Yes
Controlled for SLFS year	No	No	Yes	Yes
R-squared	0.03	0.04	0.25	0.28
Observations	330	330	330	330
F	2.84	2.96	5.18	5.01

Note: Dependent variable wage difference (teacher – non-teaching individuals). Robust standard errors in parentheses. Self-employed individuals excluded. * p<0.10, ** p<0.05, *** p<0.01

Model 4 shows that individuals with an upper secondary qualification and advanced occupation-specific qualification (ISCED 5B) constitute an average selection, whereas, for individuals with an academic qualification, positive selection applies only to those who did not previously have the option of working part-time. In contrast, subjects with an academic qualification who also had the option of working part-time are among those who earned significantly less than the comparison group in their former occupation.

The position in the wage distribution in the initial occupation of teachers does not tell us whether the career change into teaching pays off. To analyze this, we calculate the differences between wage expectations for teaching and original occupation by eliciting the respective expectations in the teacher survey.[9] The results show that the average teacher expects an annual wage benefit in teaching compared with the former occupation. This difference is significantly different from zero. However, individual results with respect to relative wage expectations can be both positive and negative. Depending on the scenario, between one-quarter and one-third of teachers expect to earn less than in their original occupation (see Table 5-4).

Table 5-4: Expected Wage Difference

	Observations	Mean	Std. Err	Median	Share neg. diff.
all teachers					
In 5 years	339	3,152	1,309**	3’000	0.32
In 10 years	339	6,101	1,620	5,000	0.27
without self-employed					
In 5 years	289	3,735	1,378***	5,000	0.31
In 10 years	289	7,396	1,685***	9,000	0.25

** p<0.05, *** p<0.01

A regression of this expected wage difference on teacher characteristics reveals that, for all scenarios, teachers with an academic tertiary A education expect to gain significantly less from a change to teaching than teachers with another educational background. The same applies for teachers from occupations in management, administration, banking, insurance and law, one of the industries with the highest wage level. Furthermore, teachers who hold a senior position in their former occupation also expect a significantly lower wage difference. Finally, tenure is negatively correlated with the expected wage difference.

[9] Descriptive statistics on wage expectation can be found in the Appendix, Table 5-6.

5.4 Conclusion

This study investigates the determinants of career change for individuals who change jobs to become a teacher as a second career. This chapter focuses mainly on the relevance of monetary factors in making people more or less likely to decide in favor of changing careers to teaching, as monetary factors are one of the most discussed levers for Swiss educational policy makers to influence the equilibrium of supply and demand in the labor market for teachers. However, an analysis of monetary factors allows us also to draw conclusions with respect to the relative importance of non-monetary factors (e.g. time off, workloads, fringe benefits, etc.) in informing the individual decision to change careers to teaching.

The framework for this study is the Swiss vocational education system, which requires that teachers of vocational subjects have a prior career in that specific field and, therefore, vocational teachers are all career changers.

The finding that the average teacher earned significantly more in his or her former occupation than comparison subjects supports the appeal for teaching. This result indicates that the average career changer does not change careers because he or she is (financially) unsuccessful or unproductive in their original occupation. Because a positive correlation between productivity in the original occupation and aptitude for teaching in vocational teaching is likely, this result has positive implications for the quality of vocational schools.

As to recruitment chances of vocational schools in the individual occupations: the higher the average wage level in an occupation, the larger the probability that individuals recruited from that occupation will rank among the low earners. Teachers need to be recruited from sectors of the rest of the economy with extremely different wage levels, but there is no major wage differential in the educational system. Therefore, as a function of wage level in the economy, equally “talented” teachers will not be available for all of the occupations taught in the VET system.

The results for all of the analyses display significant heterogeneity, some of which can be explained. Positive selection for individuals with a university degree applies only to those individuals (largely male) who did not have the option of working part-time in their former occupation, whereas the other teachers with a university degree constitute a negative selection in terms of their relative earnings in their former occupations. This analysis shows the great relevance of non-monetary factors (e.g., flexibility to arrange individual working time) in forming the decision to enter teaching. Therefore, the recently started discussion in Switzerland about highlighting non-monetary benefits in the public perception in order to attract scarce candidates can be seen as a fine step forward to improve the appeal of teaching jobs.

Although the average teacher tends to rank among the better earners in his or her original occupation, the majority of career changers expect to earn more as a teacher than in their original occupation. This finding shows that the average wage level at vocational schools can compete with average wage levels in the rest of the economy. Again, however, substantial heterogeneity exists given that between one-quarter and one-third of career changers are prepared to accept a cut in wage after changing to teaching. One probable explanation is the very high relevance of non-monetary factors that make teaching a more attractive option, at least for some individuals.

5.5 Appendix

Table 5-5: Description of Variables

Variable	Description	Matching variable
Female (dummy)	Gender, 1 for females, 0 for males	Yes
Age	Age	Yes (age categories)
Gross yearly wage in CHF	Wage in non-teaching occupation	No
Education		
Upper secondary education	Vocational training or high school	Yes
ISCED 5b	Federal PET Diploma Examination, the Advanced Federal PET Diploma (also referred to as Meisterprüfung) and PET Colleges	
ISCED 5a+6	Bachelor or Master degree at a university or university of applied sciences	
Occupational field		Yes
Agriculture, forestry and livestock breeding		
Industry and manufacturing		
Engineering and informatics		
Construction sector and mining		
Trade and transport		
Hospitality industry and service sector		
Management, administration, banking, insurance and law		
Health services, culture, science		
Children (yes/no, dummy)	1 for one or more children, 0 for no children	No
Leadership position (yes/no, dummy)	Responsible for one or more employees in the non-teaching position	No
Tenure	Tenure	No
Self-employed (yes/no, dummy)	Self-employed in the non-teaching position	No
Sideline teacher (yes/no, dummy)	Teaching as a sideline	No

Table 5-6: Descriptive Statistics "Wage Expectations"

Expected yearly gross wage (CHF)		*all teachers*		*without self-employed*	
		mean	SD	mean	SD
in teaching	in 5 years	106,789	20,948	106,998	21,382
	in 10 years	117,549	23,121	117,876	23,734
outside education	in 5 years	103,637	29,996	103,264	29,727
	in 10 years	111,448	349657	110,480	34,159
Observations		339			289

Bibliography

Abadie, A., Herr, J. L., Imbens, G. W., & Drukker, D. M. (2004). *NNMATCH: Stata Module to Compute Nearest-Neighbor Bias-Corrected Estimators*. Boston College Department of Economics.

Almada, L., McCarthy, I. M., & Tchernis, R. (2015). *What Can We Learn About the Effects of Food Stamps on Obesity in the Presence of Misreporting?* (SSRN Scholarly Paper No. ID 2563822). Rochester, NY: Social Science Research Network.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. (2011). Personality Psychology and Economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (pp. 1–181). Amsterdam: North Holland.

Altonji, J. G., & Paxson, C. H. (1986). Job Characteristics and Hours of Work. In R. G. Ehrenberg (Ed.), *Research in Labor Economics* (pp. 1–55). Greenwich: Westview Press.

Amirkhan, J. H., Risinger, R. T., & Swickert, R. J. (1995). Extraversion: A "Hidden" Personality Factor in Coping? *Journal of Personality*, *63*, 189–212.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, *91*, 444–455.

Arulampalam, W., Naylor, R. A., & Smith, J. (2012). Am I Missing Something? The Effects of Absence From Class on Student Performance. *Economics of Education Review*, *31*, 363–375.

Atrostic, B. K. (1982). The Demand for Leisure and Nonpecuniary Job Characteristics. *American Economic Review*, *72*, 428–40.

Baker, D. P., & LeTendre, G. K. (2005). *National Differences, Global Similarities: World Culture and the Future of Schooling*. California: Stanford University Press.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, *122*, 1235 –1264.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*, 9–49.

Behncke, S. (2012). How Do Shocks to Non-Cognitive Skills Affect Test Scores? *Annals of Economics and Statistics*, *108*, 155–173.

Bellei, C. (2009). Does Lengthening the School Day Increase Students' Academic Achievement? Results From a Natural Experiment in Chile. *Economics of Education Review*, *28*, 629–640.

Beresteanu, A., & Manski, C. F. (2000). Bounds for STATA: Draft Version 1.0. Northwestern University.

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a Multiyear Universal Social–Emotional Learning Program: The Role of Student and School Characteristics. *Journal of Consulting and Clinical Psychology*, *78*, 156–168.

Björklund, A., & Salvanes, K. G. (2011). Education and Familiy Background: Mechanisms and Policies. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (pp. 201–247). Amsterdam: Waxmann.

Black, S. E., & Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (pp. 1487–1541). Amsterdam: Elsevier.

Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2006). Continuity and Change in Psychopathic Traits as Measured Via Normal-Range Personality: A Longitudinal–Biometric Study. *Journal of Abnormal Psychology*, *115*, 85–95.

Blundell, R., Gosling, A., Ichimura, H., & Meghir, C. (2007). Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds. *Econometrica*, *75*, 323–363.

Boes, S. (2013). Nonparametric Analysis of Treatment Effects in Ordered Response Models. *Empirical Economics*, *44*, 81–109.

Bolli, T., & Hof, S. (2014). *The Impact of Apprenticeship Training on Personality Traits: An Instrumental Variable Approach* (KOF Working paper No. 14-350). KOF Swiss Economic Institute, ETH Zurich.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental Trajectories of Externalizing Behaviors in Childhood and Adolescence. *Child Development*, *75*, 1523–1537.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Weel, B. ter. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, *43, 972-1059*.

Bosworth, D. (1994). Truancy and Pupil Performance. *Education Economics*, *2*, 243–264.

Boyce, C. J., Wood, A. M., & Powdthavee, N. (2013). Is Personality Fixed? Personality Changes as Much as "Variable" Economic Factors and More Strongly Predicts Changes to Life Satisfaction. *Social Indicators Research*, *111*, 287–305.

Bray, M. (2007). *The Shadow Education System: Private Tutoring and Its Implications for Planners (second edition)*. Paris: UNESCO.

Bray, M. (2011). *The Challenge of Shadow Education*. Brussels: European Commission.

Brousseau, K. R., & Bruce, J. (1981). Job-Person Dynamics: An Extension of Longitudinal Research. *Journal of Applied Psychology*, *66*, 59–62.

Brunello, G., & Schlotter, M. (2011). *Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in E&T Systems* (EENEE Analytical Report No. 8). EENEE.

Burdett, K. (1978). A Theory of Employee Job Search and Quit Rates. *The American Economic Review*, *68*, 212–220.

Büttner, B., Thiel, H., & Thomsen, S. L. (2011). *Variation of Learning Intensity in Late Adolescence and the Impact on Noncognitive Skills* (ZEW Discussion Paper No. 11-007). ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Chevalier, A., Dolton, P., & Mcintosh, S. (2007). Recruiting and Retaining Teachers in the UK: An Analysis of Graduate Occupation Choice from the 1960s to the 1990s. *Economica*, *74*, 69–96.

Chingos, M. M., & West, M. R. (2012). Do More Effective Teachers Earn More Outside the Classroom? *Education Finance and Policy*, *7*, 8–43.

Clark, A. E. (2001). What Really Matters in a Job? Hedonic Measurement Using Quit Data. *Labour Economics*, *8*, 223–242.

Cobb-Clark, D. A., & Schurer, S. (2012). The Stability of Big-Five Personality Traits. *Economics Letters*, *115*, 11–15.

Considine, G., & Zappalà, G. (2002). The Influence of Social and Economic Disadvantage in the Academic Performance of School Students in Australia. *Journal of Sociology*, *38*, 129–148.

Corman, H. (1993). Who Will Teach? Policies That Matter : Richard J. Murnane, Judith D. Singer, John B. Willett, James J. Kemple and Randall J. Olsen. Cambridge, MA: Harvard University Press, 1991. XII + 187 p. *Economics of Education Review*, *12*, 273–274.

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*, 367–384.

Cortes, K. E., & Goodman, J. S. (2014). Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement. *American Economic Review*, *104*, 400–405.

Corville-Smith, J., Ryan, B. A., & Adams, G. R. (1998). Distinguishing Absentee Students from Regular Attenders: The Combined Influence of Personal, Family, and School Factors. *Journal of Youth and Adolescence*, *27*, 629–640.

Costa, P. T., & MacCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual*. Odessa, Fl.: Psychological Assessment Resources.

Cunha, F., Heckman, J. J., & Lochner, L. (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. Hanushek & F. Welch (Eds.), *Handbook of Economics of Education* (1st ed., Vol. 1, pp. 697–812). Amsterdam: North Holland.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*, 883–931.

Dahl, R. E. (2004). Adolescent Brain Development: A Period of Vulnerabilities and Ppportunities. Keynote Address. *Annals of the New York Academy of Sciences*, *1021*, 1–22.

Dahmann, S., & Anger, S. (2014). *The Impact of Education on Personality: Evidence from a German High School Reform* (SSRN Scholarly Paper No. ID 2432423). Rochester, NY: Social Science Research Network.

Dang, H.-A., & Rogers, F. H. (2008). The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources? *The World Bank Research Observer*, *23*, 161 –200.

Daniel, K., & Heywood, J. (2007). The Determinants of Hiring Older Workers: UK Evidence. *Labour Economics*, *14*, 35–51.

Deci, E., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior* (Auflage: 1985). New York: Springer.

Dee, T. S., & West, M. R. (2011). The Non-Cognitive Returns to Class Size. *Educational Evaluation and Policy Analysis*, *33*, 23–46.

De Haan, M. (2011). The Effect of Parents' Schooling on Child's Schooling: A Nonparametric Bounds Analysis. *Journal of Labor Economics*, *29*, 859–892.

De Haan, M. (2015). The Effect of Additional Funds for Low-Ability Pupils: A Non-Parametric Bounds Analysis. *The Economic Journal*, Online Version.

De Los Reyes, A., & Kazdin, A. E. (2004). Measuring Informant Discrepancies in Clinical Child Research. *Psychological Assessment*, *16*, 330–334.

Denzler, S., & Wolter, S. C. (2009). Sorting into Teacher Education: How the Institutional Setting Matters. *Cambridge Journal of Education*, *39*, 423–441.

De Paola, M., & Scoppa, V. (2014). The Effectiveness of Remedial Courses in Italy: A Fuzzy Regression Discontinuity Design. *Journal of Population Economics*, *27*, 365–386.

De Wals, S., & Meszaros, K. (2011). *Handbook on Psychology of Self-Esteem*. Hauppauge, NY: Nova Science Publishers.

Dobkin, C., Gil, R., & Marion, J. (2010). Skipping Class in College and Exam Performance: Evidence From a Regression Discontinuity Classroom Experiment. *Economics of Education Review*, *29*, 566–575.

Doeringer, P., & Piore, M. (1971). Internal Labor Markets and Manpower Analysis. *Lexington, Mass: Heat*.

Dolton, P. J. (1990). The Economics of UK Teacher Supply: The Graduate's Decision. *Economic Journal*, *100*, 91–104.

Dolton, P. J., & Chung, T.-P. (2004). The Rate of Return to Teaching: How does it Compare to other Graduate Jobs? *National Institute Economic Review*, *190*, 89–103.

Duarte, R., Escario, J.-J., & Molina, J. A. (2011). Peer Effects, Unobserved Factors and Risk Behaviours: An Analysis of Alcohol Abuse and Truancy among Adolescents. *Revista de Economia Aplicada*, *19*, 125–151.

Dunkake, I. (2006). Truants and the Family: An Empirical Study of Deviant Behavior in Early Adolescence. *Sociology of Crime, Law & Deviance*, *7*, 29–56.

Ehmke, D. T., & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft*, *8*, 521–539.

Eisner, M., & Ribeaud, D. (2007). Conducting a Criminological Survey in a Culturally Diverse Context: Lessons from the Zurich Project on the Social Development of Children. *European Journal of Criminology*, *4*, 271–298.

Endler, N. S., & Parker, J. D. (1990). Multidimensional Assessment of Coping: A Critical Evaluation. *Journal of Personality and Social Psychology*, *58*, 844–854.

Falch, T. (2010). The Elasticity of Labor Supply at the Establishment Level. *Journal of Labor Economics*, *28*, 237–266.

Falch, T., Nyhus, O. H., & Strøm, B. (2014). Performance of Young Adults: The Importance of Different Skills. *CESifo Economic Studies*, ifu005.

Falck, O., & Woessmann, L. (2013). School Competition and Students' Entrepreneurial Intentions: International Evidence Using Historical Catholic Roots of Private Schooling. *Small Business Economics*, *40*, 459–478.

Fletcher, J. (2013). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, *89*, 122–135.

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-Term Effects of Class Size. *The Quarterly Journal of Economics*, *128*, 249–285.

Geel, R., Mure, J., & Backes-Gellner, U. (2011). Specificity of Occupational Training and Occupational Mobility: An Empirical Study Based on Lazear's Skill-Weights Approach. *Education Economics*, *19*, 519–535.

Gensowski, M. (2014). *Personality, IQ, and Lifetime Earnings* (IZA Discussion Paper No. 8235), Institute for the Study of Labor (IZA).

Gerfin, M., & Schellhorn, M. (2006). Nonparametric Bounds on the Effect of Deductibles In Health Care Insurance on Doctor Visits - Swiss Evidence. *Health Economics*, *15*, 1011–1020.

Goodman, J. (2014). *Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time* (NBER Working Paper No. 20221). National Bureau of Economic Research, Inc.

Gundersen, C., & Kreider, B. (2008). Food Stamps and Food Insecurity: What Can Be Learned in the Presence of Nonclassical Measurement Error? *Journal of Human Resources*, *43*, 352–382.

Gundersen, C., Kreider, B., & Pepper, J. (2012). The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis. *Journal of Econometrics*, *166*, 79–91.

Hanushek, E. A., & Pace, R. R. (1995). Who Chooses to Teach and Why? *Economics of Education Review*, *14*, 101–117.

Hanushek, E. A., Welch, F., Machin, S., & Woessmann, L. (2011). *Handbook of the Economics of Education*. Amsterdam: North Holland.

Heckman, J. J., Humphries, J. E., Veramendi, G., & Urzua, S. (2014). *Education, Health and Wages* (IZA Discussion Paper No. 8027). Institute for the Study of Labor (IZA).

Heckman, J. J., & Kautz, T. (2012). Hard Evidence on Soft Skills. *Labour Economics*, *19*, 451–464.

Heckman, J. J., & Kautz, T. (2013). *Fostering and Measuring Skills: Interventions That Improve Character and Cognition* (NBER Working Paper No. 19656). National Bureau of Economic Research.

Heckman, J. J., & Leamer, E. E. (2007). *Handbook of Econometrics*. Amsterdam: North Holland.

Heckman, J. J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, *103*, 2052–86.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*, 411–482.

Heckman, J. J., & Vytlacil, E. J. (2000). Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect. In M. Lechner & F. Pfeiffer (Eds.), *Economic Evaluation of Labour Market Policies*. Hannheim, Germany: Physica-Verlag HD.

Heine, J. H., Sälzer, C., Bochert, L., Sibberns, H., & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Eds.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (pp. 309–346). Münster: Waxmann.

Henry, K. L. (2007). Who's Skipping School: Characteristics of Truants in 8th and 10th Grade. *Journal of School Health*, *77*, 29–35.

Heywood, J. S., Jirjahn, U., & Tsertsvardze, G. (2010). Hiring Older Workers and Employing Older Workers: German Evidence. *Journal of Population Economics*, *23*, 595–615.

Hof, S. (2014). Does Private Tutoring Work? The Effectiveness of Private Tutoring: A Nonparametric Bounds Analysis. *Education Economics*, *22*, 347–366.

Hof, S., & Strupler Leiser, M. (2014). Teaching in Vocational Education As a Second Career. *Empirical Research in Vocational Education and Training*, *6*, 1–13.

Hof, S., & Wolter, S. (2012). Nachhilfe Bezahlte ausserschulische Lernunterstützung in der Schweiz. *SKBF Staff Paper*, *8*.

Hudson, N. W., & Brent, R. W. (2012). Personality Trait Development and Social Investment in Work. *Journal of Research in Personality*, *46*, 334–344.

Jackson, C. K. (2012). *Non-Cognitive Ability, Test Scores, and Teacher Quality: Evidence from 9th Grade Teachers in North Carolina* (NBER Working Paper No. 18624). National Bureau of Economic Research, Inc.

Jackson, J. J. (2011). *The Effects of Educational Experiences on Personality Trait Development*. University of Illinois at Urbana-Champaign, Illinois.

Jacob, B. A., & Lefgren, L. (2004). Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *Review of Economics and Statistics*, *86*, 226–244.

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *Journal of Political Economy*, *87*, 972–990.

Jung, J. H., & Lee, K. H. (2010). The Determinants of Private Tutoring Participation and Attendant Expenditures in Korea. *Asia Pacific Education Review*, *11*, 159–168.

Kang, C. (2011). Family Size and Educational Investments in Children: Evidence from Private Tutoring Expenditures in South Korea. *Oxford Bulletin of Economics and Statistics*, *73*, 59–78.

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five Personality Traits in Predicting College Students' Academic Motivation and Achievement. *Learning and Individual Differences*, *19*, 47–52.

Kreider, B., & Pepper, J. V. (2007). Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors. *Journal of the American Statistical Association*, *102*, 432–441.

Kreider, B., Pepper, J. V., Gundersen, C., & Jolliffe, D. (2012). Identifying the Effects of SNAP (Food Stamps) on Child Health Outcomes When Participation Is Endogenous and Misreported. *Journal of the American Statistical Association*, *107*, 958–975.

Kuntsche, E., & Labhart, F. (2012). Investigating the Drinking Patterns of Young People Over the Course of the Evening at Weekends. *Drug and Alcohol Dependence*, 319–324.

Kunze, L., & Suppa, N. (2013). *Job Characteristics and Labour Supply* (SSRN Scholarly Paper No. ID 2278872). Rochester, NY.

Lavy, V. (2015). Do Differences in School's Instruction Time Explain International Achievement Gaps in Maths, Science and Language? Evidence from Developed and Developing Countries. *The Economic Journal*, *125*, F397–F424.

Lavy, V., & Schlosser, A. (2005). Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits. *Journal of Labor Economics*, *23*, 839–874.

Lazear, E. P. (1979). Why Is There Mandatory Retirement? *Journal of Political Economy*, *87*, 1261–84.

Lazear, E. P. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions. *The American Economic Review*, *71*, 606–620.

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, *117*, 914–940.

Leandro, P. G., & Castillo, M. D. (2010). Coping with Stress and Its Relationship with Personality Dimensions, Anxiety, and Depression. *Procedia - Social and Behavioral Sciences*, *5*, 1562–1573.

Lee, J. Y. (2013). *Private Tutoring and Its Impact on Students' Academic Achievement, Formal Schooling, and Educational Inequality in Korea* (PhD thesis). Columbia University, New York.

Lehnart, J., Neyer, F. J., & Eccles, J. (2010). Long-Term Effects of Social Investment: The Case of Partnering in Young Adulthood. *Journal of Personality*, *78*, 639–670.

Lerman, R. I. (2013). Are Employability Skills Learned in U.S. Youth Education and Training Programs? *IZA Journal of Labor Policy*, *2*, 6.

Lindqvist, E., & Vestman, R. (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, *3*, 101–28.

Lodi-Smith, J., & Roberts, B. W. (2007). Social Investment and Personality: A Meta-Analysis of the Relationship of Personality Traits to Investment in Work, Family, Religion, and Volunteerism. *Personality and Social Psychology Review*, *11*, 68–86.

Lüdemann, E. (2011). *Schooling and the Formation of Cognitive and Non-cognitive Outcomes* (ifo Beiträge zur Wirtschaftsforschung). Ifo Institute for Economic Research at the University of Munich.

Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A Random Walk Down University Avenue: Life Paths, Life Events, and Personality Trait Change at the Transition to University Life. *Journal of Personality and Social Psychology*, *101*, 620–637.

Luechinger, S., Meier, S., & Stutzer, A. (2010). Why Does Unemployment Hurt the Employed?: Evidence from the Life Satisfaction Gap Between the Public and the Private Sector. *Journal of Human Resources*, *45*, 998–1045.

Lykken, D. T., Bouchard, T. J., McGue, M., & Tellegen, A. (1993). Heritability of Interests: A Twin Study. *Journal of Applied Psychology*, *78*, 649–661.

Malcolm, H., Thorpe, G., & Lowden, K. (1996). *Understanding Truancy: Links between Attendance, Truancy and Performance*. Scottish Council for Research in Education.

Mallinson, C. (2007). Social Class, Social Status and Stratification: Revisiting Familiar Concepts in Sociolinguistics. *University of Pennsylvania Working Papers in Linguistics*, *13*.

Manski, C. F. (1987). Academic Ability, Earnings, and the Decision to Become a Teacher: Evidence From the National Longitudinal Study of the High School Class of 1972. In *Public Sector Payrolls* (pp. 219–316). National Bureau of Economic Research.

Manski, C. F. (1990). Nonparametric Bounds on Treatment Effects. *American Economic Review*, *80*, 319–23.

Manski, C. F. (1997). Monotone Treatment Response. *Econometrica*, *65*, 1311–1334.

Manski, C. F. (2007). *Identification for Prediction and Decision*. Cambridge, MA: Harvard University Press.

Manski, C. F., & Pepper, J. V. (2000). Monotone Instrumental Variables, with an Application to the Returns to Schooling. *Econometrica*, *68*, 997–1012.

Manski, C. F., & Pepper, J. V. (2009). More on Monotone Instrumental Variables. *The Econometrics Journal*, *12*, S200–S216.

Manski, C. F., & Pepper, J. V. (2011). *Deterrence and the Death Penalty: Partial Identification Analysis Using Repeated Cross Sections* (NBER Working Paper No. 17755). National Bureau of Economic Research, Inc.

Mariotta, M. (2006). Il ricorso a lezioni private nella scuola media. *Scuola ticinese*, 14 – 16.

Mariotta, M., & Nicoli, M. (2005). Il ricorso a lezioni private nella scuola media. Repubblica e Cantone del Ticino.

McCarthy, I., Millimet, D. L., & Roy, M. (2015). Bounding Treatment Effects: Stata Command for the Partial Identification of the Average Treatment Effect with Endogenous and Misreported Treatment Assignment. *Stata Journal*, *15*, 411–436.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.

McDonnell, L. M. (1995). Opportunity to Learn as a Research Concept and a Policy Instrument. *Educational Evaluation and Policy Analysis*, *17*, 305–322.

Meghir, C., Palme, M., & Simeonova, E. (2013). *Education, Cognition and Health: Evidence from a Social Experiment* (NBER Working Paper No. 19002). National Bureau of Economic Research, Inc.

Millimet, D. L., & Jayjit, R. (2015). Multilateral Environmental Agreements and the WTO. *Economics Letters*, *134*, 20–23.

Monk, R. L., Heim, D., Qureshi, A., & Price, A. (2015). “I Have No Clue What I Drunk Last Night” Using Smartphone Technology to Compare In-Vivo and Retrospective Self-Reports of Alcohol Consumption. *PLoS ONE*, *10*.

Mortensen, D. T. (1987). *Job Search and Labor Market Analysis*. In O. Ashenfelter & R. Layard (Eds), Handbook of Labor Economics (pp. 849–919). Amsterdam: Elsevier.

Neal, D. (1999). The Complexity of Job Mobility among Young Men. *Journal of Labor Economics*, *17*, 237–61.

Oberwittler, D., & Naplava, T. (2002). Auswirkungen des Erhebungsverfahrens bei Jugendbefragungen zu ‚heiklen’ Themen - schulbasierte schriftliche Befragung und haushaltsbasierte mündliche Befragung im Vergleich. *ZUMA Nachrichten*, *26*, 49–77.

OECD. (2009). *PISA Data Analysis Manual: SPSS* (Second Edition). Paris: OECD Publishing.

OECD. (2011a). *PISA 2009 Technical Report (Preliminary version)*. Paris: OECD Publishing.

OECD. (2011b). *Quality Time for Students: Learning In and Out of School*. Paris: OECD Publishing.

OECD. (2013a). *Apprenticeships and Workplace Learning* (No. 3). Paris: OECD.

OECD. (2013b). *PISA 2012 Assessment and Analytical Framework Mathematics, Reading, Science, Problem Solving and Financial Literacy: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing.

OECD. (2014). *PISA 2012 Technical Report*. Paris: OECD Publishing.

Ono, H. (2007). Does Examination Hell Pay Off? A Cost-Benefit Analysis of “ronin” and College Education in Japan. *Economics of Education Review*, *26*, 271–284.

Paulhus, D. L. (1984). Two-Component Models of Socially Desirable Responding. *Journal of Personality and Social Psychology*, *46*, 598–609.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and Denial in Socially Desirable Responding. *Journal of Personality and Social Psychology*, *60*, 307–317.

Pinkovskiy, M. L. (2013). World Welfare is Rising: Estimation Using Nonparametric Bounds on Welfare Measures. *Journal of Public Economics*, *97*, 176–195.

Podgursky, M., Monroe, R., & Watson, D. (2004). The Academic Quality of Public School Teachers: An Analysis of Entry and Exit Behavior. *Economics of Education Review*, *23*, 507–518.

Prenzel, M. (Ed.). (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

Reeve, J. (2002). Self-Determination Theory Applied to Educational Settings. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 183–203). Rochester, NY, US: University of Rochester Press.

Reid, K. (1999). *Truancy and Schools*. London ; New York: Routledge.

Renzulli, J. S., & Park, S. (2000). Gifted Dropouts: The Who and the Why. *Gifted Child Quarterly*, *44*, 261–271.

Reyes, A., & Kazdin, A. E. (2006). Informant Discrepancies in Assessing Child Dysfunction Relate to Dysfunction Within Mother-Child Interactions. *Journal of Child and Family Studies*, *15*, 643–661.

Rivkin, S. G., & Schiman, J. C. (2015). Instruction Time, Classroom Quality, and Academic Achievement. *The Economic Journal*, *125*, *F425-F448*.

Roberts, B. W. (1997). Plaster or Plasticity: Are Adult Work Experiences Associated with Personality Change in Women? *Journal of Personality*, *65*, 205–232.

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work Experiences and Personality Development in Young Adulthood. [Miscellaneous Article]. *Journal of Personality*, *84*, 582–593.

Roberts, B. W., & Chapman, C. N. (2000). Change in Dispositional Well-Being and Its Relation to Role Quality: A 30-Year Longitudinal Study. *Journal of Research in Personality*, *34*, 26–41.

Roberts, B. W., & Mroczek, D. (2008). Personality Trait Change in Adulthood. *Current Directions in Psychological Science*, *17*, 31–35.

Roberts, B. W., Robins, R. W., Caspi, A., & Trzesniewski, K. (2003). Personality Trait Development in Adulthood. In J. Mortimer & M. Shanahan (Eds.), *Handbook of Life Course*. New York: Kluwer Academic.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies. *Psychological Bulletin*, *132*, 1–25.

Roberts, B. W., & Wood, D. (2006). Personality Development in the Context of the Neo-Socioanalytic Model of Personality. In D. Mroczek & T. Little (Eds.), *Handbook of Personality Development*. Mahwah, NJ: Lawrance Erlbaum Associates.

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating Five Factor Theory and Social Investment Perspectives on Personality Trait Development. *Journal of Research in Personality*, *39*, 166–184.

Rothman, S. (2001). School Absence and Student Background Factors: A Multilevel Analysis. *International Education Journal*, *2*, 59–68.

Rothman, S. (2004). Staying Longer at School and Absenteeism: Evidence from Australian Research and the Longitudinal Surveys of Australian Youth. *International Education Journal*, *5*, 113–123.

Rotter, J. B. (1966). Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs: General and Applied*, *80*, 1–28.

Rubinstein, Y., & Weiss, Y. (2006). Post Schooling Wage Growth: Investment, Search and Learning. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (pp. 1–67). Amsterdam: North Holland.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*, 54–67.

Salusky, I., Larson, R. W., Griffith, A., Wu, J., Raffaelli, M., Sugimura, N., & Guzman, M. (2014). How Adolescents Develop Responsibility: What Can Be Learned from Youth Programs. *Journal of Research on Adolescence*, *24*, 417–430.

Sälzer, C., Trautwein, U., Lüdtke, O., & Stamm, M. (2012). Predicting Adolescent Truancy: The Importance of Distinguishing Between Different Aspects of Instructional Quality. *Learning and Instruction*, *22*, 311–319.

Shoenfelt, E. L., & Huddleston, M. R. (2006). An Evaluation of Impact on Attendance and Academic Performance. *Family Court Review*, *44*, 683–695.

Shute, J. W., & Cooper, B. S. (2015). Understanding In-School Truancy. *Phi Delta Kappan*, *96*, 65–68.

Siziya, S., Muula, A. S., & Rudatsikira, E. (2007). Prevalence and Correlates of Truancy Among Adolescents in Swaziland: Findings From the Global School-Based Health Survey. *Child and Adolescent Psychiatry and Mental Health*, *1*, 15.

SKBF. (2011). *Swiss Education Report* (English version). Aarau, Switzerland: SKBF.

Southgate, D. E. (2009). *Determinants of Shadow Education: A Cross-National Analysis* (PhD thesis). Ohio State University, Columbus.

Stinebrickner, T. R. (2001). A Dynamic Model of Teacher Labor Supply. *Journal of Labor Economics*, *19*, 196–230.

Suryadarma, D., Suryahadi, A., Sumarto, S., & Rogers, F. H. (2006). Improving Student Performance in Public Primary Schools in Developing Countries: Evidence from Indonesia. *Education Economics*, *14*, 401–429.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge, U.K. ; New York: Cambridge University Press.

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin*, *133*, 859–883.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, Grading and Student Motivation: Using Group Composition and Status to Predict Self-Concept and Interest in Ninth Grade Mathematics. *Journal of Educational Psychology*, *98*, 788–806.

Turner, J. H. (2013). Symbolic Interactionist Theories of Identity. In *Contemporary Sociologial Theory* (p. 768). SAGE Publications, Inc.

Vaughn, M. G., Maynard, B. R., Salas-Wright, C. P., Perron, B. E., & Abdon, A. (2013). Prevalence and correlates of truancy in the US: Results from a national sample. *Journal of Adolescence*, *36*, 767–776.

Veenstra, R., Lindenberg, S., Tinga, F., & Ormel, J. (2010). Truancy in Late Elementary and Early Secondary Education: The Influence of Social Bonds and Self-Control-- The TRAILS Study. *International Journal of Behavioral Development*, *34*, 302–310.

Wagner, P., Spiel, C., & Tranker, M. (2003). Wer nimmt Nachhilfe in Anspruch? *Zeitschrift Für Pädagogische Psychologie*, *17*, 233–243.

Warren, J. R., Sheridan, J. T., & Hauser, R. M. (1998). Choosing a Measure of Occupational Standing How Useful are Composite Measures in Analyses of Gender Inequality in Occupational Attainment? *Sociological Methods & Research*, *27*, 3–76.

Weissbrodt, T. (2007). Schulische Determinanten der Schulverweigerung - Befunde aus der PISA-Befragung. In Wagner (Ed.), *Juventa-Materialien. Schulabsentismus. Soziologische Anlaysen zum Einfluss von Familie, Schule und Freundeskreis* (pp. 85–104). Weinheim, Germany: Juventa.

Wells, R. (2010). An Examination of the Utility Bearing Characteristics of Occupations: A Factor Analytical Approach. *Economics Letters*, *108*, 296–298.

Wilson, V., Malcolm, H., Edward, S., & Davidson, J. (2008). “Bunking Off”: The Impact of Truancy on Pupils and Teachers. *British Educational Research Journal*, *34*, 1–17.

Zhang, Y. (2013). Does Private Tutoring Improve Students’ National College Entrance Exam Performance?—A Case Study from Jinan, China. *Economics of Education Review*, *32*, 1–28.

Selbstständigkeitserklärung

„Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Koautorenschaften sowie alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.“

Basel, den 14.03.2016 Stefanie Hof