FREIE UNIVERSITÄT BERLIN
FACHBEREICH WIRTSCHAFTSWISSENSCHAFTEN

Developing skills through non-formal learning activities: four essays in the economics of education

Inaugural-Dissertation zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaft des Fachbereiches Wirtschaftswissenschaften
der Freien Universität Berlin

vorgelegt von

Adrian Hille (M.Sc.)

2016

Erstgutachterin: Prof. Dr. C. Katharina Spieß

Zweitgutachterin: Prof. Dr. Silke Anger

Datum der mündlichen Prüfung: 19. Juli 2016

Eigenanteil und Vorveröffentlichungen

Kapitel 2: How learning a musical instrument affects the development of skills

- Koautor: Jürgen Schupp (DIW Berlin und Freie Universität Berlin)
- Veröffentlicht in *Economics of Education Review*, Volume 44, 56-82, 2015
- Vorveröffentlichungen: SOEPpaper 591, IZA Discussion Paper 7655, SFB 882 Working Paper 24

Kapitel 3: Mozart or Pelé? The effects of teenagers' participation in music and sports

- Koautoren: Charlotte Cabane und Michael Lechner (beide Universität Sankt Gallen)
- Veröffentlicht in *Labour Economics* (im Erscheinen)
- Vorveröffentlichungen: SOEPpaper 749, IZA Discussion Paper 8987, University of St. Gallen School of Economics and Political Science Working Paper 2015-09, CEPR Discussion Paper 10556

Kapitel 4: Improving access to non-formal education – evidence from the German "Educational package"

- kein Koautor
- Vorveröffentlichung: keine

Kapitel 5: How a universal music education program affects time use, behavior, and school attitude

- kein Koautor
- Im Begutachtungsprozess
- Vorveröffentlichung: SOEPpaper 810

Acknowledgements

First and foremost, I wish to thank C. Katharina Spieß, my supervisor, for her excellent and continuous support. She was always available for discussions and gave me thorough advice on all papers of this dissertation. It was an excellent idea to establish the monthly doctoral colloquium, which allowed us to discuss even early stage research ideas with her and our like-minded colleagues. It was also a pleasure to work with Katharina on some other project ideas and DIW Economic Bulletins.

I also wish to thank my second supervisor, Silke Anger, who accompanied in particular the first steps of my thesis, while she was still working at DIW Berlin. Moreover, I thank Jürgen Schupp, not only for guiding my research as coauthor for the first paper, but also for the opportunity to be part of the SOEP team. I thank Charlotte Cabane and Michael Lechner, with whom I wrote the second paper, as well as Annegret Arnold and Mila Staneva, coauthors of the two DIW Economic Bulletins that resulted as by-products of this thesis. I would also like to thank Carsten Schröder and Irwin Collier for accepting to be members of my dissertation committee.

I gratefully acknowledge financial support from the DFG through its Collaborative Research Center 882 “From Heterogeneities to Inequalities”. Being part of its A1 project “Social Closure and Hierarchization” benefitted my research a lot, because I learned to challenge the economist’s perspective once in a while.

Writing this thesis would not have been possible without the excellent working conditions provided by DIW Berlin. I am especially grateful to the DIW Graduate Center and its team, in particular Yun Cao, Helmut Lütkepohl, Juliane Metzner and Nils Saniter. I also thank all the helpful people who surrounded me at DIW, among others Herr Born for solving various computer problems, Philipp for his help with this LaTeX template, Debbie and Adam for English language editing in chapters 2 and 5, Anne for her help during her time as student assistant. And of course thank you to all SOEP colleagues, who were always willing to answer detailed question on variables or other data-related issues.

I am grateful to the *Verband deutscher Musikschulen*, as well as the *JeKi-Stiftung* for providing me with data on public music schools, as well as the list of schools participating in the JeKi program.

I also thank my colleagues and fellow doctoral students for contributing to a very amicable and friendly atmosphere. I wish to thank the GC2011 cohort and the SOEP team for being great friends and coworkers, as well as the doctoral students and postdocs attending Katharina's Promotionskolloquium for fruitful discussions. A special thank you goes to Michaela for her delicious cakes, which often kept me motivated throughout the day and which I already miss at my new job. And most importantly, I thank Alex, Elisabeth, Julia and Katharina for being incredibly nice office mates.

Finally, I thank Pauline and Juliette for putting everything in perspective!

Rechtliche Erklärungen

Erklärung gem. §4 Abs. 2 (Promotionsordnung)

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Adrian Hille
Berlin, 13. April 2016

Erklärung gem. §10 Abs. 3 (Promotionsordnung)

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe: Software LaTeXund Stata, Literatur siehe Literaturverzeichnis.

Adrian Hille
Berlin, 13. April 2016

Abstract

Chapter 2: Despite numerous studies on skill development, we know little about the effects of extracurricular music activities on cognitive and non-cognitive skills. This study examines how music training during childhood and youth affects the development of cognitive skills, school grades, personality, time use and ambition using data from the German Socio-Economic Panel (SOEP). Our findings suggest that adolescents with music training have better school grades, are more conscientious, open and ambitious. These effects are stronger among adolescents from lower socio-economic status. In order to address the non-random selection into playing music, we take into account detailed information on the child and its parents, which may determine both the decision to pursue music lessons and educational outcomes. While lacking truly exogenous variations in music activities, our results are robust to a large range of sensitivity tests. We thereby approach causality better than previous observational studies.

Chapter 3: Using data from the German Socio-Economic Panel, this paper analyses the effects of spending part of adolescents' leisure time playing music or doing sports, or both. We find that while playing music fosters educational outcomes more than doing sports, particularly so for girls and children from more highly educated families, doing sports improves subjective health. For educational outcomes, doing both activities appears to be most successful. The robustness of the results is examined with respect to the identifying assumptions, including non-affected outcomes, a formal sensitivity analysis, and instrumental variable estimation. These checks do not reveal any serious problems.

Chapter 4: Music, sports and other types of non-formal education are important leisure time activities among children in Germany and other developed countries. Previous research suggests that these activities foster the development of cognitive and non-cognitive skills. However, access to non-formal education strongly depends on socio-economic status. This paper examines whether in-kind transfers targeted at children from poor households,

as provided by the German "Educational package", increase participation in extracurricular music and sports activities. I use data from Families in Germany (FiD), a household panel study with a focus on child development. Causal effects are identified by combining a difference-in-differences estimator with propensity score matching, using eligibility for the policy as the treatment variable. I find that subsidies for music school and sports club membership fees, which are provided by the Educational package, have no overall effect on music and sports activities, but increase participation by 10 percentage points among children from two-parent families, as well as among eligible children with relatively high household incomes. The absence of overall effects is probably due to a low take-up rate. Moreover, one out of two beneficiaries was already active before the subsidy was established.

Chapter 5: It is still widely debated how non-cognitive skills can be affected by policy intervention. For example, universal music education programs are becoming increasingly popular among policy makers in Germany and other developed countries. These are intended to give children from poor families the opportunity to learn a musical instrument. Moreover, policymakers present these programs as innovative policies that are important for the personality development of young children. However, the effects of universal music education on such outcomes are not yet sufficiently studied. This paper analyses the *Jedem Kind ein Instrument* (an instrument for every child) program in the German state of North Rhine-Westphalia. To do so, data from the German household panel studies SOEP and FiD are combined with regional data on primary and music schools. Using a difference-in-differences estimator, I show that the program successfully increases music participation among disadvantaged children. It does so more effectively than the alternative policy of reducing fees at public music schools. I further find that participation reduces conduct problems and improves student teacher relationships, especially among boys.

Keywords: Non-formal education, cognitive and non-cognitive skills, inequality, policy evaluation, German Socio-Economic Panel Study

Zusammenfassung

Kapitel 2: Obwohl die Entwicklung von Fähigkeiten bereits Gegenstand zahlreicher Forschungsvorhaben war, ist bisher wenig über die Wirkung von Musik auf kognitive und nicht-kognitive Fähigkeiten bekannt. Die vorliegende Studie untersucht auf der Basis von Daten des Sozio-ökonomischen Panels, wie sich die Ausbildung auf einem Musikinstrument in Kindheit und Jugend auf die Entwicklung von Intelligenz, Schulnoten, Persönlichkeit und Ambition auswirkt. Die Ergebnisse deuten darauf hin, dass Jugendliche, die an außerschulischem Musikunterricht teilnehmen, nicht nur bessere Schulnoten haben, sondern auch gewissenhafter, offener und ambitionierter sind. Am Stärksten sind die Effekte bei Jugendlichen mit niedrigerem sozio-ökonomischen Status. Die systematische Selektion in Musikunterricht wird berücksichtigt, indem für detaillierte Charakteristika der Jugendlichen und ihrer Eltern statistisch kontrolliert wird, die sowohl die Teilnahme am Musikunterricht als auch die Outcomes beeinflussen können. Obwohl den Analysen keine statistisch einwandfreie exogene Variation in der Teilnahme an Musik zugrunde liegt, bestätigen verschiedene Sensitivitätsanalysen die genannten Effekte. Der kausale Effekt von Musikunterricht wird somit besser identifiziert als in bisherigen Studien auf Basis repräsentativer Befragungsdaten.

Kapitel 3: Mit Daten des Sozio-ökonomischen Panels vergleicht diese Studie die Wirkung der Teilnahme an außerschulischem Musik- und Sportunterricht. So wirkt sich Musik positiver auf den Schulerfolg aus als Sport, besonders bei Mädchen und Jugendlichen aus gebildeteren Elternhäusern. Sport hingegen verbessert den selbst eingeschätzten Gesundheitszustand. Jugendliche, die sowohl Sport als auch Musik ausüben, erreichen einen noch höheren Schulerfolg. Auch in diesem Kapitel hält die kausale Interpretation zahlreichen Sensitivitätsanalysen stand, unter anderem indem der Effekt auf nicht betroffene Outcomes getestet wird, sowie anhand einer formalen Sensitivitätsanalyse und mit Instrumentenvariablen.

Kapitel 4: Musik, Sport und andere non-formale Bildungsangebote sind wichtiger Bestandteil jugendlicher Freizeitaktivitäten und wirken sich positiv auf die Entwicklung von Fähigkeiten aus. Doch der Zugang zu non-formaler Bildung wird stark durch die sozio-ökonomische Herkunft bestimmt. Diese Studie untersucht anhand von Daten der längsschnittlichen Haushaltserhebung Familien in Deutschland, ob Bildungsgutscheine für bedürftige Familien, wie sie durch das Bildungs- und Teilhabepaket zur Verfügung gestellt werden, die Teilnahme an außerschulischen Musik- und Sportaktivitäten erhöhen. Die Anspruchsberechtigung für das Bildungs- und Teilhabepaket hat insgesamt gesehen keinen Einfluss auf die Teilnahme an Musik und Sport. Unter Kindern, die mit beiden Eltern leben, sowie denen, deren Haushaltseinkommen oberhalb des Medians der Leistungsberechtigten liegt, erhöht sich die Teilnahme jedoch um 10 Prozentpunkte. Nur ein Sechstel der Leistungsberechtigten nimmt den Gutschein für kulturelle Teilhabe in Anspruch, von denen etwa die Hälfte bereits vorher musikalisch oder sportlich aktiv war.

Kapitel 5: Musikalische Bildungsprogramme erfreuen sich weltweit einer immer größeren Beliebtheit unter Bildungspolitikern. Sie haben zum Ziel, Kindern aus benachteiligten Familien das Erlernen eines Musikinstrumentes zu ermöglichen. Darüber hinaus werden diese Programme besonders von Bildungspolitikern als innovative Politikmaßnahmen dargestellt, die zur Entwicklung der Persönlichkeit beitragen. Allerdings wurden solche Auswirkungen bisher kaum wissenschaftlich nachgewiesen. Diese Studie untersucht das Programm *Jedem Kind ein Instrument* (JeKi) in Nordrhein-Westfalen. Daten des Sozio-ökonomischen Panels sowie der Erhebung Familien in Deutschland werden mit Regionaldaten zu Musik- und Grundschulen kombiniert. Anhand eines Differenzen-in-Differenzen-Schätzers zeigt die Studie, dass JeKi erfolgreich die Teilnahme an Musikunterricht unter Kindern aus sozial benachteiligten Elternhäusern erhöht, und zwar deutlich besser als durch die Reduzierung von Musikschulgebühren. Darüber hinaus reduziert die Teilnahme am Programm Verhaltensprobleme und verbessert das Verhältnis zwischen Schülern und Lehrern, besonders bei Jungen.

Schlüsselwörter: Non-formale Bildung, kognitive und nicht-kognitive Fähigkeiten, Ungleichheit, Evaluierung von Politikmaßnahmen, Sozio-oekonomisches Panel

Table of Contents

List of Tables

List of Figures

1 Introduction

With the work by James Heckman and coauthors, skills have become an important research topic in economics. Numerous studies find that cognitive skills, but also personality, behavior and other non-cognitive traits are important determinants for educational and labor market success (see for example Almlund et al., 2011; Heckman et al., 2006; Heineck and Anger, 2010). By consequence, many researchers in the economics of education study how these skills develop and how they can be shaped by policy interventions.

Economic research on skill development is often studied in the context of models of skill production (for example by Cunha et al., 2010; Todd and Wolpin, 2003). In these models, children typically have some innate ability or genetic endowment. Skill development depends not only on these endowments, but – most importantly for research in social sciences – also on various environmental factors, which are often summarized as family and school inputs. The productivity of inputs into the skill production function increases with the level of previously acquired skills. By consequence, skills developed at earlier ages foster the development of skills at later ages. Moreover, both cognitive and non-cognitive skills are particularly malleable during childhood (Heckman and Kautz, 2014).

Many researchers have investigated whether school and pre-school foster skill development. Among other findings, attending school was shown to improve cognitive skills in the short-term (Carlsson et al., 2014), in the long-term (Banks and Mazzonna, 2012), or among minorities (Cascio and Lewis, 2006). Improvements in non-cognitive skills were identified as the driving mechanism for long-term positive effects of policy interventions such as the Perry Preschool Project (Heckman et al., 2013). Overall, model programs were found more successful in stimulating non-cognitive skill development than large-scale policies such as Head Start (Currie, 2001).

The role of informal education, for example through parental input, was studied as well. Among other findings, more than half of the test score gap between individuals was attributed to home investments (Todd and Wolpin, 2007). Parental inputs affect cognitive skill development especially at early ages, while non-cognitive skills can be affected at later stages of child development as well (Cunha and Heckman, 2008).

In contrast to formal schooling and informal (parental) inputs, non-formal education has been studied much less. Educational activities are categorized as non-formal if they take place in an institutional context, but do not lead to a certificate. Typical examples are extracurricular music or sports activities, but also theater lessons, dance classes, and after-school clubs.

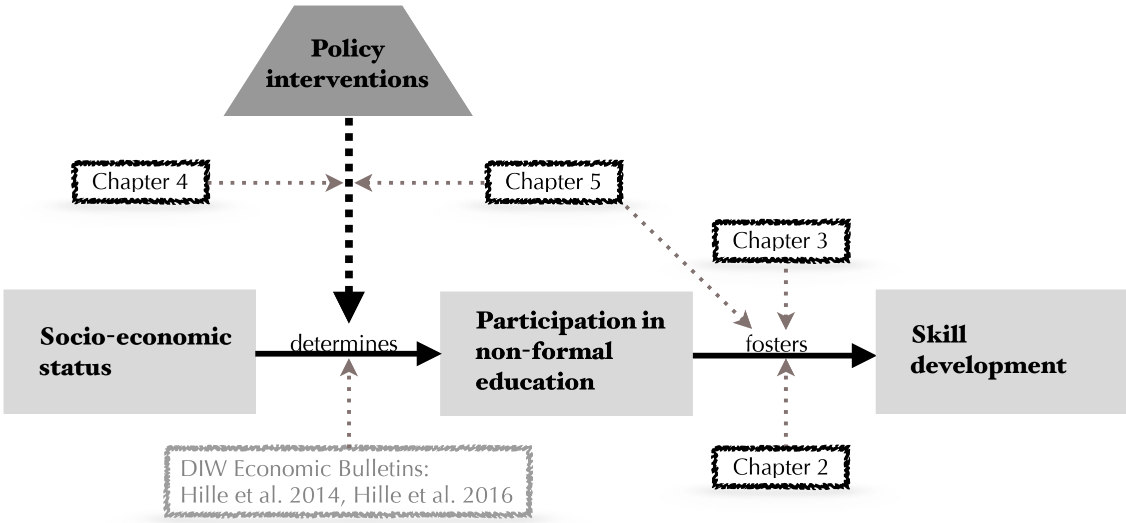
The low number of studies focusing on non-formal education is surprising. Especially in Germany, where full-day schools have only been widely established in the last decade, extracurricular music and sports activities account for a substantial share of children's leisure time. Moreover, many parents and policymakers strongly believe that these activities are beneficial for skill development. They are considered by parents as an investment into their child's human capital. It is therefore important to thoroughly investigate such claims using empirical research.

Access to non-formal education strongly depends on socio-economic background. Similar to other voluntary learning activities such as online courses, non-formal education theoretically offers disadvantaged children the opportunity to compensate for missing parental inputs or weaker school performance. In reality, unequal access to these activities might actually reinforce inequalities of educational opportunity (Emanuel, 2013; Hansen and Reich, 2015). Policymakers are increasingly aware of this and spend public funds on programs facilitating access to non-formal education for disadvantaged children. Such publicly funded programs have not yet been sufficiently evaluated.

The contribution of this doctoral thesis is twofold. In the first part, it aims to investigate whether non-formal education contributes to skill development and if yes, to which extent access to non-formal education depends on socio-economic background. Based on this, the second part explores possible policy interventions that make non-formal education more accessible to children from disadvantaged social backgrounds.

Within the area of non-formal education, this thesis focuses in particular on music education. So far, extracurricular sports activities have received some attention by economists. This is probably due to the fact that sports play a major role in American high schools and

Figure 1.1 – *Thematic overview of the chapters*



universities. Cabane and Lechner (2015), as well as chapters 2 and 3 of this dissertation provide an overview of the findings on the effects of extracurricular sports.

By contrast, the effects of learning a musical instrument have almost exclusively been studied by researchers in psychology and music education. Winner et al. (2013) and Schellenberg (2016) summarize these findings (including the results from chapter 2 of this thesis). However, these existing studies do not sufficiently discuss the issue of non-random selection, or focus on small samples of voluntary study participants (see chapters 2 and 3). Moreover, they discuss neither the socio-economic implications of their findings, nor the possible policy interventions that address these.

Figure 1.1 illustrates the topics of each chapter of this dissertation. Chapters 2 and 3 study how learning a musical instrument during adolescence affects the development of cognitive and non-cognitive skills. Cognitive skills are measured by school grades, as well as by a standardized test of fluid and crystalline intelligence. Non-cognitive skills refer not only to personality traits, such as the Big Five (conscientiousness, openness, agreeableness, neuroticism and extraversion) or the locus of control, but also to ambition, which I measure with the aim to attain an upper secondary school or university degree.

Chapter 2 studies the effects of taking musical instrument lessons by comparing adolescents with and without music training. Similar to most other studies on the effects of non-formal education, playing music is compared to a combination of all other possible counterfactual uses of leisure time. Chapter 3 starts by pointing out that such a vague comparison group might be misleading, because interpreting the effect of learning a musical instrument strongly depends on the activity that would have been carried out, had the adolescent not played music. Moreover, parents or policymakers might wonder whether other non-formal educational activities have similar effects. Therefore, chapter 3 compares the effects of learning a musical instrument to those of playing sports.

For both chapters, I use data on 17-year-olds from the German Socio-Economic Panel (SOEP). Adolescents in surveyed households answer the SOEP youth questionnaire in the year they turn 17. They answer numerous questions about their childhood and youth, providing information, for example, on music and sports activities, school grades, and educational ambitions. Moreover, cognitive skills and personality traits are assessed using standardized tests.

Studies on non-formal education face the challenge of credibly identifying causal effects. Given that these activities are usually voluntary, randomized experiments could provide reliable causal evidence (see for example Schellenberg, 2004). However, experiments are difficult to implement and typically identify short-term results for a small sample of voluntary study participants. By contrast, non-experimental studies have to overcome the potential bias resulting from non-random selection. Previous studies on sports participation tackle the selection problem using selection-on-observables, and sometimes instrumental variables, such as a change in law (Stevenson, 2010), or less credible instruments like height, distance to facilities, and school size.

Given that these potential instrumental variables are almost always too weak, it seemed more appropriate to base the identification of the effects of music training on selection-on-observables. The SOEP is an excellent data set for this purpose, as it contains numerous and very detailed characteristics on the adolescent's parents and households. To avoid unnecessary functional form restrictions, both chapters use propensity score matching to estimate the effects. Moreover, both chapters contain a large number of sensitivity tests to examine whether effects might be driven by unobserved heterogeneity.

To tackle the issue of non-random selection into music (and sports) participation, chapters 2 and 3 also contain substantial investigations on the determinants of accessing non-formal education. Two DIW Economic Bulletins were written in parallel to this thesis, which discuss the determinants of taking part in non-formal education and private tutoring, and how these determinants have developed over time (Hille et al., 2014, 2016).

The second part of this thesis starts from the observation that access to music and similar activities strongly depends on socio-economic background. Non-formal education might therefore increase the inequality of opportunities, given that it is beneficial for skill development. Therefore, two policy interventions are studied, which aim to provide access to music education for children from disadvantaged social backgrounds. They represent two approaches, which address different reasons of non-participation.

Chapter 4 studies the effects of a subsidy covering, among other child-related expenditures, music school and sports club membership fees. As an in-kind benefit, the Educational package releases the budget constraint among poor households. In theory, such a subsidy might increase participation in music or sports activities if financial constraints are an important reason not to participate. By contrast, if non-participation is related to education or preferences rather than financial constraints (as suggested among others by Weininger et al., 2015), a simple subsidy is unlikely to change behavior.

Chapter 5 studies an alternative policy approach, according to which primary schools directly offer musical instrument lessons in cooperation with local music schools. The *Jedem Kind ein Instrument* (an instrument for every child) program exists in numerous primary schools in the German federal state of North Rhine-Westphalia. Given that children are directly animated to participate in music lessons, the program might increase participation even if parents do not share the preference for artistic experiences.

Chapters 4 and 5 use data on children from Families in Germany (FiD), a longitudinal household panel study closely related to the SOEP. While chapter 4 uses FiD data only, chapter 5 makes use of the possibility to jointly analyze SOEP and FiD. In addition, data on all public music schools, as well as primary schools of North Rhine-Westphalia was merged to the sample of SOEP and FiD children for chapter 5.

To identify the effects of both policies on participation in music or sports activities for various subgroups, I estimate the difference-in-differences between eligible and non-eligible children before and after the policies were established. Chapter 4 concentrates on the question whether the subsidy increases music or sports participation among eligible children. In addition to participation of specific subgroups, chapter 5 also studies whether access to the *Jedem Kind ein Instrument* program affects non-cognitive skills, as measured by socio-emotional behavior and school attitude.

2 How learning a musical instrument affects the development of skills

2.1 Introduction

Publicly subsidized projects offering extracurricular music lessons are increasingly popular among policy makers. Having originated in Venezuela with the famous *El Sistema* (FMSB, 2013), such projects exist in many countries today. In Germany, the Federal Ministry of Education has launched *Kultur macht stark* (Culture makes you strong), a new initiative in 2013 with 50 million euros of funding per year (BMBF, 2012). On the regional level, the government of North Rhine-Westphalia provides annual support to the project *Jedem Kind ein Instrument* (An instrument for every child) in the amount of 10 million euros (JeKi-Stiftung, 2014). As stated in their official descriptions, these projects aim to improve educational opportunities, in particular for disadvantaged children and youth (e.g. BMBF, 2013).

The aim of reducing inequalities in educational opportunity through social policies promoting music education implicitly relies on the assumption that music fosters the development of cognitive and non-cognitive skills. The importance of these skills for educational and labor market success has been widely recognized (e.g. Almlund et al., 2011; Heckman and Kautz, 2012; Heckman et al., 2006; Heineck and Anger, 2010). According to findings on the "technology of skill formation", skills developed at younger ages promote later skill attainment (Cunha and Heckman, 2007; Cunha et al., 2010). In particular, the productivity of investments in subsequent stages increases as a result of previously acquired skills. Moreover, sociologists highlight that cultural capital – the familiarity with the codes and modes of conduct of particular social environments – influences success in education and the labor market (Bourdieu and Passeron, 1990; Lareau and Weininger, 2003; Lareau, 2011) and works as a mechanism for the reproduction of social inequality.

Numerous studies argue that music affects a variety of indicators of skill development (Winner et al., 2013). They argue that music induces brain reactions that stimulate the development of cognitive skills (Schellenberg, 2011). Moreover, both cognitive skills and school grades may be improved through the influence of music on personality traits such as conscientiousness, openness, and perceived control (Schumacher, 2009). A positive effect on the latter may also lead the musically trained to be more ambitious. Lareau (2011) highlights the fact that music practice, similar to other extracurricular activities, enhances educational success by sending positive signals to school teachers and by fostering children's acquisition of some elements of cultural capital. In addition, playing in an orchestra or a band can promote the development of social skills as well as the sense of belonging to a group. Finally, extracurricular activities consume time, which is then no longer available for other potentially beneficial or harmful activities (Felfe et al., 2011). Of course, leisure time occupations other than music can influence some of these outcomes similarly or even more effectively. Part of the challenge is to distinguish their differential effect.

The assumed positive effects of music, which even motivate social policy-makers, stand in contrast to a lack of causal research on this topic. Observational studies face the difficulty that the decision to learn a musical instrument is not made randomly. Causal studies must distinguish the effect of music from outcome differences related to observed and unobserved background characteristics. So far, only a small number of experimental studies are able to identify true causal effects. For example, Schellenberg (2004) finds that music lessons enhance general intelligence of children, but do not affect their social skills. This finding has been confirmed in further experimental studies for children (Neville, 2008; Nering, 2002; Bilhartz et al., 1999), but not for adults (Bialystok and DePape, 2009; Schellenberg and Moreno, 2010).[1] However, as these studies are experimental, they focus

[1] In addition, numerous studies in the fields of psychology and music education detect positive correlations between playing music and cognitive skill development. They find that music practice or training is associated with a higher IQ (Vaughn and Winner, 2000), an enhanced reading ability (Besson et al., 2007; Loui et al., 2011), increased attention (Shahin et al., 2008) and a better memory (Ho et al., 2003). Some of these relations remain after holding basic socio-demographic background characteristics constant (Schellenberg, 2006; Southgate and Roscigno, 2009). Still, most studies do not accurately address the issue of non-random selection into playing music and therefore detect correlations rather than causality (Winner et al., 2013).

on short-term music training within non-representative samples of voluntary participants. While many claim that music also benefits the development of non-cognitive skills, no study has yet proven that this is the case (Winner et al., 2013). To our knowledge, music as an extracurricular activity has not been studied by economists so far.

This paper examines how learning a musical instrument during childhood and adolescence affects the development of cognitive and non-cognitive skills using data from the German Socio-Economic Panel (SOEP). We investigate the effect of long-term exposure to music by restricting the treatment group to those who play music at age 17, have started to do so at age 8 or before, and have received music lesson outside of school. Outcomes are measured at age 17 and include cognitive skills, school grades, personality traits (Big Five and perceived control), time use, ambition and optimism about future success.

We address the non-random selection into music practice by controlling for a large number of individual and parental background characteristics. These are likely to influence the decision to engage with music, the constraints related to such a decision, as well as the willingness to carry on playing music until age 17. In particular, we control for parental income and education, household composition, the parents' personality and school involvement as well as the parents' taste for the arts. Moreover, we control for the adolescent's recommended and realized choice of upper secondary school track, as well as the predicted probability to give up music before age 17. We take these variables into account using propensity score matching.

Our findings suggest that learning a musical instrument during childhood and adolescence is associated with school grades which are one sixth of a standard deviation above those of musically inactive adolescents. Moreover, young adults with music training are more conscientious, open and ambitious. Outcome differences in school grades and personality are much stronger among adolescents from families with lower cultural capital. The effects of music are larger than those of playing sports, an activity which has been found an important input for skill development (Barron et al., 2000; Felfe et al., 2011; Pfeifer and Cornelissen, 2010; Stevenson, 2010). The causal interpretation of our findings depends on the conditional independence assumption. We discuss the plausibility of this assumption and conclude that the outcome differences we find are probably not entirely due to unobservable characteristics or reverse causality.

Our contributions to the literature can be summarized as follows. First, our study approaches causality better than previous observational studies on the effects of music. While we cannot entirely exclude the possibility that unobserved confounders drive our results, we account for more background characteristics than others have before. Moreover, sensitivity tests suggest that our results are robust to reverse causality. Second, we use a random sample of German adolescents to investigate our research question. This makes our results more generalizable than the experimental work which has been carried out mainly by psychologists. In particular, we are the first to use the SOEP to study this question, which contains parental background information even from when the adolescent was still a child. Third, our study examines the effects of music on a broader range of outcomes than previous studies have looked at. In addition to cognitive skills, we observe school grades, personality, time use and ambition. Fourth, contrary to previous observational and experimental studies in this field, we examine the effects of long-term exposure to music. We investigate the development of skills among adolescents who play music at age 17, have started to do so at age 8 or before and have taken instrumental music lessons. Our data allow us to construct alternative treatment definitions, to which our results are robust. Fifth, we examine the heterogeneity of the effect with respect to socio-economic status. Social policies promoting music education can only be effective if their treatment not only affects those from higher socio-economic backgrounds, who are likely to be involved in education-oriented leisure activities in any case. Indeed, we find that the effects of our treatment are stronger among adolescents with lower cultural capital. Finally, we compare the effect of music to the alternative of playing sports at a comparable level of intensity. In most previous studies on leisure activities, such distinctions are not explicitly made. We find that the effect of music is much stronger than that of sports.

Next, we describe why learning a musical instrument might influence educational opportunities. After a short summary of data and methodology, we present our findings. The study concludes with a discussion on the caveats of a causal interpretation of our results.

2.2 Mechanisms of the potential effect of music

Learning a musical instrument is widely believed to affect a variety of outcomes related to educational achievement (Winner et al., 2013), as summarized in Table 2.1. Any hy-

Table 2.1 – *Hypotheses: potential effects of music training*

Hypothesis	**Mechanism**
Cognitive skills	
→ Improved cognitive skills	Influence on subdomains of cognitive function, executive function or via non-cognitive skills (Schellenberg, 2004, 2011)
Non-cognitive skills	
→ Increased conscientiousness	Music requires self-discipline (Schumacher, 2009)
→ Higher perceived control	Judge ability, develop positive self-concept (Schumacher, 2009)
→ Increased openness	Contact with classical music
→ Increased ambition	Judge own ability, success and progress (Schumacher, 2009)
School achievement	
→ Improved school grades	Positive signal to school teachers (Lareau, 2011), improved cognitive skills
Cultural and social capital	
→ Enhanced cultural capital	Interaction with teacher in small group (Lareau, 2011)
→ Improved social skills	Interaction with peers and teachers (Schumacher, 2009)
→ Higher social well-being	Belonging to a group (Ormel et al., 1999; Menninghaus, 2011)
Time use	
→ Changes in time use	Crowding out of positive or negative activities (Felfe et al., 2011), structure learning and time schedule (Lareau, 2011)
Inequality	
→ Stronger effects for low SES	Efficiency of investment (Heckman and Masterov, 2007), cultural mobility (DiMaggio, 1982)
→ Stronger effects for high SES	Cultural reproduction (Bourdieu, 1986)

pothesis on such effects is based on the assumption that skills acquired through music are transferable to other domains.

Schellenberg (2011) considers three channels by which music potentially improves cognitive development. It might affect subdomains of cognitive functioning such as auditory temporal processing or visual memory. Alternatively, music training could stimulate the executive function. The executive function represents judgment and problem-solving capacities, which are particularly malleable during childhood and correlated with IQ. More indirectly, music may improve intelligence through its effect on non-cognitive skills.

Studying a musical instrument requires regular training and thereby forces students to be self-disciplined, persistent, and involved (Covay and Carbonaro, 2010). As a consequence, this may improve conscientiousness, a dimension of the Big Five personality traits. Other personality traits are likely to be affected by music as well.[2] According to Schumacher

[2] Some psychologists argue that personality is genetically determined and cannot be modified (Pervin et al., 2005), but the personalities of children have been shown to be less stable than those of adults. Heckman and Kautz (2012) highlight that the long-lasting positive effect of early childhood interventions, such as

(2009), learning a musical instrument teaches children to judge their ability to learn as well as their progress in learning. He calls this ability a positive self-concept. Similarly, Covay and Carbonaro (2010) point out that learning a musical instrument teaches a child to handle success and failure. In terms of personality traits, these experiences might affect perceived control. Perceived control indicates the extent to which someone believes to be able to influence their own destiny.[3] If music training increases perceived control, we also expect these children to be more ambitious.

In addition to skill improvements and personality changes, music may affect educational achievements through signaling effects (Lareau, 2011). If a teacher knows about a student's after-school musical activities, that teacher may reward the perceived rather than proven competence with a better grade than actually appropriate.

Improvements of social skills are possible as well. When music classes are taught in a group or an orchestra, students closely and directly interact with their peers. Typically, such interactions considerably differ from those in the classroom. Students have to learn to take over someone else's perspective, putting their own interests back for the benefit of the common goal. With other words, they have to learn to see their fellow students as partners rather than competitors (Schumacher, 2009). Furthermore, the contact with teachers in a small group may stimulate cultural capital. In particular, learning to interact with a person of authority can increase the child's sense of entitlement (Lareau, 2011).

Beyond these advantages in terms of skills and education, music training might enhance social well-being by giving individuals a sense of belonging to a group (Lindenberg, 1989; Ormel et al., 1999). Indeed, Menninghaus (2011) relates participation in the arts to the costly signal theory. He states that the possession of artistic objects or engagement in cultural activities is used to signal one's affiliation to a certain social status.

Finally, learning a musical instrument could influence educational achievement through its effect on time use. Whether the expected effect is positive or negative depends on whether playing an instrument reduces the time available for potentially beneficial or

the Perry Preschool Project, were attained through the program's positive impact on personality. Non-cognitive skills have proven to be particularly malleable at younger ages (Specht et al., 2011; Cobb-Clark and Schurer, 2012; Donnellan and Lucas, 2008).

[3] While the development of a positive self-concept might increase one's perceived control, learning a musical instrument might also be correlated with lower perceived control. Children learning a musical instrument often have parents who intervene strongly in their schedules and choices of free-time activities (Lareau, 2011). Hence, a potential positive effect on perceived control might be hidden due to a systematically lower level of perceived control among the non-random sample of musically active children.

harmful activities (Felfe et al., 2011). Still, children participating in extracurricular activities might learn to better manage their learning processes and time schedules (Lareau, 2011).

Estimating the mean effect of learning a musical instrument could hide important heterogeneities. Policies such as those mentioned in the introduction are aimed primarily at children from disadvantaged social backgrounds. Heckman and Masterov (2007) point out that such policies are among the rare which do not involve a trade-off between efficiency and fairness. DiMaggio's (1982) cultural mobility hypothesis suggests that children from disadvantaged social backgrounds benefit particularly, because these have a higher potential gain as a form of compensation for missing educational inputs from the children's families. In contrast, Bourdieu's (1986) cultural reproduction hypothesis argues that richer and more educated parents have access to better quality extracurricular activities, which stimulate skill development more successfully.

Besides understanding the effects of leisure activities on skill development, further research is needed to understand the extent to which these activities can act as substitutes. Some of the abovementioned effects are likely to result from extracurricular activities other than music as well.

2.3 Data

The German Socio-Economic Panel study (SOEP) is to our knowledge currently the best available longitudinal data set for studying the effects of learning a musical instrument. First, it contains a detailed assessment of the intensity and duration of music activities for a random sample of adolescents (Schupp and Herrmann, 2009). Second, the SOEP measures a large variety of outcomes such as school results, cognitive skills, personality, time use, and ambition. Third, given that it is a household rather than an individual survey, the SOEP allows us to directly observe numerous parental background characteristics (Wagner et al., 2007). Moreover, due to the longitudinal nature of the survey, these variables are available for when the adolescent was still a child.

Our estimation sample consists of all survey participants who have answered the SOEP Youth Questionnaire between 2001 and 2012, leading to a sample size of 3,941 observations. This questionnaire is administered to all SOEP household members in the year they turn

17. It contains youth-specific questions related to educational achievements and plans, activities during childhood and youth, as well as subjective questions on personality and opinions.

To define our treatment, we take advantage of this questionnaire's detailed assessment of music activities during youth. Young adults are asked to answer the following questions (Weinhardt and Schupp, 2011):

- *Question 16:* Do you play a musical instrument or pursue singing seriously?
  (Yes or no)

  If the answer is yes, the following further questions are asked:

- *Question 17:* What type of music do you make?
  (Classical, Pop/Rock/etc. or Folk music)
- *Question 18:* Do you do this alone or in some sort of group?
  (Alone/with teacher, in an orchestra/choir, in a band or in another type of group)
- *Question 19:* How old were you when you started? (Age)
- *Question 20:* Do you take or have you ever taken music lessons outside of school?
  (Yes or no)

With the answers to these questions, it is possible to construct a variety of treatment indicators. Compared to other data, the SOEP thus allows us to examine the effects of music at different intensity levels. In our main specification, we consider individuals to be musically active, if they (a) play a musical instrument at age 17 (answer "yes" to question 16), (b) have started to do so at age 8 or before (answer "8" or lower to question 19), and (c) who have taken music lessons outside of school (answer "yes" to question 20). Rather than simply studying adolescents who claim to be active in music at some point in time, we examine the effect of exposure to music at a minimum level of intensity. With the abovementioned questions, different treatment definitions can be constructed as well, which we will use as robustness checks. Table 2.2 describes the characteristics of music practice for various treatment definitions. The main definition described above is presented in column 1. The table shows that the characteristics of playing music do not vary that much depending on how it is defined. Still, our main treatment definition involves playing at a relatively high level of intensity.

Table 2.2 – *Characteristics of music practice*

	Various definitions of playing music						
	≤8 & 17 lessons	**17**	**17 lessons**	**>8 & 17 lessons**	**≤8 & 17**	**≤8 & 17 lessons classical**	**≤8 & 17 lessons weekly**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share playing...							
...classical music	0.47	0.30	0.38	0.29	0.44	1.00	0.46
...rock, pop or techno	0.38	0.50	0.43	0.47	0.40	0.00	0.39
...alone	0.53	0.47	0.54	0.57	0.50	0.64	0.50
...in an orchestra or choir	0.23	0.20	0.20	0.17	0.25	0.26	0.25
...in a band	0.17	0.22	0.19	0.21	0.17	0.06	0.19
...playing music daily	0.44	0.39	0.42	0.38	0.42	0.41	0.51
...taking music lessons outside school	1.00	0.74	1.00	1.00	0.90	1.00	1.00
Average starting age	6.4	9.4	8.5	10.9	6.4	6.3	6.4
Number of adolescents (total: 3,941)	417	963	709	257	465	154	360

Source: SOEP v29 (2001-2012 pooled), own calculations. Characteristics of music practice for various treatment definitions. Figures indicate the share of individuals, for whom the respective characteristic is true, except for starting age, where the age is indicated. Exact definitions: (1) Play music at age 17, started at age 8 or before, take music lessons. (2) Play music at age 17. (3) Play music at age 17, take music lessons. (4) Play music at 17, started at age 9 or later, take music lessons. (5) Play music at 17, started at age 8 or before. (6) Music at 17, started at 8 or before, have lessons, play classical music. (7) Music at 17, started at 8 or before, have lessons, play weekly.

Sport is the only other extracurricular activity which is assessed in similar detail in the SOEP Youth Questionnaire. This will allow us to compare our results to the alternative treatment of doing sports at a similar level of intensity. We consider as active in sports those who (a) play sports at age 17, (b) have started to do so at age 8 or before, and (c) regularly take part in sports competitions.

All outcomes examined in this study were taken from the SOEP Youth Questionnaire as well and are thus measured at the age of 17. In particular, we examine the effect of music training on cognitive skills, school grades, personality, ambitions and time use. Detailed information on all outcome variables can be found in Tables A.2 and A.3 of the appendix.

Cognitive skills have been measured since 2006 with a standardized test. This test consists of three subscores: analogies, figures, and mathematics operators (Schupp and Herrmann, 2009). The first assesses the individual's verbal knowledge and asks respondents to identify word pairs. To get a good score in Figures, one has to choose the correct symbol continuing a given row. Similarly, the test of mathematics ability requires individuals to insert operators in incomplete mathematical computations. In order to facilitate the interpretation of cognitive skills, all results were normalized.

In addition to directly testing their cognitive skills, the SOEP Youth Questionnaire asks young adults about their latest school grades in German, mathematics, and their first foreign language. Due to fundamental differences in educational programs, grades are not easily comparable between the three German secondary school tracks (Hauptschule, Realschule, Gymnasium). To facilitate comparisons, we normalize all school grades within each type of secondary school.

The SOEP Youth Questionnaire investigates various dimensions of personality using simplified psychologically validated items to which respondents state their level of agreement on a Likert scale. We investigate the effect of music on the Big Five personality traits (McCrae and Costa, 1999; Lang et al., 2011), which include conscientiousness, openness, agreeableness, extraversion and neuroticism. Moreover, our hypotheses suggest that practicing a musical instrument may affect perceived control. Someone is characterized by a high level of perceived control if she or he believes to be able to influence their own destiny (Specht et al., 2013). For each dimension, we use the average answer among all items. For detailed descriptions of the items, please refer to Table A.3 in the appendix.

In addition to assessing current skills and personality, the SOEP Youth Questionnaire asks about plans and worries for the future. As further outcome variables, we consider the young adult's plans to obtain an upper secondary school degree (Abitur) as well as a university degree. Moreover, respondents are asked to estimate the probabilities to find a job in their desired occupation and to be successful in their job. Finally, a measure indicating whether the individual watches TV and reads daily will allow us to examine how learning a musical instrument affects the adolescent's use of leisure time.

Due to the longitudinal nature and household dimension of our data, we are able to merge rich background information on each adolescent's family as well as information on the individual's childhood.[4] This is important because families with children who learn a musical instrument differ strongly from others. In addition to the standard socio-economic characteristics of the parents such as education, income, and household composition, we observe some important aspects that are likely to influence the decision to enroll the child into music lessons. In particular, our data contain the parents' personality, involvement

[4] We take this information from the standard SOEP household and individual questionnaires.

in the child's education as well as taste for the arts. For all parental variables, we use observations on the mother. If not available, we replace them with those for the father.[5] Time-varying variables were measured as early as available for each individual, but no earlier than age five.[6] Please refer to Table A.4 in the appendix for a list of available control variables and when they were observed.

Without considering missing values among the outcome variables, our sample contains 3,941 observations, 417 of which are treated according to the definition described above.[7] We then construct three final samples for three groups of outcome variables: those measured in every survey year (3,488 observations), cognitive skills (1,847 observations) and the Big Five personality traits (1,815 observations). Within each group, we carry out our estimations with a single sample for which we have complete information on all outcomes. The outcomes in the latter two groups were not measured every year, the sample sizes are therefore considerably smaller. Please refer to Appendix A.1 for details on the sample construction. Summary statistics for all outcome and control variables can be found in Tables A.5, A.6 and A.7 of the appendix.

2.4 Empirical approach

The decision to play and learn a musical instrument is not made randomly. This study takes selection into account by controlling for numerous observable characteristics. Outcome differences between musically active and inactive adolescents are estimated using propensity score matching. According to our treatment definition, non-random selection takes place at two stages: the decision to engage with music and enroll in music lessons, as well as the decision to carry on playing music until the age of 17.

The decision to engage with music at an early age and take up music lessons is likely to be strongly influenced by the parents. Such a decision may be motivated by utility and

[5] However, our results are also robust to including covariates for the father, in addition to or instead of those of the mother (see Appendix A.3).

[6] We do not include variables such as household income and the parents' artistic activities when the child was younger than five, because parents with very small children may be in a particular situation. About 60 percent of our sample entered the data after age five of the child. On average, individuals enter our sample when they are 8.2 years old. 53 percent of all treated enter the SOEP no more than three years after the start of the treatment. Our results are robust to using the subsample of individuals for whom background characteristics are available at age 5 or before they start to play music (see Appendix A.3).

[7] Missing covariates are replaced by 0 for binary and by their mean for continuous variables. Missing indicators are included. Our results are robust to using the subsample of individuals for whom we have complete information on all covariates, as well as on all outcomes (see Appendix A.3).

taste. In addition to the direct utility or pleasure someone derives from making music, parents are likely to consider music training as an investment in their child's future success (Eide and Ronan, 2001; Lareau, 2011). Moreover, they might enroll their child in music lessons because their own previous experience or habit of arts consumption has led them to develop a taste for the arts, as postulated in the learning-by-consuming theory (Garboua and Montmarquette, 1996). A preference for the arts depends on the educational level if we assume that more highly educated people are more able to appreciate artistic production and consumption (Lunn and Kelly, 2009). Those parents might also be more likely to enroll their children in music lessons, because the artistic activities available to children are adapted to the tastes of the more highly educated (Lunn and Kelly, 2009).

In our selection model, we address these motives as follows. First, both utility-based and taste-based motivations are related to socio-economic status (Yaish and Katz-Gerro, 2012). We control for the mother's education, qualifications, and migration background as well as her age at the child's birth. Moreover, parents are more likely to consider music lessons as an investment in children's future skills if they are more involved with the child's school activities in general. Therefore, we control for parents' contact with school, which is approximated by their disposition to help with homework, regularly meet the teacher, or be involved in other ways.[8] Furthermore, we control for the mother's personality, which might play a role in her eagerness to invest in the child's skill development. Finally, to take parental taste into account, we control for the mother's frequency of attending cultural events and being artistically active, her openness toward the arts, as well as the number of books at home.

Even though willing to send their children to music lessons, parents might face financial constraints or are lacking the means to effectively support their offspring's musical activities. We address this issue by controlling for net monthly household income. In addition, we include variables approximating the complications involved in enrolling a child in music lessons: the number of rooms per person at home, the number of siblings and birth order, whether the household lives in a rural area and federal state fixed effects. Moreover, we argue that financial constraints are less pronounced in Germany than elsewhere, given that

[8] These variables are retrospectively reported by the adolescent herself at age 17.

the association of German public music schools obliges its members to propose reduced fees to individuals with low income (VDM, 2011).[9]

This study considers adolescents as treated, if they played music at age 8 or before and continued to do so until age 17.[10] Only few researchers have studied the reasons for giving up instrumental music lessons during adolescence. Switlick and Bullerjahn (1999) conclude that major determinants are a lack of motivation, critical life events, problems at school and organizational issues. To the extent that those are not correlated with the parents' background characteristics described above, we deal with this issue by including characteristics of the adolescent herself. Lacking more precise measures, we proxy skills prior to playing music by controlling for adolescents' secondary school track. Moreover, we know the type of secondary school their teacher recommended at the end of primary school.[11] In addition, we control for gender and birth year fixed effects. In a robustness check, we estimate the probability to give up music before age 17 within a subsample, for which we observe the history of extracurricular music activities from age 12 to 17. Please refer to Section 2.5 for details. Due to the lack of pre-treatment cognitive and non-cognitive skills in our data, results might be driven by reverse causality. Please refer to Section 2.6 for a discussion of this issue.

Taking these variables into account, we estimate average outcome differences between musically active and inactive adolescents using propensity score matching. While similar results can be obtained with a simple linear regression model, propensity score matching has three advantages in our context. First, contrary to ordinary least squares, we do not need to assume a linear relation between covariates and outcomes. Especially when the distribution of covariates differs between treatment and control group, and if both groups strongly differ in size, functional form assumptions may provide results which are sensitive to minor changes in the specification (Imbens, 2015). Second, propensity score matching allows us to use a large number of control variables even with a limited sample

[9] The association of German public music schools (Verband deutscher Musikschulen) is the umbrella organization of approximately 1000 publicly funded music schools covering the entire country.

[10] Unfortunately, we are not able to identify individuals, who played music at a younger age, but gave up before answering the SOEP Youth Questionnaire at age 17. The implications of this are discussed in Section 2.6.

[11] In Germany, after primary school at age 10, students have to choose their secondary school track. They choose between upper secondary school (Gymnasium), which leads to a university entrance certificate, medium secondary school (Realschule), which provides the qualifications for vocational training and lower secondary school (Hauptschule). Teachers issue a recommendation for one of the three tracks at the end of primary school.

size, given that these are summarized in the propensity score (Huber et al., 2013). Finally, in a setting where children learning a musical instrument are likely to significantly differ from other children, the necessity to verify the existence of common support will provide us with some hint as to the comparability of treatment and control group (Imbens and Wooldridge, 2009). Given these differences and the fact that our control group is ten times larger than the treatment group, propensity score matching appears to be the best methodological choice.

The estimator was implemented in the following way. First, we estimate the probability to be treated with a probit model. This probability is called the propensity score. Ideally, all variables influencing the decision to enroll in music lessons and carry on playing until age 17 should be included in the selection model.

A table with all coefficients of the selection model can be found in the appendix (Table A.8). Overall, the selection model is able to explain about 18 percent of the variation in music practice, which is comparable to similar studies (for example Felfe et al., 2011). The coefficients are not surprising: Given that many of the covariates are correlated with each other, some of them are not statistically significant. According to recommendations from the statistics literature (Stuart, 2010), we include these insignificant coefficients in the selection model, as the aim is not to find the best model explaining the treatment but to balance observable (and if possible unobservable) characteristics in the treatment and control groups as much as possible.

Figure A.1 in the appendix shows common support between the treatment and control group. The figure indicates that untreated individuals (children who do not learn a musical instrument) are more likely to have a low propensity score, a further indicator that our selection model predicts musical practice quite well. Still, for the majority of the distribution, it is possible to find at least one corresponding control observation for each treated individual. With a caliper of one percent, only 3 out of 417 individuals in the treatment group and none in the control group remain unmatched.

Next, we find matches in the control group for each treated individual. We use radius matching with a caliper of one percent, meaning that we give equal weight to each control observation having a propensity score in the range of one percent around the corresponding treated observation. The equal weights of the control observations being matched to one treated observation are chosen to sum to one. Table A.9 in the appendix shows that all covariates are balanced after matching.

We now calculate the mean outcome difference between the treatment and the weighted control group. If our selection model is able to control for all relevant variables, this corresponds to the average treatment effect on the treated, ATT (Imbens and Wooldridge, 2009).[12] Algebraically, we estimate:

$$A\hat{T}T = \frac{1}{N_T}\sum_{i=1}^{N_T} T_i y_i - \frac{1}{N_C}\sum_{i=1}^{N_C}(1-T_i)\hat{w}_i y_i \tag{2.1}$$

where N_T and N_C are the number of treated and control observations, T_i is the treatment indicator, and y_i the outcome for individual i. Control observations are weighted with weight $\hat{w}_i$, which is obtained from matching as described above. Standard errors are clustered at the family level and estimated by bootstrap with 1999 replications.[13]

Given that we do not observe an exogenous variation in treatment participation, an interpretation of our results as causal effects of music training relies on the conditional independence assumption. To estimate the effect of music on skill production, we need to assume that the decision to learn a musical instrument is uncorrelated with any unobserved characteristics which also have an influence on the development of skills, to the extent that they are uncorrelated with the observable characteristics we control for. Section 2.6 discusses the plausibility of this assumption.

2.5 Results

2.5.1 Main results

Table 2.3 describes outcome differences at age 17 between adolescents with and without music training. Column (1) shows raw differences without controlling for any individual or parental characteristics. All other columns take selection into account using propensity score matching. As described in the previous section, we control for many variables related to the parents' motivation to enroll their children in music classes, financial and

[12] To obtain the average treatment effect, it would be necessary to separately estimate the average treatment effect on the non-treated by finding matches in the treatment group for each control observation. As the treatment group is ten times smaller than the control group, the quality of such matches would be low. To obtain robust results, we restrict our analysis to the average treatment effect on the treated.
[13] An analytical approximation would be possible, but is not recommended by Huber et al. (2015).

Table 2.3 – *Outcome differences between adolescents with and without music training or alternative activities*

	Treatment: Music (age 8 and 17, music lessons)				**Treatment: alternatives**		
	Main results		by cultural capital		*Main def. + classical*	*Sports (age 8 and 17, competitions)*	*Sample size*
	No controls (1)	*Full controls* (2)	*Low* (3)	*High* (4)	(5)	(6)	
Cognitive skills (in std. dev.)							
Average cognitive skills	0.49*** (0.07)	0.07 (0.08)	0.06 (0.13)	0.05 (0.12)	0.22 (0.14)	0.06 (0.06)	1,847
Analogies	0.59*** (0.07)	0.14 (0.09)	0.17 (0.14)	-0.01 (0.13)	0.21 (0.15)	-0.01 (0.07)	1,847
Figures	0.43*** (0.07)	0.09 (0.08)	0.14 (0.14)	0.08 (0.13)	0.21 (0.15)	0.00 (0.07)	1,847
Maths operators	0.20*** (0.07)	-0.03 (0.08)	-0.12 (0.13)	0.04 (0.13)	0.12 (0.15)	0.12^{+} (0.06)	1,847
School grades[1] *(in std. dev.)*							
Average school grade	-0.31*** (0.05)	-0.18*** (0.06)	-0.23** (0.11)	-0.03 (0.10)	-0.21^{+} (0.12)	0.05 (0.05)	3,488
German grade	-0.31*** (0.05)	-0.16** (0.06)	-0.14 (0.11)	-0.11 (0.10)	-0.16 (0.11)	0.07 (0.05)	3,488
Foreign language grade	-0.26*** (0.05)	-0.13** (0.06)	-0.09 (0.11)	-0.05 (0.10)	-0.17 (0.11)	0.11** (0.05)	3,488
Mathematics grade	-0.15** (0.06)	-0.11^{+} (0.07)	-0.28** (0.13)	0.08 (0.10)	-0.14 (0.13)	-0.06 (0.05)	3,488
Personality (in std. dev.)							
Conscientiousness	0.19*** (0.07)	0.23** (0.10)	0.23 (0.14)	0.13 (0.15)	0.16 (0.16)	-0.02 (0.07)	1,815
Openness	0.49*** (0.07)	0.26*** (0.09)	0.43*** (0.16)	0.12 (0.15)	0.10 (0.18)	-0.12^{+} (0.07)	1,815
Agreeableness	0.17** (0.07)	0.13 (0.09)	0.23 (0.15)	0.09 (0.14)	-0.06 (0.16)	0.03 (0.07)	1,815
Extraversion	0.07 (0.07)	-0.01 (0.09)	-0.04 (0.15)	0.00 (0.14)	-0.13 (0.16)	0.23*** (0.07)	1,815
Neuroticism	0.06 (0.07)	0.05 (0.09)	0.02 (0.16)	-0.00 (0.14)	0.02 (0.18)	-0.04 (0.07)	1,815
Perceived control	0.18*** (0.05)	0.03 (0.06)	0.27** (0.11)	-0.10 (0.09)	-0.02 (0.11)	0.03 (0.05)	3,488
Time use (in %)							
Watch TV daily	-19.34*** (2.75)	-11.05*** (3.32)	-12.99** (5.37)	-10.13** (5.12)	-11.40** (5.67)	1.12 (2.35)	3,488
Read books daily	15.99*** (2.69)	6.03^{+} (3.48)	14.14*** (5.44)	1.11 (5.37)	9.00 (6.08)	-4.09^{+} (2.26)	3,488
Ambition (in %)							
Aim Abitur	38.46*** (2.33)	4.13^{+} (2.15)	6.04 (4.65)	2.90 (3.32)	1.81 (3.59)	-0.38 (2.02)	3,488
Aim university	40.05*** (2.65)	8.20*** (2.93)	8.81 (5.49)	7.55 (4.59)	9.87** (4.89)	0.16 (2.16)	3,488
Job success likely	1.86** (0.93)	1.25 (1.13)	3.70^{+} (2.06)	-0.06 (1.70)	-1.54 (2.09)	2.71*** (0.88)	3,488
Desired profession likely	2.00^{+} (1.04)	2.19^{+} (1.25)	5.49** (2.24)	-0.79 (2.06)	-0.36 (2.34)	2.77*** (1.06)	3,488

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. Column (1) shows outcome differences without controlling for selection. Column (2) shows the baseline results for the full sample. Columns (3) to (4) show differential effects by cultural capital. Low cultural capital: Fewer than 200 books at home. High cultural capital: More than 200 books at home. Colums (5) and (6) show the effect of alternative activities: Play classical music (main treatment definition, restricted to those who play classical music); Sports (at age 17, started at age 8 or before, regularly participates in competitions). Propensity score matching is used to account for control variables (radius matching with caliper 0.01). The sample size is smaller for cognitive skills and some personality measures, as these have only been assessed since 2006. Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

material constraints within the household to do so, as well as characteristics of the adolescents themselves as a proxy of their willingness to carry on until age 17. Columns (2) to (6) describe the potential effect of music for different subgroups and treatment definitions. Further results showing outcome differences for various subgroups and treatment definitions can be found in Appendix A.3.

In each row, we estimate the difference between musically active and inactive adolescents with respect to the outcome specified on the left. Outcomes are grouped into five categories. Differences in cognitive skills, school grades, and personality are measured in standard deviations. Differences in time use and ambitions are stated in percentage points.

The coefficients presented in Table 2.3 describe the following picture. Without taking individual and family characteristics into account (column 1), adolescents with music training fare much better. Differences are most important with respect to cognitive skills. On average, musically active adolescents score one half of a standard deviation above their non-musical peers. This is in line with the vast amount of previous research on music, which almost exclusively concentrates on cognitive skills as outcome measures. Average school grades of musically active adolescents are almost one third of a standard above those of their non-musical peers. Note that in Germany, the grading scale runs from 1 (highest possible score) to 6 (failing). Enormous differences can also be observed with respect to personality, time use and ambition.

Once we hold observable characteristics constant as described above, outcome differences between musically active and inactive adolescents greatly reduce (column 2). However, a substantial part of these differences still remains unexplained by the a large number of covariates we control for.

Taking selection into account, the enormous advantage of music with respect to cognitive skills almost entirely disappears and becomes insignificant. Only in some specifications (among individuals who play classical music or attend upper secondary school), cognitive skills are still potentially affected by music, as described below. However, even after including control variables, a strong positive association of music and school grades remains. Adolescents who learn and play a musical instrument obtain an average school grade of approximately one sixth of a standard deviation above that of other individuals. This is a large difference, similar to the effect of introducing a central exit examination (Jürges et al., 2005). Better school grades despite a lack of improvements in cognitive skills could hint toward the existence of a positive signaling effect.

With respect to personality, adolescents with music training differ significantly as well. They are almost one fourth of a standard deviation more conscientious and open than others.[14] Learning a musical instrument is not correlated with agreeableness, extraversion or neuroticism. Contrary to what one would expect, children who learn a musical instrument are also not characterized by a higher perception of control.

If we look at time use and educational ambitions, we see systematic differences between the two groups as well. Children who learn a musical instrument are eleven percent less likely to watch TV every day. Moreover, they are eight percent more likely to aim at attending university.

In the next step, we examine the heterogeneity of these outcome differences with respect to cultural capital (columns 3 and 4). We consider families with more than 200 books at home as characterized by a high level of cultural capital. According to this definition, our sample contains 1,219 individuals with high and 2,722 individuals with low cultural capital. Among the former, 21 percent (255 adolescents) played music at age 8 and 17 and have taken music lessons, while only 6 percent (162 adolescents) of the latter did so.

Outcome differences between adolescents with and without music training are considerably higher among those with low cultural capital. In particular, differences with respect to school grades and personality are entirely driven by the latter. This supports the cultural mobility hypothesis by DiMaggio (1982), according to which children from less favorable social backgrounds have deficits to catch up and music can help them do so. An interesting observation can be made with respect to perceived control. Musically active adolescents with few books at home have a higher degree of perceived control than their peers, whereas individuals with high cultural capital have a lower degree of perceived control if they play music. Possibly, adolescents with high cultural capital, who play music, are more influenced by their parents and therefore less in control of their life.

Heterogeneities with respect to other subgroups are presented in Tables A.11 and A.12 of the appendix. Distinguishing by parental education as well as the adolescent's upper secondary school track confirms that individuals from less advantaged backgrounds benefit

[14] The higher value for openness is at least partly related to the fact that one of the three items assessing openness deals with openness to artistic experiences.

more from music. Moreover, outcome differences between musically active and inactive adolescents do not differ much by gender and living area.

Variations in the treatment definition confirm the robustness of our results and yield additional insights. Column (5) of Table 2.3 shows outcome differences between musically active and inactive adolescents. Here, we define playing music exactly as in the baseline specification – (a) play music at age 17, (b) have started no later than age 8 and (c) take or have taken music lessons outside of school –, but additionally require treated individuals to (d) play classical music. This is true for 47 percent of our original treatment group, the other half playing pop, rock, techno or any other type of non-classical music. Using this definition, only 154 adolescents (4 percent of the total sample) are treated.

Adolescents who play and learn classical music show the same outcome advantages as those playing music according to the baseline definition. They additionally score one fifth of a standard deviation above their peers in the cognitive skills test. The positive effect of music training on cognitive skills found in previous research could therefore be due to the focus of these studies on classical music.

Further results from variations in the treatment definition can be summarized as follows (see Table A.14 in the appendix). Better school grades can only be observed among musicians who take or have taken music lessons, while a higher level of conscientiousness depends on having started to play music at a young age. School grades might thus be influenced through serious involvement, while conscientiousness is shaped through long-term exposure. Unsurprisingly, all associations between music, school grades and personality are weaker if we consider individuals who played music for a shorter period. The higher aim to attend university among music participants is stable across all treatment definitions, which might be due to the fact that this outcome reflects an opinion expressed at age 17, when adolescents were active in music according to all treatment definitions.

An apparent question is whether the outcome differences presented above are specific to music. In column (6) of Table 2.3 we consider as treated those who (a) play sports at age 17, (b) have started to do so at age 8 or before and (c) regularly take part in sports competitions. Such a restrictive definition of playing sports ensures the comparability with playing music, because it describes athletic activities at a comparable level of intensity. 578 individuals (15 percent) of our sample play sports according to this definition. The results show that participating in sports competitions is associated with better mathematics skills, a higher extraversion and a more optimism about future success. However, we do not observe the strong outcome differences that characterize musically active adolescents.

2.5.2 Robustness checks

In addition to varying the treatment definition, all results presented in this section are robust to different estimation methods, specifications and sample restrictions. For example, we obtain the same results using OLS, as well as if we vary the caliper of radius matching or apply kernel matching. Moreover, our results are not sensitive to modifications in the composition of the sample studied. We do not find different outcomes if we drop individuals who entered the sample after starting to play music, as well as within the subsample of adolescents for whom we observe all outcome variables. Neither do our estimations yield different results if we control for the father's rather than the mother's characteristics, or if we include interactions between the number of siblings, birth order and household income as additional controls. Details on these robustness checks can be found in Appendix A.3.

Some individuals in our control group play music as well. They are not in the treatment group either because they did not take music lessons outside of school, started to play music later than age 8 or gave up before age 17. Irrespective of the conditional independence assumption, the existence of partly treated individuals in the control group influences the direction in which our estimates differ from the true effect. If adolescents benefit to some extent even from short-term musical experience, we do not have to worry. In this case, our estimations represent a lower bound. A more worrisome conclusion follows if short-term musical experience harms skill development. If the true effect of music training is positive when carried out from age 8 to 17, but negative for shorter periods – or for playing music without receiving lessons –, our results are overestimations even if the conditional independence assumption is valid. In this case, we would add the positive effect in the treatment group to the negative effect among the partly treated. In order to obtain the true effect, we would have to subtract these effects from each other.

Further robustness checks in Table A.17 of the appendix provide some empirical evidence according to which partly treated individuals weakly benefit from music as well. We are able to identify those who started music practice later than age 8, because they answer the relevant questions at age 17. Columns (10) and (11) of Table A.17 show the effects of music for these individuals. Here we consider adolescents as treated, who play music at age

17, but do not fulfill the requirements of the treatment definition in our main specification. Moreover, we exclude the treated individuals of our main specification from the sample. The effects of playing a musical instrument later than the age of 8 are weaker, but still positive, compared with children who start to learn a musical instrument earlier.

Unfortunately, it is more difficult to estimate the effect for those who gave up music before age 17, because the SOEP Youth Questionnaire does not ask them about past musical activities. Around 60 percent of all musically active children give up in their early teenage years. The three most important reasons are a lack of motivation, critical life events, and dissatisfaction with the teacher (Switlick and Bullerjahn, 1999). Hence, giving up music is possibly related to weaker school performance. We can test this hypothesis using the SOEP household questionnaire, which has been asking parents about their child's leisure time activities on a biannual basis since 2006. This allows us to construct a random subsample of 328 individuals, for whom we observe the complete history of musical activities since age 12.[15] We can therefore compare outcome differences between those who never played a musical instrument after age 12 to those who did so at age 12, but gave up before age 17. Similar to all other estimations, we apply propensity score matching to account for observable family background characteristics.[16]

The effects of music for this small subsample are presented in column (12) of Table A.17. Due to the small sample size, none of the outcome differences is significant. Still the direction of the effect is the same for most outcomes. However, the school grades of individuals who gave up music are lower than among those who were never involved. For the causal interpretation of our main results, this means that we might overestimate the true effect of music on school grades, because some of the partly treated individuals suffer from the stresses of demanding musical practice. However, this is at least partly balanced by other partly treated individuals in the control group, who benefit from music, as described above.

[15] The subsample is random because it consists of all observations which were interviewed in 2011 and 2012 and have been in the SOEP since 2006. These facts are exogenous to all individual or family characteristics.

[16] However, due to the small sample size, we had to omit some covariates in order to avoid collinearity. Moreover, the sample size is too small to estimate standard errors by bootstrap. We present standard errors which do not take into account that the propensity score was estimated.

2.6 Discussion: causal effects or correlations?

Even after controlling for a large number of socio-economic characteristics, we find strong differences between adolescents who learned a musical instrument during childhood and those who did not. In order to interpret these as causal effects, we must rely on the conditional independence assumption. Violations of this assumption can occur due to omitted variables or reverse causality. In the following, we discuss the extent to which such violations can be tested and are a concern to our results.

The conditional independence assumption would be invalidated if unobserved characteristics influenced the decision to learn a musical instrument and also had an impact on the outcome variables of interest. As described in Section 2.4, estimation biases resulting from selection into treatment can take place at two stages: The initial decision to take up music lessons and the decision not to give up until age 17.

We argue that our estimation satisfies the conditional independence assumption with respect to the initial decision to engage with music at age 8 or before. At such a young age, the choice of a long-term extracurricular activity such as music is strongly determined by the parents. For the parents, however, we observe a large number of characteristics, in particular their socio-economic status, personality, involvement with the child's education, and taste for the arts. All of these characteristics are strongly correlated among each other and therefore also likely to be correlated with any unobserved characteristics we might miss. The influence of unobserved characteristic invalidates a causal interpretation of our results only to the extent that these are uncorrelated with the observed characteristics we control for (Stuart, 2010).

The decision to continue music until age 17 is more likely to be based on unobserved characteristics of the child. We take this selection step into account by controlling for a variety of characteristics related to the adolescent herself. Most importantly, our matching estimator takes into account whether the adolescent attends upper secondary school track (Gymnasium), as well as the school track recommendation received from the teacher at the end of primary school. Whether the child goes to upper secondary school is one of the most important predictors of educational achievement in Germany. It is determined by a variety of background characteristics, some of which are unobservable to us. Therefore, controlling for the attendance of Gymnasium will help us capture some further unobserved characteristics of the adolescent.

To further examine the robustness of our results, we are able to estimate to probability to give up music before age 17 for a random subsample of 328 individuals. We can retrace the history of musical activities for these individuals back to age 12, as described in Section 2.5 above. 50 of them (18 percent) played a musical instrument at age 12, but gave up before age 17. This is a large number, given that among the individuals of the subsample only 43 (15 percent) have learned a musical instrument according to our treatment definition from age 8 to 17. Using all covariates and outcome variables of our main analysis, we estimate the probability to belong to the group of those who gave up music within the subsample. With the coefficients of a probit model, we predict this probability for the entire estimation sample. The predicted probability to give up music is then added as an additional control variable in our main estimation.

Table 2.4 shows outcome differences between musically active and inactive adolescents adding additional control variables.[17] Each column includes the variables mentioned in the table header as intermediate variables. As shown in column (1), our results do not differ from those of the main specification if we include the probability to give up music as control variable.

Another possible explanation for the positive association between music and cognitive as well as non-cognitive skills could be reverse causality or simultaneity. For example, musical activities do not increase ambition, but the more ambitious children tend to learn a musical instrument. Similarly, reverse causality could explain the positive correlation between music lessons and other outcomes such as conscientiousness and school grades. Ideally, we would exclude the possibility of reverse causality by controlling for pre-treatment values

[17] The propensity to give up music and other additional control variables presented in Table 2.4 can only be measured after the start of the treatment. Previous music training thus possibly has an influence on them. Therefore, we do not include these variables in the selection model for the propensity score, but add them as control variables once we estimate outcome differences between treatment and matched control group. This approach is similar to mediation analysis. Mediation analysis is usually used to differentiate between mechanisms through which the treatment affects the outcome (Imai et al., 2010; Heckman and Pinto, 2013). To identify causal mediation effects, the intermediate variable must satisfy the sequential ignorability assumption, according to which the mediator is independent of both treatment and outcome. Our aim here is to exclude, rather than identify the effect which runs through the channel of the intermediate variable. As we are solely interested in the effect which does not go through the mediator, we do not need to assume sequential ignorability.

Table 2.4 – *Outcome differences between adolescents with and without music training controlling for additional intermediate variables*

	(1)		(2)		(3)		(4)		(5)	
Intermediate variables additionally controlled for										
Probability to give up music before age 17	X		X		X		X		X	
Average cognitive skills			X							
Agerage school grade					X					
Conscientiousness							X			
Perceived control									X	
Cognitive skills (in standard deviations)										
Average cognitive skills	0.08	(0.08)			0.05	(0.08)	0.06	(0.09)	0.09	(0.07)
Analogies	0.14	(0.09)	0.08	(0.07)	0.11	(0.09)	0.11	(0.11)	0.15^{+}	(0.08)
Figures	0.10	(0.08)	0.04	(0.07)	0.09	(0.08)	0.09	(0.10)	0.12	(0.08)
Maths operators	-0.03	(0.08)	-0.09	(0.06)	-0.05	(0.08)	-0.03	(0.10)	-0.03	(0.08)
School grades[1] *(in standard deviations)*										
Average school grade	-0.18***	(0.06)	-0.19**	(0.09)			-0.19**	(0.09)	-0.17***	(0.06)
German grade	-0.16***	(0.06)	-0.18^{+}	(0.10)	-0.03	(0.04)	-0.20**	(0.09)	-0.16***	(0.06)
Foreign language grade	-0.13**	(0.06)	-0.14	(0.10)	0.01	(0.04)	-0.15^{+}	(0.09)	-0.13**	(0.06)
Mathematics grade	-0.12^{+}	(0.07)	-0.13	(0.10)	0.02	(0.05)	-0.10	(0.10)	-0.12^{+}	(0.07)
Personality (in standard deviations)										
Conscientiousness	0.24**	(0.09)	0.23^{+}	(0.12)	0.14	(0.09)			0.21**	(0.09)
Openness	0.26***	(0.10)	0.25**	(0.12)	0.27***	(0.09)	0.25**	(0.10)	0.28***	(0.09)
Agreeableness	0.14^{+}	(0.08)	0.07	(0.12)	0.13	(0.09)	0.08	(0.09)	0.13	(0.09)
Extraversion	0.01	(0.09)	-0.06	(0.12)	0.02	(0.09)	0.00	(0.11)	-0.00	(0.09)
Neuroticism	0.04	(0.09)	0.12	(0.12)	0.05	(0.09)	0.05	(0.10)	0.06	(0.09)
Perceived control	0.03	(0.06)	0.06	(0.09)	0.01	(0.06)	0.04	(0.09)		
Time use (in %)										
Watch TV daily	-11.06***	(3.26)	-7.84	(5.29)	-10.68***	(3.39)	-10.56**	(5.17)	-10.98***	(3.22)
Read books daily	5.59^{+}	(3.37)	5.80	(5.30)	4.45	(3.31)	6.39	(5.05)	5.41	(3.35)
Ambition (in %)										
Aim Abitur	4.23^{+}	(2.18)	-0.64	(3.56)	3.61^{+}	(2.10)	2.17	(3.19)	4.11^{+}	(2.14)
Aim university	8.25***	(2.97)	4.25	(4.75)	6.83**	(2.95)	5.69	(4.28)	8.12***	(3.01)
Job success likely	1.36	(1.11)	1.68	(1.72)	0.97	(1.12)	0.35	(1.62)	1.23	(1.11)
Desired profession likely	2.32^{+}	(1.26)	2.77	(1.94)	1.94	(1.28)	1.59	(1.78)	2.19^{+}	(1.27)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEPv29 (2001-2012), own calculations. Outcome differences between musically active and inactive adolescents. Same as baseline estimations (column 2 of Table 2.3), including additional control variables (intermediate variables) as indicated in the upper part of the table. Propensity score matching (radius matching with caliper 0.01) is used to account for all control variables from baseline specification, OLS to estimate outcome differences between treatment and matched control group with additional intermediate variables. The sample size is smaller for cognitive skills and some personality measures, as these have only been assessed since 2006. Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

of the outcome. Unfortunately, due to the design of our data, we can only measure the outcome variables once, at the age of 17.

Again, mediation analysis allows us to examine the robustness of our findings to successively including outcomes as control variables. For each combination of outcomes p and q, we estimate the following model:

$$Y_i^p = \alpha + \beta \cdot Music_i + \gamma Y_i^q + \varepsilon \qquad \text{for all } p, q \text{ with } p \neq q \tag{2.2}$$

where control observations (with $Music_i = 0$) receive weights obtained by propensity score matching, as described in Section 2.4. In other words, we examine whether the outcome difference in outcome Y^p between adolescents with and without music training, as estimated in the baseline model, changes once we control for outcome Y^q. We insert outcome Y^q, measured like all other outcomes at the age of 17, as a proxy for the value of Y^q at a younger age. As an example, we estimate the difference in cognitive skills at age 17 between adolescents with and without music training during their childhood, controlling for conscientiousness, also measured at age 17. All other control variables used in the baseline estimation are still accounted for by applying propensity score matching. Moreover, we include the predicted propensity to give up music as a further intermediate variable in each of these estimations.

Columns (2) to (5) of Table 2.4 provide the results for some of these mediation tests. Outcome differences between adolescents who learned a musical instrument and those who did not are very robust to including other outcomes as control variables. Even if this test is not able to perfectly exclude the risk of reverse causality, we conclude that the latter is highly unlikely to entirely explain our results. Estimations including the other outcomes as intermediate variables point in the same direction and can be provided by the authors on request.

2.7 Conclusion

The present study shows that even after controlling for a large number of parental background differences, learning a musical instrument is associated with better school grades as well as higher conscientiousness, openness, and ambition. Adolescents who have played music at age 17, have started to do so at age 8 or before and have taken music lessons outside of school obtain an average school grade of one sixth of a standard deviation above their peers. Adolescents who are enrolled in music lessons are more conscientious and open (about one fourth of a standard deviation). They are 11 percent less likely to watch TV daily and about 8 percent more likely to aim at completing upper secondary school and attending university. These results are stronger among adolescents with lower socio-economic status. All results are robust to variations in treatment definition, sample composition and estimation method.

Our analysis encounters two challenges, which could question a causal interpretation of our results. First, our results might be driven by unobserved heterogeneity. We argue that we are able to take into account the non-random decision to engage in music training at age 8 controlling for a large number of parental and individual background information. However, unobserved individual characteristics could still determine the decision to keep on playing music until age 17 rather than giving up earlier. Still, our results are robust to additionally controlling for the predicted probability to give up music before age 17, which we can predict within a subsample. Second, we examine the sensitivity of our results to reverse causality by performing mediation analysis in which we estimate the correlation between music practice and outcome p, while subsequently controlling for all outcomes q other than p. We observe that the pattern of correlation between music and cognitive as well as non-cognitive skills remains stable when we include any of the other outcome variables as controls. Reverse causality is therefore unlikely to entirely explain our results. Even though we cannot exclude the possibility that unobserved heterogeneity drives our results, we approach causality better than previous observational studies on the effects of music.

The strong effect on a variety of cognitive and non-cognitive skills indicates that extracurricular music is potentially an important input in the skill production function (Cunha and Heckman, 2008; Todd and Wolpin, 2003). More research should be carried out to confirm the findings of this paper. In our view, three challenges should determine the agenda of future research on this question. The most important is to further separate the influence of parental and individual background from the that of music. In order to do so, it would be necessary to identify a variable that increases the likelihood to learn a musical instrument without affecting the development of skills. Policy interventions and other variations in the regional availability of music lessons might be as "natural experiments" a promising way to carry out causal studies by providing a truly exogenous selection into playing music.

A second challenge will be to answer the question of the extent to which extracurricular activities are substitutable. Theoretical considerations, previous research, as well as the results of this study suggest that some types of skills might be improved through participation in extracurricular activities in general, while others are influenced particularly by music. These findings may be useful in informing policies similar to those described in the introduction that have been proposed to provide theater or sports lessons to children from disadvantaged social backgrounds. While policy makers have recognized the potential of such activities, there is still a lack of empirical research to support their implementation. Further research on the potential of different types of activities should be carried out by carefully modeling the interaction between activities that may be substitutes or complements.

Finally, further research should investigate the long-term effects of music training on outcomes such as labor market success or life satisfaction. It is possible that music has additional effects extending beyond educational achievement. Mechanisms such as signaling or an increased sense of determination might develop fully only at the entry into the labor market.

3 Mozart or Pelé? The effect of teenagers' participation in music and sports

3.1 Introduction

Music and sports are currently the most important education-oriented extracurricular activities of children in many developed countries. For example, 64% of Europeans aged between 15 and 24 regularly do sports (European Commission, 2014, 11), making this their most common leisure activity (Eurostat, 2009, 165). Among the 15- to 30-year-olds, 10% learn to play a musical instrument (European Commission, 2007, 19). Both activities also receive substantial public subsidies in many (European) countries. The literature on the effects of music and sports appears to agree that both affect child development positively compared to 'not playing them'. Unfortunately, these positive findings are unable to help a parent, child, or adolescent in deciding which of the two activities to select. Using individual-level data from Germany, we therefore investigate the differential effects of doing sports versus doing music.

In fact, a large literature studies the isolated effects of sports and music activities for schoolchildren. This research suggests that music improves cognitive skills (Winner et al., 2013), while sports improves educational achievement – as measured by school grades, secondary school completion, and university attendance, as well as labour market outcomes (Cabane and Clark, 2015; Lechner and Sari, 2015; Pfeifer and Cornelissen, 2010; Stevenson, 2010). Further tentative evidence states that music and sports may affect other outcomes such as non-cognitive skills as well (Felfe et al., 2011; Hille and Schupp, 2015). A common feature of these papers is that the explicit or implicit counter-factual, to which 'playing music' or 'doing sports' (or any aggregation of these) is compared, is vaguely defined as 'not doing the particular activity'. No study explicitly compares the effects of alternatively using the available time for doing music or doing sports (or another specific activity), although this question appears to be very relevant, given the time and budget constraints faced by parents and their children.

This is the research gap we address in this paper. We use the German Socio-Economic Panel (SOEP) (Wagner et al., 2007) to measure sports or music activities during childhood and adolescence. In our main estimations, we compare 17-year-olds doing only sports with those doing only music. We compare them with respect to educational achievements, cognitive and non-cognitive skills, subjective health measures, and measures of personality traits. To relate our results to the existing literature, we also compare playing neither or both of these activities to playing only one of them. In all comparisons, we focus on binary indicators describing participation in an activity rather than its intensity. Still, we vary the intensity of involvement to some extent by taking into account participation in extracurricular music lessons or sports competitions.

In order to improve the causal interpretability of these associations between sports or music and adolescent outcomes, we use a selection-on-observables approach and control for many relevant individual, parental, and family characteristics. This strategy exploits the fact that the SOEP is a longitudinal, very informative data set, which allows linking the records of children to those of their parents. Thus, we are able to control for a wide range of characteristics beyond those typically accounted for in selection-on-observables studies. In particular, we complement standard control variables by additional characteristics such as parents' personality and willingness to take risks. Selection on observables is also the identification strategy of most related studies so far.

In our setting of comparing activities with each other, identification based on selection on observables may be particularly compelling: Comparing the wealth of information in SOEP across different groups of participants and non-participants, we find that adolescents doing sports or music, as well as their parents and households, are similar. Thus, any differences that remain after conditioning on those variables are probably small as well. However, adolescents not doing these activities differ substantially in their characteristics from both active groups. One explanation could be that parental and individual characteristics primarily determine the decision to become active, which is costly in terms of time and money. By comparing only parents and children who already decided on being active, this substantial selection issue does not affect the results of this paper.

Some endogeneity might remain, despite the large number of parental and individual background characteristics we control for, if adolescents select into music or sports due to their innate ability or preference for one or the other activity, which might also affect some of the outcomes. This is problematic only to the extent that these individual reasons to choose sports or music are uncorrelated with the variables we control for. In addition to showing that non-random selection between music and sports is much less severe than between being active and not being active, we test the sensitivity of our results with respect to violations of the conditional independence assumption. We can show that playing music instead of sports does not affect a set of placebo outcomes, which should be unaffected by these activities, but are related to socio-economic status. Our results are also robust to a sensitivity test, as suggested by (Ichino et al., 2008), which simulates the existence of an additional influential confounder. Moreover, instrumental variables estimations do not contradict our results.

All estimations are based on semi-parametric matching estimators, which allow for flexible effect heterogeneity and avoid unnecessary functional form assumptions as much as possible. Here, allowing for individual effect heterogeneity appears to be particularly relevant, as we expect different children to react differently according to their parental background and other factors.

Our results indicate that the different activities have different effects. On the one hand, playing music fosters educational success more than sports, particularly so for girls. On the other hand, doing sports is more beneficial for health. The 'optimal' choice between music and sports therefore depends on parental and individual preferences, as well as the costs associated with each activity, which are on average somewhat higher for music lessons.

The paper is organized as follows: In the next section, we discuss the relevance of music and sports education for child development. The description of the data used is contained in Section 3.3, while Section 3.4 discusses the identification strategy and briefly reviews the estimation methods. Section 3.5 contains the results and Section 3.6 draws conclusions. Appendix B.1 contains more details on the data, while Appendix B.2 presents additional estimation results. Appendix B.3 details the estimator and its implementation. Finally, Appendix B.4 contains a formal sensitivity analysis for the chosen identification strategy, as well as results from an alternative identification strategy based on instrumental variables. Note that the full set of appendices is contained in the discussion paper version of this paper only.

3.2 How music and sports influence child development

A large literature studies the isolated effects of sports and music activities for schoolchildren. For example, many researchers in the fields of psychology and music education investigate the link between playing music and cognitive skill development. These empirical analyses find that music training is associated with a higher IQ (Vaughn and Winner, 2000), enhanced reading ability (Besson et al., 2007; Loui et al., 2011), increased attention (Shahin et al., 2008) and better memory (Ho et al., 2003). Although a majority of these studies controls for basic socio-demographic background characteristics, most of them do not consider non-random selection into playing music or taking music lessons as a major concern (Winner et al., 2013). One of the few researchers identifying causal effects of music lessons in an experimental study is Schellenberg (2004). He finds that children attending instrumental music or vocal lessons improve their cognitive ability compared to children receiving theater lessons and to children not attending such lessons in general. Further experimental studies confirm this finding for children (Bilhartz et al., 1999; Nering, 2002; Neville, 2008), but not for adults (Bialystok and DePape, 2009; Schellenberg and Moreno, 2010). While many proponents of music lessons claim that music also benefits the development of non-cognitive skills, so far there is no robust empirical evidence on such effects (Winner et al., 2013). Researchers in economics and sociology rarely studied the effects of music on child development. Southgate and Roscigno (2009) and Hille and Schupp (2015) are two exceptions in this respect. By accounting for selection on observables to a substantial extent, both studies find that music fosters the development of both cognitive and non-cognitive skills (compared to not playing music).

There is also substantial research on the influence of sports participation on skill development. Studies carried out with US data find that sports participation improves educational achievements (Anderson, 2001; Eide and Ronan, 2001; Lipscomb, 2007; Rees and Sabia, 2010; Stevenson, 2010). For Germany, Felfe et al. (2011) show that childhood sports activities improve school grades and non-cognitive skills. Moreover, participating in sports increases the probability to complete upper secondary school and to attend university (Pfeifer and Cornelissen, 2010). Furthermore, labor-market outcomes are positively affected by sports participation in the US and Canada (e.g. Cabane and Clark, 2015; Ewing, 2007; Lechner and Sari, 2015; Long and Caudill, 1991), but also in Europe (e.g. Lechner, 2009; Lechner and Downward, 2013). This literature conjectures that the main channels responsible for these effects are an increase in human capital (cognitive and non-cognitive skills) as well as health improvements. While the papers mentioned appeared in economics journals, other disciplines, such as epidemiology, sociology, and sport science investigate the impact of sports participation on children as well. In a recent survey reviewing research papers in these fields, Singh et al. (2012) find a positive relationship between physical activities and performance at school, for example.

There are also papers considering various leisure time activities including music and sports to study the role of extracurricular activities in general (e.g. Covay and Carbonaro, 2010; Feldman Farb and Matjasko, 2012). Typically, these studies compare individuals who are involved with at least one extracurricular activity compared to those who are not. Similarly, Del Boca et al. (2012) consider music and sports along with any other activity that "improves the child's human capital" as a substitutable form of child input. Obviously, one drawback of that approach is that the estimated effects are implicitly averaged over possibly heterogeneous activities.

Which differences should we expect between the influence of sports and music on skill development? Comparing existing research provides only limited hints about differences in their effects. The reason for that is threefold. First, outcome variables are usually not comparable between studies of extracurricular music and sports activities. While the music literature focuses on channels, measured by skills, the sports literature concentrates rather on outcomes such as educational and labor market success. In this paper, we compare the effect of both activities on a common set of outcomes. We analyze the differential impacts of sports and music on children's educational achievements, as well as on cognitive skills, non-cognitive skills, and health.

Second, a direct comparison between the results found in studies on each activity might be misleading, because participants in sports and music activities are not necessarily comparable. The question is whether different effects are due to differences in the activity or differences in the composition of adolescents enrolled. Indeed, the decision to take up musical or athletic activities is not random. For example, some adolescents might learn team capabilities better by playing football rather than by playing violin (in an orchestra) because of their different socio-economic background. Moreover, as Section 3.3 will show, boys are more likely to play team sports and girls are more likely to play a musical instrument. Therefore, girls might be more likely to learn team capabilities when playing in an orchestra, while boys might acquire the same skills while playing football.

Third, the composition of the control group differs between studies on music and sports due to lacking information on the counter-factual use of time. Implicitly, these studies compare those active in music or sports to all other individuals, irrespective of how they spend their time. Therefore, the control group in studies on sports participation includes people who play music, and vice versa. Hence, comparing the estimated effects would require information on the alternative use of time.

While previous research on causal effects of music and sports can only provide us with limited guidance on expected effect differences, we can derive hypotheses from theoretical considerations. Music and sports may affect child development through a variety of channels. Some of these may be similarly activated by both activities. Schellenberg (2011) discusses mechanisms according to which music training influences sub-domains of cognitive functioning, which leads to improvements in cognitive skills. Singh et al. (2012) list three potential physiological effects that link physical activity to benefits in cognition. The first one is an increase in blood and oxygen flow to the brain. The second is a reduction of stress and a mood improvement due to increasing levels of norepinephrine and endorphins. Finally, sports could lead to an increase in growth factors related to new nerve cells creation and to support of synaptic plasticity. Thus, both music and sports are likely to affect cognitive skills.

In addition, as a potential link between music or sports and non-cognitive skills, Schumacher (2009) highlights the potential of these activities to teach individuals judge their own abilities and development. The ability to perform may improve attitudes towards school (Eccles et al., 2003), or at least raise awareness that hard work leads to success (Winner et al., 2013). Moreover, playing music or sports after school usually involves close interactions with teachers and peers. Thereby, adolescents engaged in these activities might acquire social skills as well as cultural capital (Lareau, 2011; Schumacher, 2009). Finally, even in absence of a causal link between music or sports and skill development, being engaged in these activities might send a positive or negative signal to schoolteachers

(Lareau, 2011). Thus, the adolescent could receive school grades that deviate from their actual level of competence due to extracurricular engagement.

Since no study has directly compared music and sports, we do not yet understand whether music and sports activate these mechanisms to the same extent. For example, sports are more often carried out in teams (see Section 3.3). Therefore, social skills might be acquired more effectively by doing sports than playing music. In contrast, musically active adolescents are more likely to receive instruction from a teacher than sports participants, although trainings supervised by coaches are also frequent in adolescent sports. In particular, musically active adolescents are more likely to take part in lessons alone or in a small group, similar to non-team or small team sports. Hence, the particular skills acquired through the close interaction with a teacher might be developed more effectively among music participants.

Finally, the acquisition of some capabilities is specific to either music or sports. For example, playing sports improves the level of fitness and leads to a better health status (e.g. Felfe et al., 2011). Consequently, athletically active adolescents might be better able to concentrate, learn skills and succeed in school. By contrast, the occupation with music increases musical self-efficacy (Ritchie and Williamon, 2011) and induces the formation of a musical self-concept (Spychiger et al., 2009). Few studies focus on these activity-specific channels. Rather, the findings of previous research are consistent with both the activity specific and the general channels presented above. Nonetheless, in most cases, these results are misleadingly presented as being activity-specific.

3.3 Data

3.3.1 The German Socioeconomic Panel

This paper uses data from the German Socio-Economic Panel (SOEP). The SOEP is a representative sample of German households interviewed every year since 1984. Interview questions cover numerous aspects of life, including socio-demographic background, work, housing, as well as opinions and attitudes. The number of households interviewed has been increasing constantly and has attained almost 15,000 in 2014 (for a detailed description of the SOEP, see Wagner et al., 2007). In the 2000s, the SOEP developed a particular focus on children, which led to the creation of several specialized questionnaires addressing child development. One of these, the youth questionnaire, constitutes the main data source for this paper. Since 2000, all adolescents in SOEP households receive specific survey questions in the year they turn 17 years of age. In their answers to the youth questionnaire, they share their educational achievements and plans, their objectives for career and private life, their relation to the parents, as well as opinions and personality traits (see Weinhardt and Schupp, 2011, for details on this questionnaire).

Most relevant for this study, the youth questionnaire contains information on the participation in extracurricular sports and music activities (see Table 3.1). For each activity, individuals answer five questions describing their involvement during youth. In addition to information about participation in music and sports, the youth questionnaire contains questions related to educational achievement and plans, as well as cognitive skills, non-cognitive skills, and opinions. This information is used to create outcome variables for the empirical analysis below. In particular, educational success is measured using adolescents' school types and their plans to attend university. Moreover, adolescents provide their latest school grades in the subjects of mathematics, German and foreign language.

Cognitive skills were measured with standardized tests in three categories: word analogies, figures, and mathematics operators (Schupp and Herrmann, 2009). In addition to cognitive skills, the SOEP youth questionnaire measures a variety of non-cognitive skills. These include the Big Five personality traits: conscientiousness, openness, extraversion, neuroticism and agreeableness (McCrae and Costa, 1999; Lang et al., 2011). Moreover, we have information on the individual's willingness to take risks. More than cognitive skills, non-cognitive skills are considered malleable during childhood and youth (Cobb-Clark and Schurer, 2012; Donnellan and Lucas, 2008; Heckman and Kautz, 2012; Specht et al., 2011). Some health measures, as well as a variable indicating university attendance, stem from the individual questionnaire, which survey participants answer for the first time when they are 18 years old.

In addition to this substantial amount of information, the household structure of the SOEP allows us to observe many socio-demographic background characteristics of family members. These characteristics potentially determine the selection between the relevant activities. We obtain this information by combining the data on 17-year-olds with information from the household and individual questionnaires of the adolescent's parents and

Table 3.1 – *Music and sports in the SOEP youth questionnaire*

Question in SOEP youth questionnaire	**Possible answers**
MUSIC	
15. How often do you carry out the following leisure time activity: ...making music	Daily, weekly, monthly, less often, never
16. Do you play a musical instrument or pursue singing seriously?	Yes, no
If 'yes' in question 16:	
17. What type of music do you play?	Classical, pop/rock/techno/funk/ rap/hip-hop, other folk or light music
18. Do you do this alone or in a group?	Only alone or with a teacher, in an orchestra or choir, in a music group/band, in other group
19. At which age did you start?	Free text
20. Do you have music lessons outside of school?	Yes, no
SPORTS	
15. How often do you carry out the following leisure time activity: ...doing sports	Daily, weekly, monthly, less often, never
21. Do you play sports?	Yes, no
If 'yes' in question 21:	
22. Which is the most important type of sports you play?	Free text
23. Where do you play this sport?	Sports club, commercial provider (e.g. fitness studio), at school, other organization (e.g. adult education center), with others (but not in an organization), alone
24. At which age did you start this sport?	Free text
25. Do you regularly participate in sports competitions?	Yes, no

Source: SOEP v29, 2001-2012.

their household. Among others, these questionnaires contain information about parental education and income, as well as household composition and the parents' personality.

3.3.2 Selection of the estimation sample

Our sample used in the empirical analysis consists of all adolescents who answered the SOEP youth questionnaire between 2001 and 2012. We drop less than 1% of the observations, for whom we have no answers to the questions on athletic and musical involvement during their youth. The data from the youth questionnaire is merged with data on the adolescent's household and parents. This leads to a sample of 3,835 individuals, consisting of 12 cohorts of 17 year-olds (see Table B.2 for further details on the selection of the sample).

3.3.3 Definition of 'doing sports' and 'doing music'

Using the answers to the questions about musical and athletic involvement described in Table 3.1, we define being musically or athletically active in the following way (for details see also Table B.1). We consider adolescents as musically active if they match the following criteria: i) they state playing a musical instrument when answering the youth questionnaire (at age 16 or 17), and ii) they have started to play music at the age of 14 or earlier. With the latter, we aim to capture a somewhat substantial exposure to the activity (3 years at least). As robustness checks (see Section 3.5), we also consider more restrictive activity definitions that require following paid music lessons outside of school or playing at least on a monthly basis. This definition applies to 80% of those who play music. For sports, we define two levels of sports intensity. For both levels, we require the adolescents to be active in sports at age 16 or 17 (i.e. when answering the youth questionnaire) and to have started their main sports no later than at age 14. With the aim to capture another, more structured and intense dimension of sports, we separately examine those who regularly take part in sports competitions, in addition to having started to play sports at or before age 14 and still doing so at age 17.

Unfortunately, the data do not contain information on former music or sports participation for adolescents who gave up the respective activity before answering the SOEP youth questionnaire at age 17. However, for a subsample of 752 individuals in the latest cohorts of survey respondents, we can reconstruct the history of extracurricular music and sports participation since age 12 or 14. To do so, we use information from the SOEP household questionnaire, in which parents state their children's extracurricular activities biannually since 2006. Section 3.5 presents the results of this analysis.

With the definitions of music and sports participation described above, we are able to define the different groups and resulting comparisons as detailed in Table 3.2 In a first step, both activities are compared to each other. However, in particular sports participation may be heterogeneous with respect to the formality and intensity of involvement. Therefore, a second estimation compares adolescents active in music with those who regularly take part in sports competitions. As shown below, this more or less equalizes the effort and time investment with which the activities are carried out. In the empirical results, it turned out that the two different types of sport activities led to few differences in the results.

Table 3.2 – *Comparisons of different activities*

	'Active' group	'Comparison group'
Main comparison – Music vs. sports: playing music only versus playing sports only		
Play music	Yes	No
Play sports (competitively and non-competitively)	No	Yes
Comparison 2 – music vs. sports (competitive): music only versus sports at a high intensity level only		
Play music	Yes	No
Play sports (only competitively)	No	Yes
Comparison 3 – both vs. sports: playing music and sports versus playing sports only		
Play music	Yes	No
Play sports (competitively and non-competitively)	Yes	Yes
Comparison 4 – both vs. music: playing music and sports versus playing music only		
Play music	Yes	Yes
Play sports (competitively and non-competitively)	Yes	No
Comparison 5 – sports vs. none: playing sports only versus playing neither sports nor music		
Play music	No	No
Play sports (competitively and non-competitively)	Yes	No
Comparison 6 – music vs. none: playing music only versus playing neither sports nor music		
Play music	Yes	No
Play sports (competitively and non-competitively)	No	No
Comparison 7 – both vs. none: playing music and sports versus playing neither sports nor music		
Play music	Yes	No
Play sports (competitively and non-competitively)	Yes	No

Note: For exact definitions of the activities, see Table B.1.

Therefore, the results concerning sports with competition are referred to Appendix B.2. To complete the picture, further comparisons contrast adolescents who engage in both activities, as well as those who do neither music nor sports with those who do only one of the two.

3.3.4 Descriptive statistics

In Germany, non-formal extracurricular activities have a long tradition and are organized predominantly in some type of formal organization. Concerning music, there exists a countrywide network of 929 public and numerous private music schools (VDM, 2014). For sports, there are 91,000 non-commercial sports clubs, which are organized in a well-structured network and cover the entire country (Deutsche Sportjugend, 2014).

Doing sports and exercise is the most popular leisure time activity in Germany. 80% of girls and 90% of boys do sports when they are children, a share that decreases slightly until early adulthood (Grgic and Züchner, 2013). Music is also an important extracurricular activity. According to the German Youth Institute, 25% of all Germans of age 18 to 24 played a musical instrument in 2012. At a younger age, music is even more prevalent. In the same year, music was played by 36% of those aged 13 to 17, and by 44% of those aged 9 to 12 (Grgic and Züchner, 2013).

Youth music and sports activities have steadily increased in the last ten years. While only 10% of German adolescents attended paid music lessons outside of school in 2001, this number rose to 18% in 2012. Similarly, for sports, the share of adolescents of the same age, who regularly took part in sports competitions, has risen from 29% to 34% (Hille et al., 2014). The engagement with music and sports remained constant after most federal states increased the number of hours students must spend in upper secondary school in a recent school reform (Dahmann and Anger, 2014).

Table 3.3 describes the characteristics of music and sports activities in our estimation sample. The table indicates details such as the average starting age or the share of individuals carrying out their activity in a team or in a non-formal context for the subgroups of adolescents who play music only, sports only, both or none of the two activities.

Playing music is associated with a more thorough commitment than sports, both in terms of time and formal instructions. About 40% of the musically active state that they play every day, almost 80% take lessons. For sports – competitive and non-competitive combined – these numbers are only 29% and 64%. By contrast, adolescents who regularly take part in sports competitions engage at a similar level as those who play music. Around 80% of them take part in formal instructions, a share similar to those who play music. Furthermore, the frequency, at which the respective activity is carried out, is comparable between music and competitive sports participants.

One would think that engaging in both music and sports reduces the level of commitment an adolescent is willing to dedicate to each of these activities. However, as described in column (5) of Table 3.3, adolescents active in both activities are just as likely to play music and sports daily or to take part in formal lessons as those who only participate in one activity. Thus, for them, music and sports seem to be rather complements than substitutes.

Some adolescents, who are neither active in music nor sports (according to our definition), still play these activities to some extent, as shown in column (1). Around 6% or 7% state to do music or sports every day, and even 16% receive sports instructions. These

Table 3.3 – *Characteristics of musical and athletic activities*

	Neither music nor sports		**Music only**	**Sports only (competitive and non-competitive)**	**Competitive sports only**	**Music + sports (competitive and non-competitive)**	
	Music	*Sports*				*Music*	*Sports*
	(1)		(2)	(3)	(4)	(5)	
Average starting age			8.8	9.1	8.6	8.6	9.1
Share (%)...							
...doing sports/music in team	2	11	53	63	79	52	59
...with lessons/instructions	1	16	79	72	93	78	76
...playing sports/music daily	7	6	40	29	37	38	29
Number of adolescents	1374		333	1640	884	501	

Note: SOEP v29, 2001-2012, own calculations. Column (4) is a subgroup of column (3). Otherwise, each column contains distinct individuals, who carry out either nothing (1), music (2), sports (3 with subgroup 4) or both (5). Given that playing music or sports according to our definition involves being active since age 14 or longer, some inactive adolescents played music or sports at age 17 as well. This explains positive percentages in column (1). Table B.1 shows the exact definitions of sports and music participation. Separate tables by gender can be found in the appendix (Table B.5).

adolescents are not in our treatment groups, because they start to become active at age 15 or 16, only shortly before answering the SOEP youth questionnaire.

In addition to differences related to the level of intensity, adolescents who play sports start to become active at a later age and are more likely to play in a team than those who play music. Team sports are particularly prevalent among boys and among those who regularly participate at competitions. While only slightly more than half of the musically active adolescents play music with other people (in an orchestra or a band), almost 80% of those who do sports competitively take part in team sports.

As expected, adolescents who engage with music differ with respect to background characteristics from inactive individuals, as well as from those who engage with sports. Table 3.4 illustrates these differences. The share of girls is higher among adolescents who play music or both activities than among those who only play sports. Moreover, adolescents who play sports usually have less advantaged social backgrounds than musically active youth, but more advantaged backgrounds than individuals who play neither sports nor music. For example, in 47% of the families of musically active adolescents, at least one parent has a university degree, while this is true only for 31% of the athletically active families and for 19% of the parents among inactive individuals. Approximately 24% of

Table 3.4 – *Selected descriptive statistics by treatment status*

	Not active	One activity: *Sports (competitive and non-competitive)*	One activity: *Sports (competitive only)*	One activity: *Music*	Both activities
Female (%)	55	41	35	59	56
Recommendation for upper secondary school (%)	25	38	44	55	68
Single parent household (%)	26	22	22	17	17
Migration background (parent, %)	24	20	18	14	17
Parent with university degree (%)	19	31	35	47	51
Av. monthly labor market income of parents present in household (€)	1181	1443	1512	1681	1920
Household net overall wealth (in 100,000 €)	1.46	2.05	2.39	2.64	3.02
ISEI socio-economic status (highest among parents, scale from 0 to 90)	39	45	46	52	54
Household lives in rural area (%)	28	24	24	29	22
Hours working (mother)	21	22	23	23	21
Willingness to take risks (mother, 0 to 10)	4.0	4.2	4.3	4.1	4.0
Conscientiousness (mother, 0 to 1)	0.87	0.87	0.87	0.86	0.85
Extraversion (mother, 0 to 1)	0.72	0.72	0.72	0.72	0.71
Neuroticism (mother, 0 to 1)	0.59	0.59	0.60	0.59	0.58
Openness (mother, 0 to 1)	0.65	0.65	0.65	0.68	0.67
Agreeableness (mother, 0 to 1)	0.80	0.81	0.81	0.80	0.80

Note: SOEP v29, 2001-2012, own calculations. The exact definitions of each activity are given in Table A.1.2. Tables presenting statistics for all covariates and outcomes can be found in the appendix (Tables B.3 and B.4).

inactive adolescents have a parent with migration background. This share is considerably lower especially among music participants. Similarly, parental labor market income is almost twice as high among adolescents who play both music and sports than among inactive adolescents.

Table 3.4 also shows that differences in individual and parental background characteristics are much smaller between music and sports participants than between active and inactive adolescents. In particular, more competitively active adolescents come from a more favorable social background than those who are less involved. Individuals, who are engaged in sports at a high intensity level (who regularly take part in sports competitions), are more similar to those who play music than those who do sports at a lower level of intensity. Most remarkably, adolescents active in both music and sports have wealthier and more educated parents than all other groups of adolescents.

With respect to parental non-cognitive skills, adolescents who play music, sports or none of the two are rather similar. The Big Five personality traits of mothers and fathers are almost identical along all activity types. The only notable difference concerns the parental willingness to take risks: it is higher among parents whose children engage in competitive sports.

3.4 Econometrics

3.4.1 Identification

This paper investigates the differential impact of practicing sports or music during childhood on the development of various skills, health, and personality traits. The challenge for identifying causal differences is the fact that the decision to play music rather than doing sports is not random. Thus, we have to disentangle the effect of participating in a specific activity from the influence of differences in socio-demographic and other background characteristics, given participation in at least one of those activities. In other words, we need to take into account selection effects involved in comparing adolescents carrying out one or the other activity. Since we are not aware of any specific exogenous institutional features useful for identification (like in a difference-in-differences or regression discontinuity design), the potential identification strategies are based on either assuming selection on observables or on instrumental variables type of assumptions. We discuss both options in turn.

As described above, SOEP contains rich socio-demographic information on adolescents and their parents. Therefore, selection on observables is an attractive identification strategy. The main requirement is that we are able to control for all variables affecting treatment (sports or music participation) and outcomes simultaneously, given participation in at least one of the activities. If we assume that the decision is a two-stage process, in which the first stage is a decision to be active or not, and the second stage is about the type of activity, the potentially selective decision to take up music or sports practice rather than not being active plays no role for identifying the effects of such comparisons. Even if the decision to become active and the choice of activity are made simultaneously (i.e. the latter choice is not made conditional on being active), music and sports participants are much more similar in terms of observed and unobserved characteristics than active and inactive individuals are.

To justify selection on observables as a strategy to identify causal effects, we need to discuss the determinants of choosing between music and sports, or even carrying out both at once. The choice of engaging with either activity can be motivated by taste, as well as expected costs and gains. Moreover, there might be constraints hindering a child to pursue the activity of interest. Finally, we only observe adolescents who play music or sports at age 17. Since we consider adolescents as active if they started their activity at age 14 or earlier, being in the 'treatment group' also depends on the adolescent's willingness to pursue the activity until age 17. In the following, we discuss the driving factors for each of these dimensions, as well as how they might differ between music and sports. Moreover, we explain how our identification strategy considers these potential confounders.

The choice to start playing music or/and sports can be considered as a common decision by child and parents. The decision most likely depends on utility gains and taste (Hille and Schupp, 2015). In addition to current pleasure, the former includes the desire to invest in the child's future skills (Eide and Ronan, 2001; Lareau, 2011). A priori, neither parents nor children can judge which activity more effectively provides pleasure or stimulates skill development. Therefore, they can only form their opinion according to their own experience and taste, as well as the information obtained from others. These factors are approximated by socio-economic status (Garboua and Montmarquette, 1996).

Even more than sports, highbrow cultural activities such as playing music can be used by individuals from higher social classes as a costly signal to assert status (Menninghaus, 2011; Ormel et al., 1999). A similar argument could justify the desire of parents to enroll their child for two activities rather than one. Similarly, adolescents from particular socio-economic groups may be more likely to play music, because the offer of artistic activities for children might be adapted to the tastes of the more highly educated (Lunn and Kelly, 2009).

To consider these determinants, we control for the parents' socio-economic background, which reflects both their eagerness to invest in the child's future skills, as well as their taste. We hold socio-economic background constant by controlling for the parents' level of education and income, as well as detailed information about the job they carried out when the adolescent was 17 years old. The latter includes both parents' work hours, their sector of activity, whether their job required training, as well as their socio-economic status. In addition to these objective indicators, parenting style is likely to affect the motivation to invest in the child's skill development. We proxy parenting style by controlling for the Big Five personality traits of mother and father, as well as an indicator of their willingness to take risks. Finally, since ethnicity is an important determinant of socio-economic status and taste, we include the parents' migration background.

Even though willing to enroll their child at the local music school or sports club, parents might face constraints hindering them to do so. Gustafson and Rhodes (2006) point out that extracurricular activities require parental support, both financially and logistically. Parents might face financial constraints if they cannot or do not want to afford the costs for lessons, club membership, as well as musical instruments or sports equipment. Moreover, parents might not be able to provide the necessary logistic support, especially if they have other children or work full-time (Lareau, 2011). In addition, the child's position in the birth order plays an important role for the educational investments made by parents (Black et al., 2005).

For most parents in Germany, financial constraints are unlikely to be the crucial determinant of the decision to sign up at a music school or sports club. Among families with children, 57% report that they have regular expenses for non-formal educational activities. On average, these families spend 51 Euros per month for the athletic and musical activities of their children (Schroeder et al., 2015). Concerning expenditures for music, fees vary strongly between music schools and account for 47% of the overall budget of German public music schools in 2013 (MIZ, 2015a). However, the German association of public music schools (*Verband deutscher Musikschulen*) stipulates that their members should provide reduced fees for children from poor families (VDM, 2011). In contrast, membership fees in sports clubs are very low on average. For adolescents, the median membership fee was just 3.10 Euros per month in 2013 (Breuer and Feiler, 2015). In addition, the *Educational package* – a policy introduced by the German federal government in 2011 – subsidizes membership fees in sports clubs or music schools for poor families in the amount of 10 Euros per month (BMAS, 2015a).

The extent to which the above-mentioned constraints vary between music and sports depends on factors such as household income, the number of children in the household, and birth order, but also on the institutional context in which the activity is carried out. We control directly for the first three factors, whereas the latter cannot be observed in our data. Depending on living area and time, as well as the organizational structure of the local music school or sports club, material investments might be more substantial for music or for sports. We approximate these factors by including state and birth year fixed effects, as well as an indicator of whether the household lives in a rural area.

To be observed as playing sports or music, adolescents in our sample need to pursue these activities until they answer the SOEP youth questionnaire at age 17. In addition to motivation and constraints, the willingness not to give up before age 17 therefore constitutes an important dimension in the non-random selection into one of the activity groups. Whether an individual gives up music or sports during adolescence is likely to depend on encouragement by the parents, as well as own motivation, the opportunity costs of time, and the perceived returns in general. According to Farrell and Shields (2002) and Raudsepp (2006), parents as role models can motivate children to play music and sports, and might determine whether they pursue these activities until age 17. Whether parental encouragement leads the child to prefer music to sports probably depends on parental taste, which we control for as explained above.

Moreover, carrying out music or sports throughout adolescence requires motivation and time. Adolescents who struggle at school are unlikely to be able to dedicate time to these activities in their leisure time. The individual characteristics determining whether adolescents can pursue their extracurricular activities are difficult to take into account, given that we observe only few individual characteristics before the age of 17. To approximate these characteristics, we control for gender and an indicator whether the adolescent has received a recommendation for upper secondary school by her primary school teacher (in grade 4, at age 9 to 10). The latter serves as a proxy for ability. Furthermore, if the remaining unobserved reasons for giving up sports and music before the age of 17 are similar, this does not pose a threat to identification.

Several reviews investigate the determinants and correlates of physical activity among children and adolescents in empirical studies (Craggs et al., 2011; Van Der Horst et al., 2007; Sallis et al., 2000). They underline two points: i) the determinants vary between children (aged 4 to 12 years old) and adolescents (aged 13 to 18 years old), and ii) the identification of the determinants of a change in physical activity levels often suffers from reverse causality. Our study focuses on individuals who decided to engage or to quit sport while being children and then adolescents. Therefore, we are interested in determinants of sports and of a change of the sports activity for both age categories. Previous findings on these determinants are consistent with the arguments described above. Light (2010) suggests that friendship is an important determinant of the child's decision to remain in a competitive club. We cannot control for such individual characteristics because of the risk of reverse causality. Moreover, we believe that these unobserved factors are similar for music and sports.

Theoretically, we could imagine children choosing a particular activity due to their innate ability, which gives them a comparative advantage in either sports or music and determines some of the outcome differences we find. For example, children who are good at languages or mathematics might prefer music to sports. They might also obtain better German or mathematics grades. A positive coefficient would therefore reflect innate differences and rather than a consequence of doing music instead of sports.

We argue that after conditioning on a wide range of standard and non-standard individual and family characteristics, musically and athletically active adolescents are rather similar in terms of observed, but also unobserved characteristics. This is even more plausible, given that music and sports participants are much more (observationally) similar than active and inactive individuals.

Further tests examine the sensitivity of our results to the conditional independence assumption. First, we estimate the effect of playing music rather than sports on a set of outcomes, which should be affected by neither of these activities but are correlated with socio-economic status. These outcomes capture whether adolescents have their own room at home, the amount of pocket money they receive each week, and whether the share of foreign students in their school class is larger than 25%. These variables were measured at age 17. Second, we conduct the test suggested by Ichino et al. (2008), which examines the sensitivity of our results with respect to the inclusion of a strong additional confounder. Third, for a small subsample with information on the history of music and sports participation since age 12 or 14, we compare the characteristics of musically and athletically active individuals with those who gave up the respective activity before age 17. This test allows us to verify whether the potential existence of adolescents, who played an activity earlier in life, in the control group affects our results.

An alternative identification strategy is instrumental variables (IV). A crucial characteristic of a valid instrument is that it affects outcomes only through its effect on music and/or sports participation (the so-called exclusion condition). Indeed, some of the commonly used instruments (e.g. height or school characteristics) may be questionable in this respect. Therefore, we use an alternative instrument, namely parental artistic activities, which arguably fulfills the requirements for an IV. They are a strong predictor for adolescents' music participation and – conditioning on further covariates – unlikely to influence the outcomes directly. However, this instrument is not available for the entire sample, which implies a further reduction of the number of observations, which in turn drastically reduces the precision of our estimates. Therefore, we will use this IV as a robustness check only.

3.4.2 Estimation

Our study compares the effects of different activities with each other. This corresponds to the multiple treatment setting discussed in Imbens (2000) and Lechner (2001). Therefore, the sample reduction results of Lechner (2001) apply. Thus, to estimate the effects of one activity compared to another by a selection on observables approach, participants in activity groups other than those two under explicit consideration are deleted for the purpose of this particular estimation. For example, for the estimation of the effect of being active in sport compared to music, individuals who do both activities play no role.

An alternative to this procedure of splitting the multinomial treatment problem into several pairwise comparisons is to jointly model the selection problem by a multinomial logit or probit model. However, as shown by Lechner (2001), such an approach is more restrictive because the estimation of a (tractable) multinomial choice will always require more assumptions than flexibly modeling pairwise comparisons. However, with respect to the assumptions needed for non-parametric identification, both approaches are equivalent. For performing the pairwise comparisons, we use the propensity-score-matching estimator proposed by Lechner, Miquel and Wunsch (2011). This estimator performed well in the large-scale simulation study by Huber et al. (2013). It is described in detail in Appendix C. Such semi-parametric estimators are based on estimating a parametric model (e.g. probit) for the probability of belonging to one of the groups compared to another, conditional on the above-mentioned control variables. The relation between the outcomes, activity types, and confounders, however, are left unspecified (non-parametric). Therefore, such estimators have the advantage that they allow for very flexible effect heterogeneity (contrary to regression models, for example).

For each outcome variable, having four groups of activity (music, sports, competitive sports and both) leads to up to 6 different estimates for each treatment effect, like the average treatment effect (ATE), the average treatment effect on the treated (ATET) and the average treatment effect on the non-treated (ATENT). However, an implication of treating the estimation problem as many single pair-wise problems is that the reference population for these effects is specific to each single comparison. For example, the standard ATE refers to the union of the respective treatment and control groups in the particular comparison. However, the characteristics of those groups will not be the same for different comparisons. Therefore, we follow the approach advocated and implemented by Lechner and Wunsch (2009) in a similar situation to keep the 'target' or 'reference' distribution the same for the different comparisons. In this case, there is an additional matching step. Matching is performed such that the implied weighting scheme leads to matched covariate distributions of treated and controls in all comparisons that resemble those in the 'target' population. Thus, the various estimation results presented below always refer to the same population (and thus the same distribution of confounding characteristics) and thus are comparable in that sense. In our case, a natural candidate for the target population is the union of sports and music participants.

3.5 Results

3.5.1 Propensity scores

We investigate the differential effects of music and sports in this paper, not only by examining the consequences of playing music instead of doing sports, but also by comparing these effects to outcomes of adolescents active in both, as well as those who play neither activity. All estimations were carried out separately for each possible comparison between the groups of adolescents playing nothing, music only, sports only, and both activities. Detailed definitions of the treatment and control group for each comparison are given in Table 3.2. While our identification strategy is most credible with respect to the direct comparison of musically and athletically active youth, these additional comparisons provide us with further insights.

Table 3.5 shows the average marginal effects of the estimation of the propensity score for three of these comparisons for selected covariates (for the other comparisons and covariates, the propensity score estimation is referred to Appendix B.2). Generally, the results confirm the (unconditional) summary statistics of Section 3.3. Adolescents who engage with music tend to come from more advantaged social backgrounds. Still, parental education plays a much smaller role in choosing between sports and music than in the general decision to

Table 3.5 – *Selected results of propensity score estimation*

	Music only vs. sport only (1)		**Music + sport vs. sport only** (2)		**Music + sport vs. music only** (3)	
	Marginal effects	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*
Female	0.106***	0	0.096***	0	-0.035	31
Recommendation for upper secondary school	0.044**	2	0.148***	0	0.141***	0
Birth order	-0.033***	1	-0.040***	0	0.004	88
Number of siblings in SOEP	0.023***	1	0.032***	0	0.011	49
Household lives in rural area	0.039*	10	0.011	65	-0.055	23
Monthly labour market income of parents who live in household	-0.001	88	0.005	61	0.017	33
Working hours (mother)	-0.001	22	-0.002**	2	-0.001	59
Working hours (father)	0.001	13	0.002**	3	0.001	71
Willingness to take risks (mother)	-0.008*	7	-0.008*	7	0.005	61
Willingness to take risks (father)	-0.011**	2	-0.011**	2	0.002	80
At least one parent with						
...university degree	0.004	90	0.048*	7	0.035	54
...upper secondary schooling degree	0.039	14	0.033	21	-0.038	47
...migration background	-0.023	36	0.027	29	0.078	13
Big Five (mother)						
- Agreeableness	-0.162**	3	-0.129*	9	0.130	45
- Conscientiousness	0.010	91	-0.092	30	-0.243	19
- Extraversion	-0.081	22	-0.088	24	-0.041	78
- Neuroticism	-0.025	67	-0.042	48	-0.058	62
- Openness	0.142**	2	0.120*	8	-0.065	63
Big Five (Father)						
- Agreeableness	0.079	32	0.124	12	0.054	71
- Conscientiousness	0.048	57	0.120	18	0.179	35
- Extraversion	-0.072	30	-0.077	30	0.059	68
- Neuroticism	0.157***	1	0.087	15	-0.099	44
- Openness	0.141*	6	0.172**	3	-0.034	82
Efron's R2 (in %)	11		16		7	
Number of observations	1973		2141		834	

Note: SOEP v29, 2001-2012, own calculations. Probit model estimated. Average marginal effects presented. Inference is based on 4999 bootstrap replications. Significance levels: * 10% ** 5% *** 1% level. The exact definitions of each activity (dependent variable) are given in Tables 3.2 and B.1. The following covariates are also included in this specification: Constant term, year of birth, single parent household, dummies for the federal states, household net overall wealth, age of mother at birth, indicators for mothers working in services, fathers working in services, mothers working in manufacturing or agriculture, fathers working in manufacturing or agriculture, at least one parent with vocational degree, ISEI socio-economic status ('higher' for both parents), current job did not require training (mother), current job did not require training (father), indicators for missing values in willingness to take risks (mother), and willingness to take risks (father). The full results are given in Appendix B (Tables B.6 and B.7).

become musically or athletically active. In particular, parental income is not statistically significant. Other parental characteristics play only a minor role. For example, a higher willingness to take risks among parents is associated with a lower probability that their child plays music rather than sports. The strong effect of openness both for mothers and for fathers can be explained with the fact that openness towards artistic experiences is used as one of the items assessing this personality dimension.

Two important predictors of playing music rather than sports are the adolescent's gender and whether she or he received a recommendation for upper secondary school at the end of primary school. Girls are 10% more likely to play music instead of or in addition to sports than boys are, all other covariates held constant. As a proxy for prior abilities, the recommendation for upper secondary school is especially important in determining who carries out two activities rather than one. The probability to play music and sports rather than just one of the two is 14% higher among adolescents who received such a recommendation.

Using the coefficients from Table 3.5, we predict the probability for each individual to belong to the respective activity group. In order to eliminate differences in covariate distributions between the groups, we match both groups on the propensity score and on gender (to fully remove all gender differences that seem to be particularly important in this application), as described in Section 3.4 After matching, all covariates are balanced (see Table B.7). The cut-off for the common support with respect to the target distribution is the 99% quantile of the respective propensity score distribution. Different cut-off values are considered as robustness checks (more details are given below).

3.5.2 Effects of music compared to sports

Table 3.6 shows the average effects of playing music rather than sports, as well as the comparison of each activity to playing both. Table 3.7 also contains the results of comparing inactive adolescents to those playing music and/or sports. All results are re-weighted with respect to the characteristics of adolescents who are active in music or sports or both, as described in Section 3.4. Appendix B.2 contains the results for an additional comparison of music versus competitive sports, a subgroup of sports participants that is even more similar to music participants in terms of observed (and unobserved) background characteristics (see Section 3.3). These results turned out to be very similar to the ones that

Table 3.6 – *Average effects of music vs. sports*

	Music only vs. sport only (1)		Music + sport vs. sport only (2)		Music + sport vs. music only (3)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17						
Cognitive skills - Average	0.012	90	0.233***	1	0.230**	5
- Analogies	0.054	58	0.216***	1	0.169	15
- Figures	0.053	59	0.276***	0	0.229**	5
- Math	-0.045	63	0.070	45	0.123	30
School grades - Average	0.034	66	0.151**	3	0.123	20
- Maths	-0.119	14	0.113	12	0.233**	2
- German	0.172**	2	0.113	11	-0.051	59
- 1st for. lang.	0.063	42	0.103	13	0.046	63
Attends upper secondary school	0.081**	4	0.135***	0	0.063	18
Aim to enroll at university	0.106***	1	0.187***	0	0.087*	6
Attends university at age 20	0.024	48	0.084***	1	0.065	13
Personality at age 17						
Big 5: Conscientiousness	-0.114	21	0.059	51	0.165	14
Extraversion	-0.165*	9	0.128	13	0.294***	1
Neuroticism	0.103	31	0.013	88	-0.090	47
Openness	0.407***	0	0.349***	0	-0.058	63
Agreeableness	0.048	60	0.206**	2	0.152	17
Willingness to take risk	-0.211**	4	-0.181*	6	0.021	87
Subjective health and life style at age 18						
Satisfaction with health	-0.119	45	-0.143	28	-0.050	80
Current health situation	-0.138**	4	-0.048	39	0.083	30
Currently smoking	-0.007	88	-0.110***	0	-0.107**	5
Other leisure activities at age 17						
Watching TV daily	-0.099***	0	-0.115***	0	-0.016	71
Playing computer games daily	-0.075**	2	-0.044	14	0.029	47
Reading daily	0.067**	5	0.127***	0	0.062	14
Aggregate outcome indices (Anderson, 2008)						
Cognitive skills & school grades	-0.020	75	-0.084	16	-0.016	83
Big Five	0.050	39	0.148***	0	0.130*	6
Non-cognitive skills	0.015	76	0.107***	1	0.116**	5
Outcomes that should not be affected						
Own room at home	0.041**	5	0.017	36	-0.014	55
Weekly pocket money (Euros)	-0.330	73	-0.793	48	0.782	66
25% or more foreign students in school class	-0.048	20	0.004	91	0.022	63
Attending church monthly	0.007	88	0.126***	0	0.144***	0

Note: SOEPv29, own calculations. Effects presented are average treatment effects for the target population. Inference is based on 4999 bootstrap replications. Significance levels: * 10% ** 5% *** 1% level. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalized to mean zero and variance 1 (higher value of grades is better). Aggregate outcome indices are generated following Anderson (2008) and refer to groups from the standard outcomes in the upper part of the table. All other outcome variables are binary, except for 'satisfaction with health' (0 "worst" to 10 "best"), 'current health situation' (1 "bad" to 5 "very good"), and 'amount of weekly pocket money' (Euros).

Table 3.7 – *Average effects of doing nothing vs. music or sports*

	Sports only vs. doing nothing (1)		**Music only vs. doing nothing** (2)		**Music + sport vs. doing nothing** (3)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17						
Cognitive skills - Average	0.080	19	0.194**	3	0.445***	0
- Analogies	0.063	37	0.332***	1	0.427***	0
- Figures	0.169***	1	0.190*	9	0.429***	0
- Math	0.033	61	0.049	61	0.205	13
School grades - Average	0.124***	0	0.230***	1	0.287***	0
- Maths	0.076	17	-0.006	95	0.206***	0
- German	0.144***	0	0.390***	0	0.270***	0
- 1st for. lang.	0.089**	4	0.213**	2	0.217***	0
Attends upper secondary school	0.043***	1	0.155***	0	0.200***	0
Aim to enroll at university	0.043**	2	0.153***	0	0.218***	0
Attends university at age 20	0.040**	4	0.069*	9	0.170***	0
Personality at age 17						
Big 5: Conscientiousness	0.160***	1	-0.052	74	0.251***	1
Extraversion	0.152**	3	-0.004	96	0.231**	3
Neuroticism	-0.061	48	0.031	81	0.036	73
Openness	0.036	59	0.414***	0	0.329***	1
Agreeableness	0.066	26	0.086	53	0.301***	0
Willingness to take risk	0.197***	1	0.035	78	0.033	78
Subjective health and life style at age 18						
Satisfaction with health	0.204**	2	0.180	25	0.072	76
Current health situation	0.139***	0	0.014	84	0.052	51
Currently smoking	-0.033	22	-0.090	11	-0.148***	0
Other leisure activities at age 17						
Watching TV daily	0.013	63	-0.091**	2	-0.082***	1
Playing computer games daily	-0.022	21	-0.067*	6	-0.080***	1
Reading daily	-0.017	42	0.091**	2	0.114***	0
Aggregate outcome indices (Anderson, 2008)						
Cognitive skills & school grades	-0.054	13	-0.072	26	-0.087*	10
Big Five	0.083**	2	0.112	11	0.242***	0
Non-cognitive skills	0.110***	0	0.104*	8	0.221***	0
Outcomes that should not be affected						
Own room at home	0.025**	2	0.070***	0	0.056***	0
Weekly pocket money (Euros)	0.570	34	-0.009	98	0.772	67
25% or more foreign students in school class	-0.016	41	-0.063	12	-0.041	23
Attending church monthly	0.044*	8	0.055	23	0.199***	0

Note: SOEPv29, own calculations. Effects presented are average treatment effects for the target population. Inference is based on 4999 bootstrap replications. Significance levels: * 10% ** 5% *** 1% level. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalized to mean zero and variance 1 (higher value of grades is better). Aggregate outcome indices are generated following Anderson (2008) and refer to groups from the standard outcomes in the upper part of the table. All other outcome variables are binary, except for 'satisfaction with health' (0 "worst" to 10 "best"), 'current health situation' (1 "bad" to 5 "very good"), and 'amount of weekly pocket money' (Euros).

included all types of sports and were thus removed from the tables in the main body of the text.

Column (1) of Table 3.6 presents effects of playing music only compared to doing sports only: Musically active adolescents obtain better school grades in languages, scoring about one sixth of a standard deviation above athletically active adolescents. This tendency of music leading to better skills exists for most of the school achievement and cognitive skills variables, although these differences are not always statistically significant. Interestingly, mathematics grades and competencies seem to be better among sports participants, although the coefficients are not statistically significant. Furthermore, adolescents who play music are 8 percentage points more likely than sports participants to attend upper secondary school, and 10 percentage points more likely to aim at going to university. The advantage of music participation in terms of school achievement also appears when contrasting the outcomes of each activity to those of not being active (see Table 3.7).

When comparing both activities to doing only one of them (columns 2 and 3 of Table 3.6), the results for educational outcomes show that doing both activities jointly is more beneficial than just a single one. For example, adolescents playing both music and sports outperform those who only play sports by more than one fifth of a standard deviation in the cognitive skills test. This effect is similarly positive in comparison to those who play music only. In addition, individuals active in music and sports have better school grades, although only one of the four grade effects is significantly different from zero for each comparison. However, comparing adolescents who play both music and sports to those who only do one activity involves a more important risk of selection based on unobservable characteristics than the single-activity comparisons described above. Being this much engaged in extracurricular activities requires particular investments. Students with weak school performance or students who have work after school will probably not find the necessary time to play music and sports. Let alone the particular financial and emotional support that is necessary from parents and teachers.

As described in Section 3.4, adolescents who play music and sports carry out each activity with a time commitment similar to those who are only active in one of them. Irrespective of potential benefits from music or sports, playing both activities might impede school achievement due to the large amount of time spent on this extracurricular involvement. However, the above-mentioned results suggest that on average doing both activities does not put so much strain on the adolescents as to affect negatively their educational success, but quite the opposite.

With respect to personality, we observe several statistically significant differences between musically and athletically active adolescents. The positive effect of music on openness is statistically significant and substantial at more than one third of a standard deviation. This is related to the fact that openness to artistic experiences is one of the three items assessing this personality dimension in our data. Musically active adolescents are less willing to take risks than those who play sports. Indeed, the comparison to inactive individuals (Table 3.7) shows that music and willingness to take risks are statistically unrelated. Finally, adolescents who do sports in addition to music are more than one quarter of a standard deviation more extrovert than those who play music only.

The third block of results in Table 3.6 shows that being athletically active leads to improved health outcomes. Adolescents who play sports in addition to or instead of music are more satisfied with their health and report a better subjective health status. Indeed, the comparison to non-participation (Table 3.7) reveals that these effects result from a positive effect of sports, compared to no effect of music participation. Of course, given our data, it remains a somewhat open, although maybe philosophical question whether this gain in subjective health translates into objective health measures of this young population (which is of course healthier than the population at large). Considering additional effects of doing two activities versus one, it turns out that adolescents who do both activities are 10% less likely to smoke than adolescents doing only one of those activities are.

Finally, adolescents active in music and sports differ with respect to their leisure time occupations, as described in the next panel of Table 3.6. Adolescents who play music are more likely to read and less likely to watch TV or play computer games every day than those doing sports. If we believe that playing computer games and watching TV is on average not supportive of educational success while reading fosters it, then these results are in line with the findings for the educational outcomes and may provide at least a partial explanation of these findings.

Given the many outcome variables we look at, some of the statistically significant results might appear simply by chance. Anderson (2008) considers such random rejections of the null hypothesis in the context of multiple hypothesis testing as a major problem, which does not receive the attention it deserves by economists. A simple test consists in reducing the number of outcomes by aggregation. To aggregate a group of outcomes into one index, Anderson (2008) suggests taking a weighted average of normalized outcomes, weighting each by the sum of the row entries in the inverted covariance matrix of all outcomes contained in the outcome group. This procedure gives more weight to outcomes that add new information to the aggregate index, i.e. that are less correlated with the other outcomes.

The second-to-last block of Table 3.6 shows the effects of music versus sports for three aggregated outcome indices constructed according to Anderson (2008). The first summarizes cognitive skills and school grades, the second is a weighted mean of the Big Five personality traits, and the last is an index of all six non-cognitive skills presented in the table. In the comparison of doing music rather than sports, none of the aggregated outcome indices appears significant, which means that the outcomes we find for cognitive and non-cognitive skills might be random results of multiple hypothesis testing. However, the size and direction of the coefficient for the first two outcome groups (cognitive skills/school grades and Big Five) confirm the findings discussed above.

3.5.3 Heterogeneity

Next, we investigate heterogeneities in the effects for different subgroups of the target population with respect to gender and socio-economic background (for detailed results see Tables B.11 to B.19 in Appendix B.2). According to the literature, as well as the descriptive statistics shown in Section 3.3, male and female sports are different. Our main results consider this specificity by matching adolescents by gender (but of course subsequently averaging their effects). In the following, we estimate outcome differences between musically and athletically active adolescents for each gender separately to understand gender-specific differences between both activities. Before going into detail, note however that most of the following comparisons, documented in Appendix B.2, suffer from a loss of power, the magnitude of which varies between small and substantial. This loss is due to the reduced sample sizes within the additional strata.

With respect to cognitive skills and school grades, music seems to be more beneficial for girls, while sports may rather help boys. Girls see their language-related skills and grades improve by one fourth of a standard deviation when playing music, while athletically active boys obtain better results in mathematics. Male adolescents drive other differences between music and sports only. The positive effect of music on ambition is stronger for males than for females. Finally, male sports participants increase their degree of extraversion compared to male musicians. This difference cannot be observed among women.

In order to investigate outcome differences between music and sports for different levels of socio-economic status, we estimated these differences in various subgroups. These subgroups stratify the full sample by parental labor market income, parental education, as well as the adolescent's abilities, as measured by the recommendation for upper secondary school received from their primary school teacher. In summary, we find that adolescents with richer or more highly educated parents, as well as adolescents with higher ability, have an advantage from playing music rather than sports for a variety of outcomes. For them, music is associated with higher cognitive skills and school grades. This is even more the case if the adolescent is active in two activities rather than one. In other words, while music and sports are similarly beneficial for children from disadvantaged social backgrounds, those from richer and more highly educated families do better with music than with sports. They also do better with two rather than one activity. One possible explanation relates to Bourdieu's cultural reproduction hypothesis. According to Bourdieu and Passeron (1990), parents from higher socio-economic status have access to better quality cultural education, leading their children to benefit more strongly from playing music.

3.5.4 Robustness and sensitivity tests

One advantage of propensity score matching in comparison to OLS is related to the possibility to verify the common support between treatment and control group. If both groups are very different in their covariates, linear estimations might be sensitive to minor changes in the specification (Imbens, 2015). All our estimations have common support along the distribution of the propensity score between treatment and control groups as well as the target population. Even though we control for a large number of individual and parental background characteristics, we can identify control observations along the entire distribution of propensity scores in the treatment group, and vice versa.

Our estimations are robust to various modifications in treatment definition and estimation method. For example, all effects are slightly stronger if we restrict the sample of musically active adolescents to those who take music lessons outside of school. Moreover, our results are robust to the alternative definitions of being active in music or sports, in which we disregard those who state to be active less than once a month. The results of these estimations are presented in Tables B.20 and B.21 in Appendix B.2.

We also checked the robustness of the results with respect to the particular method used for the bias correction step in the matching estimator. It turns out that, here, the bias correction does not matter. Similarly, varying the criteria to define the common support from 99% to 97% does not change the results in any relevant way (see Table B.22).

Panel attrition and item non-response do not appear to threaten our results either. We check this by creating additional outcome variables measuring non-response. The effects on these missing indicators are small and statistically insignificant (see Table B.23). In addition, all results are robust to including survey weights (see Table B.24).

We believe that the large number of informative and relevant individual and family characteristics included in our estimations is likely to make the conditional independence assumption (CIA) plausible. Thus, our estimates have a causal interpretation. Still, we cannot entirely exclude the possibility that unobserved characteristics determine the choice between music and sports and simultaneously affect adolescents' outcomes. Although, of course, the CIA is untestable empirically (in a strict statistical sense), we run four plausibility tests to shed some light on the potential validity of this assumption and the effects of deviating from it.

First, we examine several outcomes that should be affected by neither music nor sports, which are, however, correlated with socio-economic status. These include whether adolescents have their own room at home, the weekly amount of pocket money they receive, the share of foreign students in their class, as well as whether they regularly attend church. Measured at age 17 (post-treatment), these outcomes are not included as explanatory variables in our propensity score estimation.

The last block of Table 3.6 shows the effects of music versus sports, as well as both activities versus only one of them on these additional outcomes (the effects of being active versus doing nothing are shown in Table 3.7. As expected, these outcomes are not affected in the comparison of adolescents playing music rather than sports, with the exception of a weakly significant coefficient for the share of foreign students in the adolescent's school class. Especially the comparison of music and competitive sports passes this test. By contrast, the large and strongly significant coefficient for church attendance among children who play both rather than only one activity indicates that these comparisons of very active youth might suffer from endogeneity. As Table 3.7 shows, the comparison of active with inactive adolescents also suffers much more from potential endogeneity.

Second, we run the formal sensitivity test developed by Ichino et al. (2008) and implemented by Lechner and Downward (2013). This test is based on a simulation exercise to evaluate the sensitivity of the results with respect to the inclusion of a strong confounder that is correlated with treatment and outcomes, but uncorrelated to previously included covariates. It turns out that the results appear not to be sensitive to the investigated misspecifications (the details of the procedure used and the results are discussed in Appendix B.4).

Third, we used an instrumental variable to model the choice between music and sports. As instrument, we took an indicator of whether either parent played a musical instrument in their youth. This instrument arguably has no direct impact on children's outcome other than influencing the parents' taste with respect to their child's leisure activity conditional on the covariates (for a more detailed discussion, see Appendix B.2). The empirical results do not contradict our main findings. However, this test is not very powerful, because the IV results are subject to substantial sampling noise.

A final test addresses one major weakness of our data: we only observe information on music and sports participation among adolescents who carried out the respective activity until age 17. Therefore, all comparison groups defined by not doing a specific activity contain some individuals, who played that activity earlier in life, but gave up before answering the SOEP youth questionnaire at age 17. Fortunately, it is possible to reconstruct the history of sports and music participation back to age 12 or 14 for a subsample of 752 adolescents. Indeed, 15% of sports and 45% of music participants previously played the other activity as well. In general, dropouts are situated between active and inactive individuals in terms of socio-economic background. Compared to those who continue to play until age 17, drop-outs are characterized by weaker school grades and a lower level of conscientiousness. However, these characteristics are similar between those who give up music and those who give up sports. Please refer to Appendix B.4 for more details.

3.6 Conclusions

Using data from the German Socio-Economic Panel, this paper analyses the effects of spending part of adolescents' leisure time playing music or doing sports, or both. Concentrating on these competing leisure time activities explicitly recognizes that doing one potentially beneficial activity may crow out another one due to lack of sufficient free time. Thus, our estimated effects have a more obvious and useful interpretation than those found in previous studies, comparing each activity separately to 'not doing that particular activity'.

There is also an econometric advantage, because the decision to become active at all becomes irrelevant. If selection into music or sports can be thought of as a two-stage process, with first, a decision to spend time and money on being active, and second, the decision between music and sports, then only the second stage is of potential concern in our setting. Thus, the credibility of identification may be greatly improved. This argument is supported by the fact that, even if the decision to play music or sports is made in one rather than two steps, differences in a wealth of observable variables between active groups are much smaller than between active and inactive adolescents. The fact that active adolescents appear to be rather homogeneous with respect to (rich) observable characteristics, irrespective of the activity they carry out, suggests that remaining differences on unobservables are small, if not absent, as well.

The main results indicate that music and sports may indeed have quite different implications. Playing music appears to foster academic performance and academic ambition more than doing sports, in particular for girls and children from more highly educated families. It also goes along with additional time spent reading and less time spent watching TV or playing computer games. Doing sports, on the other hand, improves adolescents' subjective health. Since better academic performance, as well as better health usually improve future labor market success (and life satisfaction), parents (and children) have to consider their preferences with respect to these outcomes when choosing between music and sports. Since both effects should be valuable to policymakers, our study does not imply favoring one activity over the other with regard to public funds. A further aspect is related to costs: Although we are not aware of any data that are sufficiently precise to feed into a cost-benefit analysis, the descriptive evidence presented in Section 3.4 suggests that taking music lessons is the more expensive alternative. Finally and interestingly, when comparing the effects of doing both activities versus doing only one of them, positive effects appear for doing both. However, this latter comparison is more likely to suffer from endogeneity than the main results.

An extensive robustness analysis based on non-affected outcomes, a formal sensitivity analysis, as well as a semi-parametric instrumental variable estimation suggests that our results are robust with respect to several possible violations of key identifying assumptions, as well as various plausible values of the tuning parameters of the estimator. An interesting avenue for future research consists in understanding the reasons ('mediators') behind our findings. Such an endeavor, however, requires larger samples than the ones available for the current study.

4 Improving access to non-formal education – evidence from the German "Educational package"

4.1 Introduction

In Germany and other developed countries, many children take part in extracurricular activities. For example, 36% of German 13- to 17-year-olds play a musical instrument, and 85% play sports. More than 60% do these activities in institutional contexts such as music schools or sports clubs (Grgic and Züchner, 2013). Categorized as non-formal education, activities at these institutions are carried out on a voluntary basis and usually do not lead to a certificate.

A recent body of literature in the economics of education and other disciplines suggests that playing music or doing sports benefits a variety of education-related outcomes. Music training fosters cognitive skill development (Schellenberg, 2004; Winner et al., 2013) and is likely to affect non-cognitive skills as well (Hille and Schupp, 2015; Schellenberg, 2016). Playing sports enhances educational and labor market outcomes (e.g. Cabane and Clark, 2015; Lechner, 2009; Stevenson, 2010), an effect that might partly be driven by improvements in health (Cabane et al., 2015).

However, access to non-formal education strongly depends on socio-economic status. Among German adolescents from the highest quintile of the household income distribution, 80% take music lessons, prepare for sports competitions, or play theater at least once every week. By contrast, only 48% of those from the lowest income quintile do so. Similarly, among adolescents of mothers with upper secondary school degrees, 73% take part in at least one non-formal educational activity on a weekly basis, compared to only 54% among individuals with less highly educated mothers (Hille et al., 2014). Given the benefits of these extracurricular activities for skill development and educational success, unequal access might be an important source of inequalities of educational opportunity.

Inequalities in access to non-formal education are also problematic from a political perspective, given that these activities are strongly subsidized in many countries. In Germany, for example, about 90,000 sports clubs, as well as 929 public and numerous private music schools cover all regions of the country and receive substantial amounts of public funding (Deutsche Sportjugend, 2014; VDM, 2014; BDPM, 2015). Thanks to these subsidies, average sports club membership fees are as low as 3.10 Euros per month (in 2013, Breuer and Feiler, 2015). At approximately 38 Euros per month, courses offered by public music schools are considerably more expensive (MIZ, 2015a,b). However, the German Public Music Schools Association (*Verband deutscher Musikschulen*) requires its members to reduce fees for low-income households (VDM, 2011).

This paper examines a policy, which addresses this challenge by providing subsidies for participation in non-formal education directly to low-income households. Established in 2011, the Educational package (*Bildungs- und Teilhabepaket*) covers membership fees or equipment costs related to activities such as music or sports, up to an amount of 10 Euros per month. In addition, the policy offers support for school-related expenditures like school material, class trips, transportation, meals, and private tutoring. Households are eligible for these subsidies if they receive basic income support for job-seekers, social assistance, social housing assistance, or a complementary child allowance.

The Educational package might increase participation in non-formal education both by relaxing the household's budget constraint and by affecting the trade-off between costs and marginal utility. The relative importance of these income and substitution effects depends on fee levels, marginal utility and household income. I expect these effects to be strongest among higher-earning eligible families and those with higher educational levels (see Section 4.2). Not only might these families value non-formal education more strongly, they might have fewer difficulties with the application procedure and self-target into participation (Currie and Gahvari, 2008).

To empirically investigate these hypotheses, I use data from *Families in Germany* (FiD), a longitudinal household survey covering child development information (Schröder et al., 2013). The data is particularly suited for the present study, given that poor and vulnerable households are overrepresented. Children are considered treated by the Educational package, if their household is eligible for the package in 2012 or 2013. In addition to pre- and post-treatment information on participation in various non-formal educational activities, the data contains detailed individual and parental background characteristics.

The Educational package might lead children to start an extracurricular activity, or to increase the time spent with an activity that was already taken before. FiD data only allows me to study changes at the extensive margin, because the amount of time spent with extracurricular activities is unknown. As outcome variables, I consider several indicators of whether children play music or sports outside of school.

Eligibility for the Educational package is related to socio-economic characteristics, which also affect participation in non-formal education. I identify causal effects of the policy by estimating the difference-in-differences between outcomes of eligible and non-eligible children before and after the reform. To make the underlying common trend assumption more plausible, I use propensity score matching to reweight the control group such that it is more comparable to the treatment group. Moreover, one specification restricts the sample to children who were eligible for the Educational package before the policy was implemented. I identify causal effects under the assumption of common trends in treatment and matched control group. Moreover, I assume that effects are not due to another policy shift in the time window studied.

Overall, the Educational package does not affect participation in extracurricular music or sports activities. Potential explanations are, first, that less than one in six eligible children takes advantage of the subsidy. Information deficits, complicated application procedures, as well as a lack of interest in non-formal education are possible reasons for non-participation. Second, about half of the beneficiaries already played music or sports before the Educational package was introduced. Whether they increased participation due to the subsidy is unobservable in the data I use.

Studying different subgroups, I find that children in families with a household income above the median among eligible households increase music or sports participation by 10 percentage points. Before the reform, 40% of these children played music or sports. The reform thus increased participation by 25%. Similar effects are found for children living with both rather than one parent. These findings confirm the abovementioned hypothesis, according to which better-off eligible children are more affected by the policy.

Subsidies increase enrollment in an activity only if financial constraints are an important determinant of non-participation. Whether this is true has been widely debated. Proponents of a material resources hypothesis (Bennett et al., 2012; Damelang and Kloß, 2013) were contradicted especially by Lareau (2011) and Weininger et al. (2015), who show that taste and preferences are relevant predictors for participation in non-formal education as well. This paper contributes to this debate and provides evidence in favor of both hypotheses. Financial constraints cannot be the only constraint for participation, otherwise the Educational package would have a larger overall effect. At the same time, I show that financial constraints might explain non-participation for particular subgroups, who are at the margin of taking music or sports lessons.

This paper also contributes to the debate on cash versus in-kind transfers (Currie and Gahvari, 2008). Part of the voucher subsidizing non-formal education was previously paid to families in cash. The Educational package successfully transfers funds towards investing in children's non-formal education. However, the overall welfare effect of the policy also depends on whether families cut other child-related expenditures as a consequence (Bargain and Donni, 2011).

Finally, this paper contributes to the literature by providing first causal evidence for an important policy that affects 2.5 million children in Germany. Contrary to its official evaluation, commissioned by the German Federal Ministry of Labour and Social Affairs (BMAS, 2015b), I observe sports and music participation not only after, but also before the policy was introduced. In BMAS (2015b), as well as in another study by Apel and Engels (2012), retrospective information on these outcomes was only collected for the subgroup of children who benefitted from the subsidy. Apel and Engels (2012) do not observe a control group of non-eligible children at all.

In the next section, I describe the Educational package in more detail and discuss the potential effects it may have. Sections 4.3 and 4.4 present the data and method used in the empirical estimations. Results are presented in Section 4.5, before Section 4.6 concludes.

4.2 Institutional background and potential effects

4.2.1 Institutional background

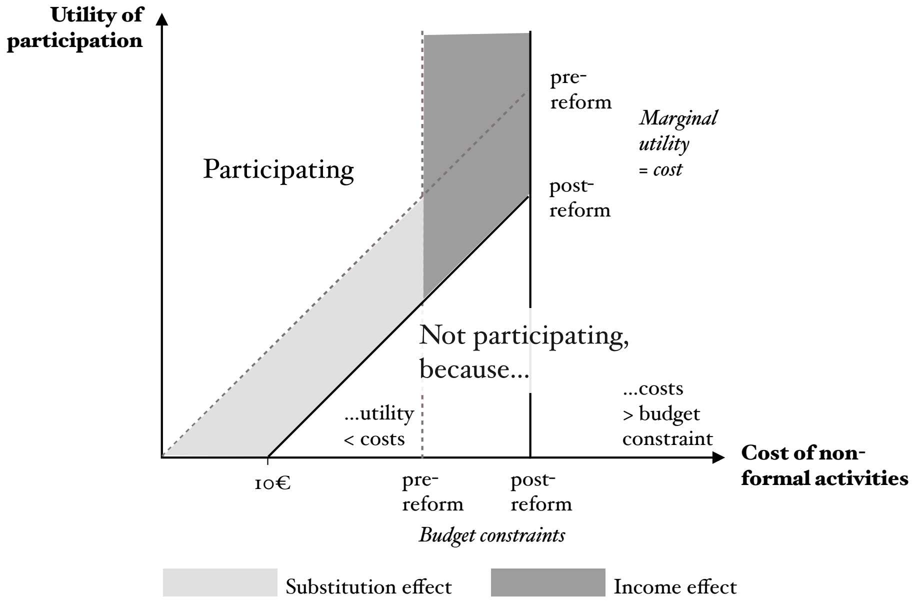
In a widely debated judgment from 2010, Germany's federal constitutional court determined that transfer payments to low-income families were not sufficiently based on actual needs. According to the court ruling, means-tested benefits ignored important aspects, such as educational, as well as non-recurring expenditures, especially for children (BVerfG, 2010). Following requirements of the ruling, Germany's federal government established the Educational package (*Leistungen für Bildung und Teilhabe*), which came into force on April 1st, 2011, with retrospective effect starting on January 1st, 2011. Children are eligible if their household receives basic income support for job-seekers (*Arbeitslosengeld II*), social or housing assistance (*Sozialhilfe* or *Wohngeld*), or a means-tested child allowance (*Kinderzuschlag*). These are transfer payments to families who are unable to cover their basic costs of living by their own means. Approximately 2.5 million children qualify for the Educational package every year (Damelang and Kloß, 2013).

The policy consists of various subsidies that cover costs of school equipment, school trips, school meals, transportation to school, private tutoring, and participation in non-formal education. While support for school equipment is directly transferred to most eligible households in the amount of 100 Euros per school year, all other subsidies are subject to application. Sports club or music school membership fees, or similar expenditures for non-formal educational activities, are subsidized in the amount of 10 Euros per month (BMAS, 2015a).

Subsidized non-formal education includes not only sports club or music school membership fees, but also other types of structured extracurricular activities like visits to museums, theater groups or scouts. As of August 2013, costs for equipment are covered as well (BMAS, 2014, 254). Families can apply for these subsidies by filing an application, in which they provide contact details about the institution at which the activity is carried out, and join a proof of enrolment. Subsidies are directly transferred to the institution or provided as a voucher. If this is impossible, expenditures are reimbursed. Unused subsidies can be transferred to another month within the calendar year.

Does a monthly subsidy of 10 Euros cover the costs of non-formal education? According to Breuer and Feiler (2015), average membership fees in German youth sports clubs

Figure 4.1 – *How the Educational package affects participation in non-formal education*



are as low as 3.10 Euros per month (not counting costs of equipment). Based on this figure, as well as on average expenditures found in the Federal Statistical Office's survey of income and expenditure, the federal government considered 10 Euros as sufficiently generous to carry out even more than one activity (Deutscher Bundestag, 2010, 106). However, average monthly fees at German public music schools are approximately 38 Euros per month[18], although public music schools are required to reduce fees for low-income households (VDM, 2011). Thus, 10 Euros per month might be insufficient in some cases. Moreover, social assistance payments for children were reduced by 3.58 Euros, which were previously dedicated to "extracurricular courses" and "membership fees without purpose of gain" (Deutscher Bundestag, 2010, 106). With other words, the voucher provided by the Educational package partly replaces a previously existing cash transfer.

4.2.2 Potential effects

Figure 4.1 describes the mechanisms through which the Educational package might increase participation in non-formal education. The x-axis displays costs associated with an activity, while the y-axis represents its utility. Utiliy can be derived not only from the

[18] This figure was calculated by dividing total fee revenues (MIZ, 2015a) by the total number of students (MIZ, 2015b)

direct pleasure of doing activities like music or sports, but also from expected benefits in terms of skill development.[19] Without subsidies, the dashed diagonal line illustrates the decision to engage with non-formal education. If marginal utility is equal or higher than costs, individuals decide to become active. At a certain cost level, represented by the dashed vertical line, the budget constraint becomes binding and the household can no longer afford the activity, irrespective of its marginal utility. The same reasoning applies for the decision to increase the amount of time spent with non-formal education, with other words, it applies for decisions both at the extensive and the intensive margin.

I expect the Educational package to have a substitution, as well as an income effect. Both follow from a 10 Euros reduction in the costs faced by participating families. The subsitution effect, represented by the light gray area in Figure 4.1, increases participation among children, whose marginal utility was previously lower than costs. Now, marginal utility is compared to subsidized rather than actual costs.

The income effect, represented by the dark gray area in Figure 4.1, is a consequence of relaxing the budget constraint. For some families, non-formal education might have been worthwhile in terms of utility already before the reform. However, they were unable to afford its costs. Given that expenditures for non-formal education are now subsidized, families can afford activities at a cost exceeding their household's budget constraint by 10 Euros.

The relevance of substitution and income effects varies as a function of costs, the household's budget constraint, as well as utility of participation, and can be summarized as follows. Everything else equal, the Educational package increases participation in non-formal education to a small extent if costs are low, as only the substitution effect plays a role. If costs are between the pre- and post-reform budget constraint, it does so to a larger extent, because both income and substitution effects become relevant. By contrast, the reform has no effect at all if costs are larger than the post-reform budget constraint, given that households are unable to afford non-formal education altogether.

Given a sufficiently high level of utility, the Educational package increases participation in non-formal education among households that did not enroll their children previously, because costs exceeded the households's ability to pay by up to 10 Euros. For these

[19] For a discussion of the motivation and determinants to play music or sports, see e. g. Hille and Schupp (2015) or Cabane et al. (2015).

families, the income effect releases the budget constraint. In addition, the substitution effect leads to an increase in participation among children, whose utility is below costs with a difference equivalent to 10 Euros.

Unfortunately, the data used in this study observes neither costs, the budget constraint, nor utility. I therefore test the hypotheses stated above by examining heterogeneous effects by household income and monthly household expenditure for culture, which could be considered as proxies for the budget constraint and utility of participation. Moreover, I am unable to observe the amount of time dedicated to non-formal education. Therefore, I can only study changes at the extensive margin.

Theoretically, the other components of the Educational package – subsidies for school equipment, school trips, transportation to school, school meals, and private tutoring – might affect participation in non-formal education as well. If household income is substitutable between different types of child- and education-related expenditures, these other subsidies relax the budget constraint. By contrast, time spent in other subsidized activities such as school trips or private tutoring, might no longer be available for music or sports. Not separately measurable, these indirect effects are part of the treatment effect I identify in the empirical estimations below. Given that they affect participation in both directions, these indirect effects are probably small on average.

4.3 Data

To my knowledge, Families in Germany (FiD) is currently the only data set that combines all features necessary to study how the Educational package affects participation in non-formal education.[20] Created in 2010 as part of a large evaluation of family policies in Germany, FiD is a household panel study, which is closely related to the German Socio-Economic Panel (SOEP). FiD contains samples of single parent and low-income families, families with many children, as well as cohort samples of children born in 2007, 2008, 2009,

[20] Another possible choice of data could have been the *Panel Arbeitsmarkt und Soziale Sicherung (PASS)*, a study conducted by the German Institute for Employment Research. PASS contains a sample of households receiving basic income support for job-seekers (*Arbeitslosengeld II*), as well as a sample representative of the German population. While the sample size of eligible households interviewed both before and after the introduction of the Educational package is rather high, outcome variables – children's extracurricular activities – were only introduced in PASS questionnaires after 2011. Restrospective information on participation in non-formal education was collected for beneficiaries of the Educational package. However, due to non-random selection into eligibility, a study of the policy's causal effects requires data including pre- and post-treatment information both for the treatment and control group.

and 2010. The data can be reweighted to be representative of the German population. In total, 4,000 households are interviewed every year. Using household, individual and parent-child questionnaires, information is collected on all household members, including children (Schröder et al., 2013).[21]

For this study, I restricted the data to children, who were part of a realized FiD household in 2010, before the Educational package was introduced, as well as in 2012 or 2013. The sample is further restricted to children celebrating their third birthday no later than in 2010, given that children below the age of 3 rarely participate in non-formal education (Schober and Spieß, 2012). Of these 5,000 children aged 3 to 16, around 700 had to be dropped because parents did not answer the questionnaire items on extracurricular activities described below. Answers can be missing both for exogenous or endogenous reasons (insufficient questionnaire space or refusal to answer). However, children with and without information on non-formal education do not seem to systematically differ in terms of observable characteristics. The final sample consists of 4216 children observed between 2010 and 2012 (born between 1997 and 2007), among whom 3457 were also observed in 2013. Please refer to Table C.1 in the appendix for details on how the final data set was constructed.

FiD is an appropriate data source for this paper, because of its focus on low-income and multi-child households, which leads to an overrepresentation of families who are eligible for the Educational package. Eligibility is used as the treatment variable and was constructed from the four conditions described in Section 4.2: receiving basic income support for job-seekers, social assistance, housing assistance, or the means-tested complementary child

[21] As of 2014, FiD households were integrated into the SOEP, and both data sets are distributed under the common label SOEP.

Table 4.1 – *Patterns of eligibility in estimation sample*

Sample 2010-12				**Sample 2010-13**			
Pattern of eligibility				*Pattern of eligibility*			
2010	*2012*	*N*	*%*	*2010*	*2013*	*N*	*%*
No	No	2608	61.8	No	No	2103	60.7
Yes	No	486	11.5	Yes	No	455	13.1
No	Yes	210	5.0	No	Yes	175	5.1
Yes	Yes	917	21.7	Yes	Yes	729	21.1

Source: FiD v4, not weighted, own calculations. Sample sizes for the sample of all children aged 3 to 16. Table C.4 in the appendix shows the same information on schoolchildren only.

allowance. Every survey year, FiD households are asked whether they received these subsidies. In the main specification, I consider children as treated if their household was eligible for the Educational package in 2012 or 2013. Table 4.1 shows that in both years, 26% to 27% of the sample was eligible. Moreover, more than 15% changed eligibility status after 2010.

FiD is particularly well suited to address this paper's research question, not only because of its focus on low-income households, but also because it contains yearly information on extracurricular activities. Parents are asked to describe their children's participation in non-formal education, distinguishing between 1) pre-school activities, as well as those that take place 2) at or 3) outside of school. In 2012 and 2013, the first category also distinguishes between activities at and outside of daycare centers. Within each context, possible answers include sports and music, as well as some category-specific options (e.g. scouts, youth centers, or the Red Cross).

The list of activities covered by the Educational package is only vaguely defined. Therefore, it is not obvious which outcome variable to prefer. I concentrate on music and sports, given that these are played by many children in Germany (Grgic and Züchner, 2013; Hille et al., 2014) and were surveyed for all age groups. In the full sample of children aged 3 to 16, the main outcome variable measures whether children play sports outside of school or kindergarten. For robustness, an additional outcome variable measures participation in sports or music at or outside school or kindergarten. Activities at school or kindergarten (school clubs) are unlikely to be affected by the subsidy, given that they are typically free of charge. Almost no child in the sample below school age plays music outside of school. Music is therefore only considered among the outcomes described next.

Two further outcome variables are restricted to schoolchildren. They include whether children play sports outside of school, as well as whether they play music or sports outside of school. As one the most important extracurricular activities especially among children from less advantaged households (Cabane et al., 2015), sports are considered in a separate variable. For robustness and similar to the full sample described above, I also measure whether children play music or sports at or outside of school.

Considering schoolchildren separately can be justified with two reasons. First, as described above, questions on pre-school extracurricular involvement were modified in 2012. Comparisons of pre- and post-treatment outcomes in the full sample might therefore show

Table 4.2 – *Participation in non-formal education*

	2010	2012	2013
All children aged 3 to 16			
Sports outside school or kindergarten (in %)	49	54	56
Music or sports, at or outside school or kindergarten (in %)	62	72	73
Number of observations	4221	4221	3462
Schoolchildren (up to age 16)			
Sports outside school (in %)	56	55	55
Music or sports outside school (in %)	64	63	62
Music or sports, at or outside school (in %)	74	72	72
Number of observations	2432	2432	1902

Source: FiD v4, own calculations. Figures for 2011 are not displayed, as the data for that year is not used in this study.

inconsistent variations over time. Second, parents of school-aged children are more likely to know about the Educational package (BMAS, 2015b, 281) and to take advantage of its subsidies (BMAS, 2015b, 298). Not only do they receive funding for school equipment, but they are also regularly informed by teachers about the possibility to finance school trips, as well as about the other components of the Educational package.

Table 4.2 shows the share of participating children for each of these outcome definitions by survey year. Overall, involvement was very constant over the time period studied. The sharp jump in participation within the sample of all children results from the change in questionnaire design described above. While this discontinuity is independent of treatment status and therefore irrelevant for the identification of causal effects, we should keep it in mind when interpreting the results.

In addition to providing information on eligibility for the Educational package and participation in extracurricular activities, FiD contains numerous individual, household, and parental background characteristics, which can be used as control variables. Similar to other studies, I include the following individual-level covariates: gender, birth year, migration background, school type and grade, as well as whether the child attends a full-day school. In the sample of all children, I further include a variable indicating whether the child attends school. On the household level, I take into account whether children live in a single parent home, whether at least one parent has an upper secondary school or university degree, as well as the log of monthly net household income. In addition, I control for the households' monthly expenditures for culture, leisure, and education, as well as whether they live in a large city. With respect to the mother, I control for labor

Table 4.3 – *Summary statistics of covariates by eligibility (2013)*

	All children		**Schoolchildren**	
	Non-eligible	*Eligbile*	*Non-eligible*	*Eligbile*
Individual-level covariates (measured 2010)				
Female	0.49	0.48	0.50	0.51
Migration background	0.31	0.42	0.27	0.38
Attends school	0.54	0.58	1.00	1.00
Primary school	0.33	0.40	0.62	0.68
Lower secondary school track	0.03	0.04	0.06	0.07
Middle secondary school track	0.06	0.05	0.10	0.08
Upper secondary school track	0.08	0.03	0.14	0.04
Comprehensive secondary school	0.02	0.03	0.04	0.05
Other school	0.48	0.46	0.04	0.07
School grade	2.02	2.00	3.73	3.41
Full-day school	0.12	0.16	0.21	0.28
Birth year	2002.50	2002.27	1999.94	2000.05
Household-level covariates (measured 2010)				
Nb of children in household	2.50	2.72	2.55	2.77
Single-parent household	0.21	0.40	0.30	0.46
One parent with Abitur	0.38	0.09	0.34	0.07
Parent with university degree	0.32	0.08	0.28	0.07
Log monthly net household income	7.90	7.46	7.86	7.49
Monthly HH exp: education	26.38	10.98	29.79	11.07
Monthly HH exp: culture	15.86	9.42	18.27	10.20
Monthly HH exp: leisure	64.31	30.41	67.77	33.94
Rural area (population < 20k)	0.51	0.38	0.51	0.38
Large city (population > 100k)	0.24	0.30	0.24	0.29
Federal state	7.38	7.65	7.30	7.55
FiD subsample	63.96	64.10	64.52	64.40
Missing: parental education	0.04	0.02	0.04	0.02
Missing: household expenditure	0.05	0.04	0.05	0.04
Mother-related covariates (measured 2010)				
Work hours	14.58	6.74	16.34	8.13
Mother not working	0.40	0.68	0.33	0.61
Age at birth	30.42	27.71	29.99	27.58
Missing: mother's work hours	0.02	0.00	0.02	0.01
Number of observations	2558	904	1377	525

Source: FiD v4, own calculations. For the table distinguishing by eligibility in 2012, please refer to Table C.3 in the appendix.

market status, work hours and age at birth. Finally, all regressions include federal state and FiD subsample dummies. The exact definitions and summary statistics of all variables can be found in Tables C.2 and C.5 of the appendix.

Eligibility for the Educational package is correlated with a lower socio-economic status, which is obvious given that the policy addresses low-income families. Table 4.3 shows that eligible children are more likely to have a migration background and to live in single-parent households. Moreover, they are less likely to attend upper secondary school and to have parents with upper secondary school or university degrees. Furthermore, eligible children have more brothers or sisters, live in poorer households and more often in cities. Mothers of eligible children work less often and are younger. Similar to findings from Marcus et al. (2015), eligible children are more likely to attend full-day schools.

4.4 Method

4.4.1 Identification strategy

As discussed above, families are eligible for the Educational package if they receive welfare benefits targeted at low-income households. Eligibility is therefore not randomly determined. A simple comparison of music or sports participation by eligibility would not reveal the policy's causal effects, but rather reflect socio-economic group differences.

Time-constant unobserved heterogeneity between eligible and non-eligible children can be eliminated using a difference-in-differences estimator. However, we still require the common trend assumption to identify causal effects. With other words, we must assume that music and sports participation in treatment and control group would have evolved similarly over time, had the Educational package not been established. Given that access to these activities strongly depends on socio-economic status (Weininger et al., 2015; Hille and Schupp, 2015), its evolution might do so as well.

I propose two modifications to make the difference-in-differences estimator more plausible. First, I use propensity score matching to reweight the control group, such that it resembles the treatment group. Propensity score matching is similar to ordinary least squares in the sense that it relies on the conditional independence assumption and uses observable covariates to make treatment and control group more similar. But unlike ordinary least squares, propensity score matching does not require any particular functional form assumption (Imbens, 2015) and allows for more flexible effect heterogeneity (Huber et al., 2013). In a robustness check, I use entropy balancing (Hainmueller, 2012), a similar approach that reweights the control group such that covariates have the same mean and variance in both groups. However, entropy balancing is less flexible than propensity score matching. It assumes constant rather than heterogeneous effects, because it does not assign precise matches from the control group to each treated observation.

Second, I refine the difference-in-differences estimator by restricting the sample to children who would have been eligible for the Educational package in 2010, before it was introduced. This approach was suggested by Lechner (2009). It reduces the selection problem by making treatment and control group more similar before the policy was implemented. Of course, families eligible for the Educational package both in 2010 and 2013 are somewhat different from those who manage to get back into employment and lose eligibility. Therefore, this approach still requires propensity score matching to make treatment and control group more comparable.

The estimator described so far identifies causal effects under two assumptions. First, as mentioned above, I assume a common trend, according to which participation in non-formal education would have evolved similarly in treatment and control group, had the policy not been established. Given that my estimation sample only contains one pre-treatment time period (2010), it is not possible to test this assumption. Propensity score matching and the restriction of the sample by pre-treatment eligibility makes treatment and control group more comparable and thereby the common trend assumption more plausible.

Second, I can only attribute potential effects to the Educational package, if no other policy shift occurred between the pre- and post-treatment observation year and affected treatment and control group differently. The estimation sample allows me to measure post-treatment outcomes in 2012 or 2013. On the one hand, both identifying assumptions are more realistic when examining outcome changes between 2010 and 2012. On the other hand, the Educational package was not taken up very much in the first two years of existence (see BMAS, 2015b, or Section 4.5). It might therefore be more interesting to study outcome changes between 2010 and 2013. However, the risk of capturing alternative policy shifts is larger when examining a longer time frame.

The other benefits from the Educational package can be seen as one obvious policy change that took place simultaneously. Subsidies for school trips, transportation, lunch and tutoring affect my treatment effects irrespective of the time period studied. To single out the causal effect of subsidies for cultural participation, I have to assume that the other benefits do not affect participation. Alternatively, as discussed in Section 4.2, I consider the influence of these other components to be part of the treatment effect.

4.4.2 Estimation

I specify the model described above in first differences, using the change in music or sports participation as the dependent, and eligibility in 2012 (or 2013) as the explanatory variable. We can derive this specification from a two-period difference-in-differences model with individual fixed effects:

$$Y_{it} = \beta_0 + \delta t + \alpha T_{it} + \theta_i + \varepsilon_{it} \qquad \text{with} \quad t \in \{0, 1\} \tag{4.1}$$

where $t = 0$ corresponds to the pre-treatment (2010) and $t = 1$ to the post-treatment year (2012 or 2013). Y_{it} is the outcome – participation in music or sports – for individual i in year t, t a time dummy, T_{it} an interaction between time and eligibility, and θ_i the individual fixed effect. In this model, α is the treatment effect, which can be interpreted as the causal effect of being eligible for the Educational package under the common trend assumption.

I obtain the first-difference specification by taking the difference between post- and pre-treatment year:

$$(Y_{i1} - Y_{i0}) = \delta + \alpha(T_{i1} - T_{i0}) + (\varepsilon_{i1} - \varepsilon_{i0}) \tag{4.2}$$

where the difference $T_{i1} - T_{i0}$ equals 1 for treated and 0 for untreated individuals, and is therefore nothing else than the treatment indicator. The coefficient of the previous time dummy becomes the constant:

$$\Delta Y_i = \delta + \alpha T_i + \nu_i \tag{4.3}$$

To implement the propensity score matching estimator, I use a probit model to predict eligibility for the Educational package and control for numerous individual, parental and household covariates, all measured in 2010 (see Section 4.3). I then carry out radius matching with a caliper of 2%. Standard errors were clustered at the household level and obtained using bootstrap. Please refer to Appendix C.3 for more details on the matching estimator.

The estimator described so far identifies the average treatment effect for the treated. Propensity score methods also allow estimating the average treatment effect on the non-treated, as well as the overall average treatment effect. To obtain the former, the treatment group is reweighted in order to resemble the control group. The average treatment effect is the average of the effect on the treated and non-treated, weighted by the share of the respective group in the sample. However, this study is interested in the effect of the Educational package, which explicitly addresses low-income households. We are thus interested in the treatment effect on the treated.

4.5 Results

4.5.1 Take-up of the Educational package

Before studying the effects of the Educational package on participation in non-formal education, this subsection provides descriptive evidence on take-up of the subsidy. For this purpose, I use information available in FiD survey waves 2012 and 2013, in which parents indicate for each of the policy's components whether it was used. Moreover, I compare the findings from FiD with those from the official evaluation, commissioned by the Federal Ministry of Labour and Social Affairs, which uses data from the low-income labor market panel (PASS) of the German Institute for Employment Research (BMAS, 2015b).

When reweighting the FiD sample used in this paper to be representative of the German population, I find that 15% of the eligible children applied for the subsidy covering non-formal education in 2013. This corresponds to the figure found in the official evaluation (BMAS, 2015b, 288), which confirms the quality of the data used in this study.

Take-up of the subsidy for music school and sports club membership fees is especially low in comparison to take-up of the other subsidies covering school trips and school lunches. Transportation to school and private tutoring are the only two components of the Educational package that were demanded even less frequently. However, these are probably not relevant for every child. Moreover, tutoring is only covered for children at risk of failing their school year (BMAS, 2015a).

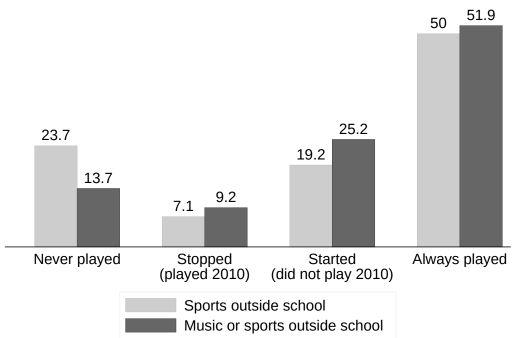
Why do five out of six eligible children forego the subsidy for extracurricular activities? While FiD data does not contain any information that would help answering this question, several plausible reasons are dicussed in the official evaluation of the Educational package. First, only 71% of families without and 90% of families with schoolchildren knew about the existence of the policy in 2013 (BMAS, 2015b, 282). Among them, some might have known about the content in general, but not about the specific subsidy related to extracurricular activities.

Second, even if families knew about the possibility to obtain funding for non-formal education, they might not have considered it worthwile to apply. Among those who did not apply for any of the subsidies covered by the Educational package in 2013, 47% simply did not know which government agency was in charge. About 19% stated that the application procedure was too complicated (BMAS, 2015b, 301). Moreover, the amount of 10 Euros might not cover the costs of music or sports lessons, as well as the necessary equipment (see Section 4.2). Approximately 24% named the lack of appropriate course offers near their home as a further reason for not applying (BMAS, 2015b, 301). Finally, given that the subsidy for non-formal education is usually paid directly to institutions, families might not fully value it, because it is not directly paid out.

Third, as discussed in the introduction, financial constraints are only one reason why children do not play music or sports. Indeed, 54% of the non-applying households justified their decision by stating that they did not require the subsidy (BMAS, 2015b, 301). In 2012, about 30% of the eligible households did not apply with the explicit justification that their child is not member of an association and the subsidy therefore not needed (BMAS, 2014, 240).

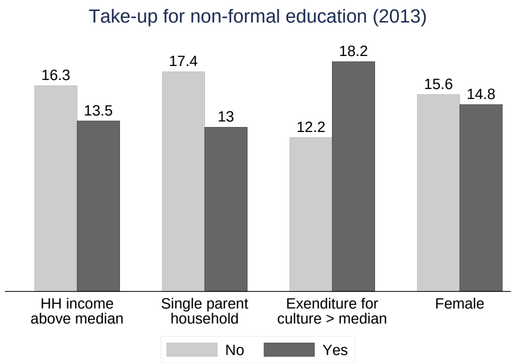
A high take-up rate of the Educational package is not a sufficient condition for an increase in music or sports participation. Children who take advantage of the subsidy might already have practised their activity before. Indeed, among its beneficiaries in

Figure 4.2 – *Pattern of participation in music and sports among beneficiaries of the subsidy for non-formal education (2013)*



Source: FiD v4, weighted, own calculations. Figures for 2012 are found in Figure C.5.

Figure 4.3 – *Take-up of the subsidy for non-formal education by income, household composition, cultural expenditures, and gender (2013)*



Source: FiD v4, weighted, own calculations. Figures for 2012 are found in Figure C.5.

2013, Figure 4.2 shows that 52% had already played music or sports in 2010 before the policy was introduced. While 25% started to become active after receiving the financial contribution, a considerable share never played or even stopped to play music or sports. Apparently, other qualified extracurricular activities such as visits to museums play an important role as well, and might even replace music or sports. In an evaluation covering the policy's first year of implementation, Apel and Engels (2012) find that 78% of the subsidy's beneficiaries had already been member of their association previously.

As discussed in Section 4.2, I expect a high take-up rate among households with a comparatively high income and both parents present. For 2013, Figure 4.3 shows that – contrary to these hypotheses – children with a household income above the median took advantage of the Educational package less often than those living in poorer households. By contrast, with a share of 13% of the eligible children, single parent households benefitted from the subsidy somewhat less often, as expected. Similarly, at 18%, the take-up rate was 6 percentage points higher among children with higher household expenditures for culture, indicating that preferences are likely to play an important role.

4.5.2 Selection model

Families are eligible for the Educational package if they receive basic income support for job-seekers, social assistance, housing assistance, or supplementary child allowances. Eligibility for these welfare payments depends on the household's economic situation. Table 4.4 shows the marginal effects of a probit model predicting eligibility in 2013. As described in

Table 4.4 – *Selection into eligibility for educational package: Probit model (2013)*

	All children		Schoolchildren	
	Marginal effect	*P-value in %*	*Marginal effect*	*P-value in %*
Individual-level covariates				
Female	-0.02	22	-0.00	96
Migration background	0.07	0	0.09	0
Attends school	0.08	27		
School type (Reference: primary school)				
Lower secondary school	-0.02	63	-0.03	53
Medium secondary school	-0.06	17	-0.08	12
Upper secondary school	-0.11	2	-0.12	2
Comprehensive school	-0.06	29	-0.05	38
Other	0.04	43	0.05	38
School grade	-0.02	9	-0.04	1
Full-day school	0.00	89	0.01	79
Household-level covariates				
Nb of children in household	0.10	0	0.10	0
Single-parent household	0.09	0	0.11	0
One parent with Abitur	-0.13	0	-0.17	0
Parent with university degree	-0.02	61	0.02	68
Log monthly net household income	-0.35	0	-0.31	0
Monthly HH exp: education	-0.00	34	-0.00	34
Monthly HH exp: culture	0.00	53	0.00	94
Monthly HH exp: leisure	-0.00	23	-0.00	34
Rural area (population $<$ 20k)	-0.06	2	-0.08	3
Large city (population $>$ 100k)	0.04	21	0.01	71
Mother-related covariates				
Mother's work hours (in 10h/week)	-0.04	0	-0.01	0
Mother not working	-0.01	79	-0.02	63
Age at birth	-0.00	2	-0.00	9
Missing: mother's work hours	-0.25	1	-0.29	1
Number of observations	3462		1902	
Efron's R-square (in %)	37.0		35.7	

Source: FiD v4, own calculations. Marginal effects are calculated at the mean for each covariate. The following covariates are not displayed in the table: birth year fixed effects, federal state fixed effects, FiD sample fixed effects, missing indicators. For the full model, as well as for a model estimating eligibility in 2012, please refer to Tables C.6 and C.7 in the appendix.

Section 4.3, covariates were measured in 2010. The model was separately estimated for the sample of all children, as well as for schoolchildren only. A table including all coefficients, as well as the selection model for 2012 are referred to the appendix (Tables C.6 and C.7).

The characteristics we observe in the data predict eligibility relatively well. R-squares in Table 4.4 are around 36% to 37%. As expected for any welfare payment, household income and parental education are the most important determinants of eligibility. All other characteristics held constant, a 10% increase in monthly household income is associated with a 3.5 percentage points decrease in the probability of being eligible for the Educational package. If at least one parent has an upper secondary school degree, children are 13 percentage points less likely to be in the target group for the subsidy. Similarly, eligibility is 11 percentage points lower among children attending upper secondary school. Given that eligibility is positively correlated with unemployment, the probability to be eligible is 4 percentage points lower for every 10 additional weekly work hours of the mother.

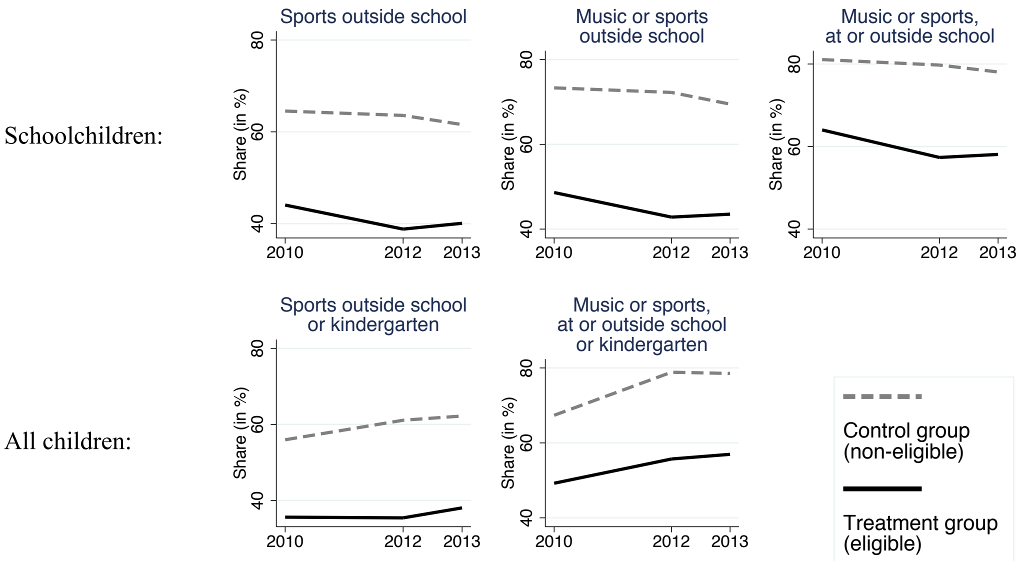
Migration background and household composition play a role, too. Children with a migration background are between 7 and 9 percentage points more likely to be among those eligible for the Educational package. Living in a single parent home and having an additional sibling are each associated with an increase in eligibility by around 10 percentage points. Finally, households living in rural areas are 6 percentage points less likely to be in the target group for the Educational package. All coefficients are similar in both samples and years.

Matching is only possible if treated and untreated individuals exist along the entire distribution of the propensity score. Given its high R-square, the selection model might predict eligibility so well that it is impossible to find control observations with similar propensity scores for some members of the treatment group. As illustrated by the common support graphs plotted in Figure C.1 of the appendix, the majority of non-eligible children are indeed concentrated around very low propensity scores. To ensure that control observations do not receive excessive weights in the matching procedure, all treated observations with a propensity score higher than the highest score within the control group were deleted. Fortunately, this applied to only 5 observations in the sample of all children, and 31 (eligibility 2013) or 46 (eligibility 2012) observations in the sample of schoolchildren. After reweighting the control group to match the treatment group, covariate differences between both groups are no longer significant, as shown in Tables C.8 and C.9 of the appendix.

4.5.3 Average treatment effects

Before discussing the effects of the Educational package in a regression framework, descriptive graphical evidence provides some hints about the evolution of participation in non-formal education in treatment and control group. For each year, Figure 4.4 depicts the share of children attending sports or music activities by eligibility for the Educational

Figure 4.4 – *Participation in non-formal education by eligibility and year*



Source: FiD v4, own calculations.

package. Here, eligibility is determined for every survey year, some children may therefore change groups between years.

The first important observation from Figure 4.4 is that participation in non-formal education is about 20 percentage points lower among eligible than among non-eligible children. This is not surprising, given that characteristics determining eligibility are opposite to those that characterize participation in extracurricular music or sports activities.

Moreover, Figure 4.4 clearly illustrates that the difference between treatment and control group becomes larger from 2010 to 2012. The coefficient of a difference-in-differences estimator between these years would therefore be negative. Although the difference in time trends is unlikely to be statistically significant, a potential explanation could be that subsidies for music and sports participation were taken up very slowly in the first two years after the Educational package was established.

In 2013, however, eligible children seem to catch up. Their participation in non-formal education is still much lower than that of non-eligible children, but the difference between both groups becomes smaller, especially in comparison to 2012. Whether the overall difference in evolution between treatment and control group is positive and significant can be tested using regression analysis.

Table 4.5 – *Regression results (first-differences models; treatment: eligibility 2013)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Regression 1: Simple first differences model					
Treatment effect	0.02	0.03	0.01	-0.03	-0.00
	(60)	(34)	(72)	(30)	(93)
N	1871	1871	1871	3457	3457
Regression 2: Subgroup of children eligible in 2010					
Treatment effect	0.04	0.05	0.02	0.01	0.02
	(42)	(31)	(64)	(81)	(53)
N	656	656	656	1180	1180
Regression 3: Control group reweighted using propensity score matching					
Treatment effect	-0.03	-0.02	-0.05	-0.02	-0.01
	(50)	(68)	(31)	(53)	(87)
N	1871	1871	1871	3457	3457
Regression 4: Propensity score matching, subgroup of children eligible in 2010					
Treatment effect	-0.00	0.01	0.04	0.02	0.04
	(97)	(90)	(61)	(74)	(47)
N	650	650	650	1130	1130

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors of propensity score matching results were computed with 999 bootstrap replications.

Table 4.5 shows the regression coefficients of the difference-in-differences estimator described in Section 4.4, in which I compare the evolution of participating in music or sports activities by eligibility for the Educational package. As described in Section 4.4, I estimate first-difference models using eligibility in 2013 as the treatment. Outcome differences are taken between 2010 and 2013, given that the graphical analysis presented in Figure 4.4 already showed that the policy is unlikely to have an effect in 2012 already. For completeness, results of the difference-in-differences estimator taking the difference between 2010 and 2012 are shown in the appendix (Table C.15).

Table 4.5 shows the results of four variations of the difference-in-differences estimator. Regressions 1 and 2 are simple first-differences models that take no covariates into account. While Regression 1 includes all children or schoolchildren, Regressions 2 restricts the sample to those who would have been eligible in 2010 already. Regressions 4 and 5 reweight the control group using propensity score matching as described in Section 4.4, regression 5 being again restricted to the subsample of children who would have been eligible for the Educational package in 2010.

While none of the treatment effects in Table 4.5 are significant, we can still make two observations. First, coefficients in the subgroup of children who would have been eligible for the Educational package in 2010 are slightly higher than in the overall population. This is consistent with the selection effect discussed in Section 4.4, according to which eligible children are less likely to engage with non-formal activities, and therefore probably experience a smaller growth in participation as well. When comparing only children who would have been eligible before the subsidy was implemented, selection plays a less important role. At the same time, the coefficients illustrate that moving out of eligibility is not associated with an immediate start of music or sports activities.

Second, when reweighting the control group using propensity score matching, all coefficients become smaller. This is somewhat surprising, as it goes against the untestable hypothesis, according to which the participation trend is smaller among eligible than among non-eligible children, had the Educational package not existed. If this was true, we would have expected the treatment group to increase sports or music participation relative to the control group, once the control group is reweighted to be comparable in terms of observable characteristics.

4.5.4 Heterogeneous results

As pointed out in Section 4.2, the Educational package is likely to affect only certain subgroups among the eligible, namely those who are more advantaged with respect to material and cultural resources. The regression results presented in the following confirm this hypothesis. Table 4.6 shows the effects of the Educational package and divides the sample into subgroups according to four characteristics, which aim to model heterogeneity in terms of material resources and preferences.

To distinguish families by material resources, I separately estimate the effects of the Educational package for households with an income above or below the median among eligible households, as well as among children living with both or only one of their parents. Single-parent households are usually considered particularly at risk of poverty. Heterogeneities with respect to preferences were examined by separately estimating the model for households with expenditures for culture above or below the median among eligible households, as well as by gender. While expenditures for culture might be considered as a revealed preference for education-oriented leisure activities in general, gender strongly

Table 4.6 – *Heterogeneous results (schoolchildren, treatment: eligibility 2013)*

	Activities outside school among schoolchildren			
	Sports (1)	*Music or sports* (2)	*Sports* (3)	*Music or sports* (4)
	...by household income (among eligible)			
	Below median		*Above median*	
Treatment effect (First difference)	-0.05	-0.04	0.07	0.10**
	(31)	(41)	(14)	(4)
N	516	516	1355	1355
Treatment effect (First difference & propensity score matching)	-0.10	-0.09	0.10	0.09
	(31)	(34)	(19)	(20)
N	506	506	1325	1325
	...by household composition			
	Both parents present		*Single parent household*	
Treatment effect (First difference)	0.10**	0.11**	-0.05	-0.03
	(3)	(1)	(29)	(55)
N	1239	1239	631	631
Treatment effect (First difference & propensity score matching)	0.07	0.06	-0.06	-0.05
	(31)	(36)	(62)	(70)
N	1218	1218	604	604
	...by household expenditure for culture			
	Below median		*Above median*	
Treatment effect (First difference)	-0.02	-0.02	0.03	0.05
	(77)	(74)	(49)	(20)
N	509	509	1362	1362
Treatment effect (First difference & propensity score matching)	0.02	0.02	-0.05	-0.02
	(88)	(85)	(52)	(73)
N	498	498	1363	1363
	...by gender			
	Male		*Female*	
Treatment effect (First difference)	0.01	0.03	0.02	0.03
	(81)	(43)	(59)	(55)
N	936	936	935	935
Treatment effect (First difference & propensity score matching)	0.01	-0.01	-0.05	-0.04
	(92)	(91)	(58)	(61)
N	924	924	936	936

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). The median monthly household income among eligible children in the estimation sample in 2013 was 2000 Euros. Median household expenditures for culture among eligible children were 2.67 Euros per month. Standard errors were computed with 999 bootstrap replications. Full results for all outcome variables are presented in Tables C.11 to C.14, results for 2012 in Tables C.16 to C.19 of the appendix.

influences the choice between music and sports, given that girls are much more likely to prefer music (Cabane et al., 2015).[22]

For each subgroup, Table 4.6 shows the difference-in-differences coefficient for the two most relevant outcomes among schoolchildren, as well as with and without reweighting the control group using propensity score matching. Results for the other outcomes, as well as for the evolution of participation in music and sports between 2010 and 2012, are shown in the appendix (Section C.4).

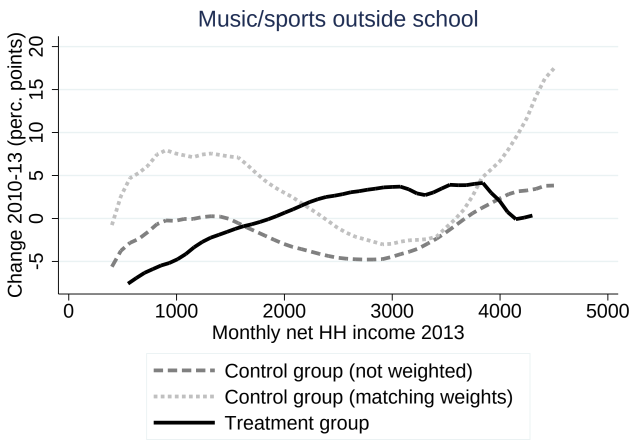
The Educational package seems to increase participation in extracurricular music and sports activities among children with a household income above the median of eligible families, as well as those living with both parents. In both cases, the change in music or sports participation among eligible children exceeds that of non-eligible children by 10 percentage points. However, as soon as the difference between treatment and control characteristics is reduced using propensity score matching, these coefficients slightly drop and become statistically insignificant.

Contrary to what would have been expected, Table 4.6 does not reveal a significantly positive treatment effect for children in households with higher expenditures for culture. While coefficients for these children are larger than for those with lower household cultural expenditures, treatment effects are not significantly different from zero. A potential explanation is that for these children, financial constraints are not the reason not to sign up for music or sports lessons. Alternatively, these children might have increased participation at the intensive, but not the extensive margin. Similarly, no difference in the effect of the Educational package can be observed by gender.

Figure 4.5 displays the results of a local polynomial regression, describing the difference in participation between 2010 and 2013 for treatment and control group along the distribution of monthly net household income. Differences are shown both for the weighted and unweighted control group. The difference-in-differences estimator presented above is positive if the difference for the treated (the solid black line) is higher than that of the untreated. Figure 4.5 plots changes only for participation in music or sports activities

[22] It would also have been interesting to conduct these subgroup estimations by parental education. However, only a low number of eligible parents have an educational level higher or lower than an intermediate school degree (*Realschulabschluss*). At least one subgroup among the treated would thus have been too small to obtain sufficiently precise estimates. By consequence, these results are not presented here, but can be provided by the author on request.

Figure 4.5 – *Changes in music or sports participation by household income (local polynomial estimation)*



Source: FiD v4, own calculations. Figure C.2 shows the kernel density for the entire distribution of household income. Epanechnikov kernel with a bandwidth of 500 (household income). Local polynomial regressions on the difference between 2010 and 2012 are shown in Figure C.3. Children with a household income larger than 4500 Euros were deleted from the sample due to the small density in the treatment group.

outside school, given that the graph looks very similar for other outcomes. Graphs for other outcomes can be provided by the author on request.

The local polynomial regression confirms the finding that participation in extracurricular music or sports activities increases especially among eligible children with a higher household income. The treatment group change lies above that of the control group for household incomes larger than approximately 1700 Euros (without reweighting the control group) or 2200 Euros (taking covariates into account). For incomes above approximately 3800 Euros, the treatment group lies below the control group again, however, only very few treated households are at that income level (see kernel density presented in Figure C.2) and not too much weight should be given to the extreme ends of the distribution.

Overall, the heterogeneous results presented above show that the Educational package increases participation in music and sports activities among children who live in households with relatively fewer material constraints, while no effect was found when distinguishing by cultural preferences. To put this result in perspective, it is important to keep in mind that all eligible households are among the poorest in Germany, otherwise they would not be eligible for welfare payments. One possible explanation is that the subsidy is sufficiently high to change behavior among the less financially constrained among the eligible. On average, however, eligible households might be unable to afford music and sports participation for their children despite the Educational package, even in the subgroup of those who value cultural activities more highly.

4.5.5 Robustness checks

I conducted several robustness checks examining the sensitivity of the results presented above with respect to measurement error and estimation method. As described in Section 4.3, answers from the FiD household questionnaire allow me to construct numerous indicators of extracurricular involvement that might be affected by the Educational package. Sports, as well as music or sports activities outside of school were considered to be the most relevant outcomes. However, next to these, Table 4.5 presents the overall effects of the policy on three additional outcome variables, which are similar in size and significance: 1) music or sports activities at or outside of school; 2) sports outside of school or kindergarten; as well as 3) music or sports, at or outside of school or kindergarten.

Tables C.11 to C.14 show the results for all five outcome variables, distinguishing by the four heterogeneity categories described in the previous subsection. Here we can see that some of the abovementioned heterogeneous effects disappear. However, this might result from the fact that the two main treatment definitions most closely reflect the aim of the Educational package, namely to subsidize extracurricular activities taking place outside of school.

As discussed in Section 4.4, it is not obvious in which year post-treatment outcomes should be measured. Given that the policy was only introduced in April 2011 and is unlikely to affect participation immediately, 2012 or 2013 seem appropriate choices. The advantage of 2012 is the lower risk of capturing alternative policy changes due to the shorter duration since pre-treatment measurement (2010). Measuring post-treatment outcomes in 2013 increases the likelihood of finding effects, due to the longer time of adaptation to the new policy. While the main paper focuses on results for 2013, all estimations were also carried out for post-treatment outcome measurements in 2012, as presented in Tables C.15 to C.19, as well as Figures C.3 to C.4 in the appendix.

In another regression, I pooled the model for 2012 and 2013, which allows me to see how the treatment effect varies over time. In the full sample, as well as in each of the subgroups analyzed in the previous subsection, coefficients were very similar to the regression models concentrating on either 2012 or 2013. Results for the full sample are presented in the appendix (Table C.10), heterogeneous results can be provided by the author on request.

Further robustness tests (appendix, Table C.20) show that the results are not sensitive with respect to the estimation method. I reestimated the propensity score models, changing the caliper from 2% to 1% and using kernel instead of radius matching. Moreover, I used entropy balancing (Hainmueller, 2012; Hainmueller and Xu, 2013) instead of propensity score matching. Finally, both the simple first-differences estimator, as well as the propensity score model were reestimated including survey weights that are provided with FiD data to make the sample representative of the full population of German children. All models produce coefficients similar to those of the main results presented above.

4.6 Conclusion

The present study examines the effects of the Educational package – a subsidy covering the costs of music school or sports club membership for low-income households – on participation in extracurricular music and sports activities. I find that the policy has no effect on the overall sample of eligible children. Potential explanations are that less than one in six eligible children takes advantage of the subsidy, and that more than half of its beneficiaries already played music or sports before the Educational package was introduced. The policy might also have no effect if financial or material constraints are not the main determinant of non-participation.

When examining various subgroups, I find that those with a household income above the median among eligible households increase their participation in extracurricular music or sports activities by approximately 10 percentage points. The same is true for children living with both of their parents. This is in line with the hypothesis that better-off eligible children benefit more, given the interaction between utility, costs and ability to pay in the decision of becoming active. No significant effect difference was found when dividing the sample by cultural preferences.

Among the most important reasons for the low take-up rate of 15% are a lack of information, the complicated application procedure, and the insufficient amount of the subsidy. Knowledge about the possibilities offered by the Educational package has been strongly increasing in the first three years of its existence (BMAS, 2015b). Still, it is astonishing that 40% of the non-beneficiaries, which is equivalent to approximately 35% of all eligible children, receive paid sports or music lessons outside of school and forego a subsidy they are entitled to.

We can expect the subsidy to be taken up by more children in 2014 and onwards. As of August 2013, the application procedure was simplified and costs of equipment were covered as well (BMAS, 2014, 254). Unfortunately, with the inclusion of FiD into the German Socio-Economic Panel, questions on take-up of the different components of the Educational package were not administered in the 2014 survey wave. At least the SOEP survey wave of 2014 therefore does not allow to study the effects of these modifications to the policy.

A further reason for the low take-up rate cannot be easily addressed by modifying the policy itself. As pointed out by Weininger et al. (2015), participation not only depends on financial and material constraints, but also to a large extent on preferences. If music or sports benefit skill development and educational attainment, as previous research suggests, it might be an important policy goal to increase awareness of these findings and to encourage participation among children from all socio-economic backgrounds. Simply reducing financial or material barriers might not be sufficient. A promising, yet costly approach is to directly bring non-formal educational activities to schools, as discussed in the following chapter.

5 How a universal music education program affects time use, behavior, and school attitude

5.1 Introduction

The importance of personality traits and other non-cognitive skills for educational and labor market success is widely acknowledged (e.g. Almlund et al., 2011; Heckman and Kautz, 2012; Heineck and Anger, 2010). Non-cognitive skills are particularly malleable during childhood (Heckman and Kautz, 2014), although it is still widely debated how they can be successfully affected by policy interventions. Model programs are regularly found to be more successful than large-scale policies, partly due to the fact that they are typically more carefully implemented and evaluated (Currie, 2001). For example, the Perry Preschool Program has long-lasting positive effects on non-cognitive skills (Heckman et al., 2013), in contrast to Head Start, a publicly funded program that covers all of the United States of America (e.g. Barnett, 2011).

As one possible intervention for non-cognitive skill development, universal music education programs are increasingly popular among policymakers around the world. Most existing programs are inspired by the famous Venezuelian youth orchestras, known as *El Sistema*, which were established in 1975 by José Antonio Abreu (Creech et al., eds, 2013). These programs, organized by local communities or schools, aim to provide deprived children access to music education. In addition to the regular school curriculum, participating children learn to play a musical instrument for several years. Lessons are taught in small groups or orchestras, with a duration between one or two hours per week in developed countries (Hille, 2010; JeKi-Stiftung, 2014) to several hours per day in the original *El Sistema* program (Creech et al., eds, 2013).

In Germany, the country in the focus for this study, the most prominent example is the *Jedem Kind ein Instrument* program (an instrument for every child, henceforth abbreviated as JeKi). Created in 2007 as a pilot project, JeKi reaches, as of 2015, almost all primary schools in 3 of Germany's 16 federal states: North Rhine-Westphalia, Hamburg, and Saxony. With an annual subsidy of 10 million Euros from North Rhine-Westphalia's federal state government, for instance, and further amounts from other administrations, the program receives substantial public funding (JeKi-Stiftung, 2014).

While cultural partaking is formally stated as the primary objective for JeKi and similar programs, policymakers praise the benefits of universal music education for personality development (e.g. BSB, 2009). *El Sistema* aims at nothing less than "to affect social change through the provision of musical ... opportunities for young people from poor and vulnerable communities" (Creech et al., eds, 2013, p. 17). In their 2013 coalition agreement, the German federal government's ruling parties (CDU, CSU and SPD) stated that, "cultural education is essential for the personality development of young people, as well as their social skills" (CDU, 2013, p. 90). If universal music education has the potential to foster non-cognitive skills, the increasing political willingness to invest in such policies might be an innovative approach to improve educational opportunities, especially for children from disadvantaged social backgrounds.

However, political hopes related to the wide range of benefits that universal music education is believed to have stand in stark contrast to a lack of empirical research on this topic. While (voluntary) extracurricular music training is found beneficial for cognitive skills and is likely to also improve non-cognitive skills (Hille and Schupp, 2015; Schellenberg, 2004; Winner et al., 2013), few researchers study the effects of universal music education policies on non-music related outcomes. In an extensive meta-study of 85 research papers evaluating 44 programs in 19 countries, Creech et al., eds (2013) find mostly qualitative evidence, according to which program participants had higher self-esteem, better listening skills, better social skills, a more pro-social behavior, higher academic aspirations, and better academic achievement. However, most of these effects were found for small and selective samples. Almost all studies were of low academic quality or carried out by program administrators themselves. To my knowledge, not one sufficiently takes endogenous selection into account (Creech et al., eds, 2013).

For the German JeKi program, Roden et al. (2014) find that program participants improved their working memory in comparison to students receiving an intensified natural science curriculum. Adding a second comparison group of non-treated individuals, Roden et al. (2012) find improvements in verbal, but not visual memory. As highlighted by the authors, these papers are based on small and selective samples (50 to 75 observations), which may question the plausibility of the (implicitly made) common trend assumption. The effects on non-cognitive skills are also studied for JeKi. Nonte and Schwippert (2014) find no effects of the program on outcomes such as the self-concept of ability or the feeling of being integrated at school. However, their study also suffers from a small sample size and lacks data on the 65% of children who drop out of the program. Moreover, as they only use data from schools that were willing to participate in the evaluation, they are likely to overestimate the true effect. At the same time, Nonte and Schwippert (2014) only compare program participants to children in schools with an intensified sports curriculum, which might affect these skills as well.

This paper addresses this gap in the literature and answers two research questions related to the effects of the JeKi universal music education program in North Rhine-Westphalia, the most populous German federal state. First, I examine whether JeKi achieves its stated goal to increase access to music education for children from disadvantaged social backgrounds. I compare the effects of the program to those of a much simpler policy, which consists of reducing fees at public music schools. Second, I address the above-mentioned claims by policymakers and investigate whether JeKi also affects children's socio-emotional skills, as measured by the Strengths and Difficulties Questionnaire (SDQ), as well as their attitude toward school, teachers, fellow students and studying. To derive potential hypotheses on how JeKi might affect these traits, I describe a skill production model (Todd and Wolpin, 2003), in which school and family inputs can be more or less productive (Falck et al., 2015). The model illustrates that the potential influence of universal music education is composed of direct effects resulting from program participation, as well as indirect effects of crowding out alternative activities, which children would have carried out in absence of the program.

For the empirical estimations, I use a unique combination of three data sets. As the main data source, I use a repeated cross-section of 8- to 10-year-old children from the German Socio-Economic Panel (SOEP) and Families in Germany (FiD). For a random sample of German households, SOEP and FiD contain a variety of child development measures, as well as detailed information about the parents, including education and labor market participation (Schröder et al., 2013; Wagner et al., 2007). This individual-level data is geographically merged with statistics on the five closest primary schools, as well as the closest public music school.

To identify the causal effects of JeKi participation, I take advantage of the regional and temporal variation in the implementation of the program. Given that participating schools were not chosen randomly, I use a difference-in-differences estimator to eliminate potential time-constant unobserved heterogeneity between participating and non-participating schools. Moreover, to eliminate bias resulting from self-selection into JeKi schools, I estimate intention-to-treat rather than actual treatment effects. To do so, I consider children as treated, if they live close to several schools offering the JeKi program in the year of school enrollment. Furthermore, I include not just numerous individual and regional control variables, but also county fixed effects.

My findings suggest that JeKi successfully reaches children who did not previously benefit from extracurricular music activities. The possibility to attend a primary school offering the JeKi program increases the likelihood of learning a musical instrument at age 9 by 30 percentage points among children from disadvantaged social backgrounds. This amounts to an increase of 115%, given that, on average, only 26% of these children are musically active. Moreover, the possibility to participate in the JeKi program reduces behavioral problems and leads children to have a better relation to their school teacher. These effects are larger for boys and among children from wealthier households. Boys also appreciate school and studying more as a consequence of the JeKi program. All results are robust to variations in measurement and estimation method, as well as to sensitivity tests such as replacing the treatment by placebo reforms and measuring pre-treatment outcomes.

This paper adds to the literature by studying the effects of a large-scale publicly funded universal music education program on music activities and socio-emotional behavior. Contrary to the most closely related work by Nonte and Schwippert (2014), I use a random sample of students living in all areas of the program. The program is, therefore, analyzed in a random rather than self-selected subsample of participating primary schools. Moreover, I thoroughly discuss the potentially non-random selection of students and schools into the program, which is insufficiently taken account of in virtually all previous studies on such policies (Creech et al., eds, 2013). My study proposes an empirical design that credibly identifies causal effects.

Moreover, I contribute to the more general literature on the effects of music training and non-formal education activities. Previous research on the effects of music is either experimental and uses data on self-selected study participants (Schellenberg, 2004), or analyzes observational data and faces the issue of non-random selection into music practice (e.g. Cabane et al., 2015; Covay and Carbonaro, 2010; Hille and Schupp, 2015; Southgate and Roscigno, 2009). By studying arguably exogenous variation in music participation in a randomly drawn observational data set, this study is the first to combine the advantages of both approaches.

Finally, I discuss effect heterogeneities with respect to socio-economic background, an important issue neglected in much of the existing research on music training, which is carried out in disciplines such as psychology, musicology and educational sciences.

The following section describes the JeKi program in more detail, followed by a discussion of its potential effects on non-cognitive skills (Section 5.3). Sections 5.4 and 5.5 present the data and method used for the empirical estimations. Section 5.6 discusses the results and their robustness with respect to measurement, estimation and identifying assumptions.

5.2 Institutional background: Music education in Germany and the JeKi program

Learning a musical instrument is among the most popular education-oriented leisure time activities in Germany. Approximately 44% of 9- to 12-year-olds regularly play music. While many quit until early adulthood, 25% are still musically active between 18 and 25 years of age (Grgic and Züchner, 2013). Serious commitment to music is constantly increasing: the share of 17-year-olds taking musical instrument lessons and playing music at least weekly has risen from 10% to 18% between 2001 and 2012 (Hille et al., 2014). Even the substantial increase in the length of school days caused by a recent educational reform did not decrease the commitment to extracurricular musical activities (Dahmann and Anger, 2014; Grgic and Züchner, 2013).

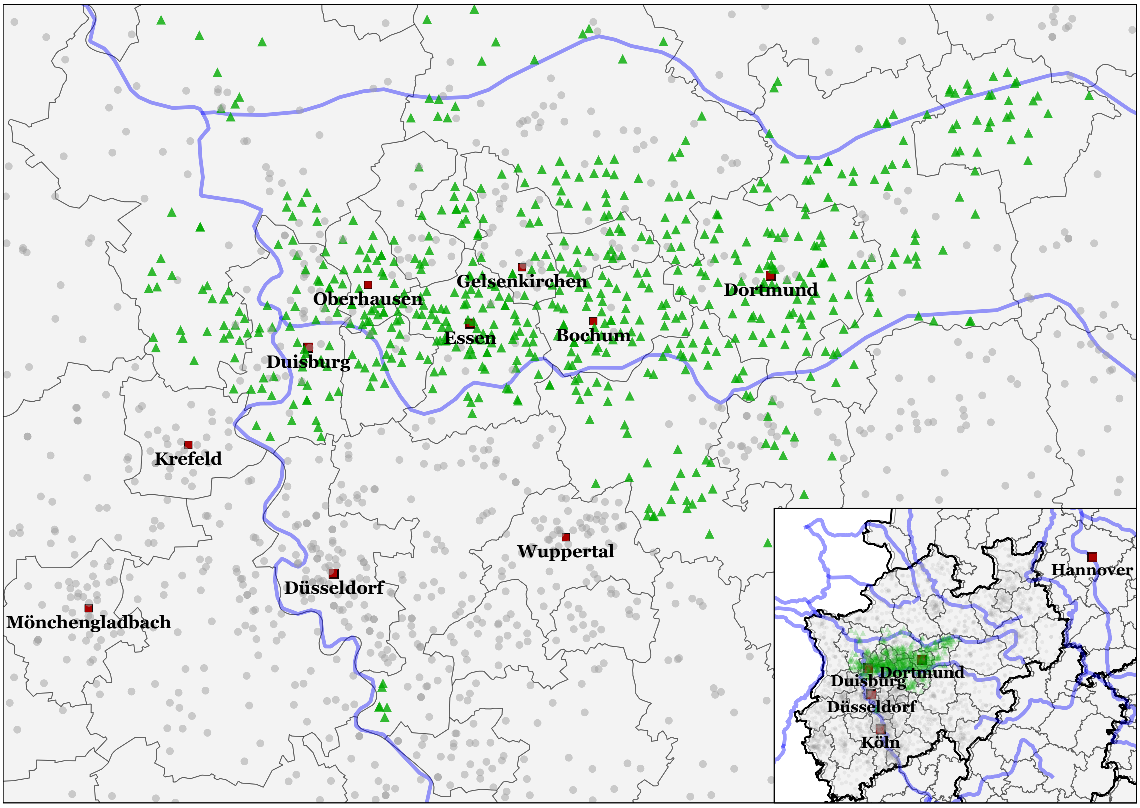
Typically, musical instrument lessons in Germany take place on a voluntary basis and outside the general school system.[23] A network of 929 public and at least 350 private music schools covers the entire country (VDM, 2014; BDPM, 2015). In addition to offering music lessons for individuals and groups, many of these cooperate with general schools to offer voluntary music projects in addition to the regular school curriculum. Partly due to the large expansion of all-day schools, the number of such cooperations between music and primary schools increased by 63% from 2006 to 2013 (MIZ, 2014). In 2013, 4,347 of all 15,749 primary schools officially cooperated with public music schools (MIZ, 2014; Statistisches Bundesamt, 2015).

Despite the broad availability of music education both within and outside the general school system, participation in music activities depends highly on socio-economic background. Children with richer or more highly educated parents are considerably more likely to learn a musical instrument in their leisure time (Hille and Schupp, 2015). In multivariate regressions, parental education predicts participation in music activities more strongly than income (Hille et al., 2014). These socio-economic differences are, therefore, unlikely to solely reflect financial constraints. The German Association of Public Music Schools (Verband deutscher Musikschulen) requires its members to offer reduced fees for low-income families (VDM, 2011). Moreover, the *educational package* – introduced by the German federal government in 2011 – subsidizes sports club or music school membership for welfare recipients with 10 Euros per month (BMAS, 2015a), an amount that is rather low in comparison with the average costs of extracurricular music education.

To make music education more accessible for children from all social backgrounds, several German federal states have established universal music education projects. North Rhine-Westphalia's JeKi – the project in the focus of this study – is currently the largest of these programs in Germany. It was first created as pilot project in 2007 and subsequently expanded to cover large parts of the Ruhr area. With 576 primary schools in 41

[23] The regular school curriculum also includes music education, which may include some active music participation.

Figure 5.1 – *Distribution of JeKi projects in 2011*



Primary schools with (triangles) and without (dots) the JeKi program.
Source: Data from IT.NRW (2011), see Section 5.4, own representation.

municipalities, approximately 20 % of North Rhine-Westphalia's primary schools, JeKi reached its full scale in the 2011/12 school year.[24]

Participating primary schools were chosen as follows. If located in the program's target area, public music schools and municipalities could jointly apply for participation (JeKi-Stiftung, 2015). Among the submitted proposals, the program's coordinating body – the JeKi foundation – selected those that meet some previously defined "program standards" (JeKi-Stiftung, 2011). In other words, music schools and municipalities had to take the initiative to create JeKi projects. Only after their municipality had successfully applied could primary schools choose to participate in JeKi.

Figure 5.1 shows a map of North Rhine-Westphalia's primary schools. Primary schools participating in the JeKi project are represented by triangles, all other primary schools by dots. As described above, all of the federal state's JeKi schools are concentrated in the

[24] In 2015, the federal state government expanded a modified version of JeKi – now called JeKits – to the entire federal state; see www.jedemkind.de.

Table 5.1 – *Description of "Jedem Kind ein Instrument" North Rhine-Westphalia*

Coverage by year				
2007	**2008**	**2009**	**2010**	**2011**
34 municipalities 223 primary schools 7,100 students (first-graders)	41 municipalities 370 primary schools 19,600 students (first-graders)	42 municipalities 522 primary schools 27,700 students (first-graders)	42 municipalities 641 primary schools 31,150 students (first-graders)	42 municipalities 659 primary schools 32,754 students (first-graders)

Description by grade				
	Grade 1	**Grade 2**	**Grade 3**	**Grade 4**
Content	Basic concepts: rhythm, melody	Instrument lessons (small group)	Instrument lessons (small group) Ensemble lessons Performance (end of year)	Instrument lessons (small group) Ensemble lessons Performance (end of year)
Duration	1 extra hour per week with regular school teacher	1 extra hour per week with music school teacher	2 extra hours per week with music school teacher	2 extra hours per week with music school teacher
Participation	Mandatory Participation: 100%	Voluntary Participation: 77%[a]	Voluntary Participation: 49%[a]	Voluntary Participation: 35%[a]
Costs (lessons)	Free	20 €/month[b] (reduction possible)	35 €/month[b] (reduction possible)	35 €/month[b] (reduction possible)
Costs (instruments)	Free	Free	Free	Free

Sources: JeKi-Stiftung (2014); JeKi-Stiftung (2013), own summary.
[a]The share of participating students is taken from Busch and Kranefeld (2013).
[b]100% fee reductions for welfare recipients. 14% of the participants do not pay these fees (JeKi-Stiftung, 2008).

Ruhr area. However, the geographically concentrated distribution of projects is unlikely to simply capture a "Ruhr effect." Numerous schools in that area did not participate in the program, even after the program had reached its full scale in 2011. Moreover, JeKi's expansion between 2005 and 2011 did not follow any systematic geographical pattern.

JeKi involves its participants throughout primary school, from grade 1 to grade 4 (see Table 5.1). The program is in addition to the regular school curriculum. In the first year, students spend an additional hour each week with their regular primary school teacher covering basic musical topics such as rhythm and melody. Musical instrument lessons start in grade 2. Participating children choose their preferred instrument and receive group lessons through grade 4. In the final two years, students additionally play in an orchestra and take part in concerts at the end of the school year. The choice of instruments offered can vary between schools, but it typically includes common string and wind instruments, as well as, in some cases, less typical ones such as mandolin or the Turkish bağlama (JeKi-Stiftung, 2011).

Participation in the first year of the program is mandatory. The program content of grades 2 through 4 is voluntary, with 70% of children taking at least one year of instrument lessons. In grades 3 and 4, participation drops to 49% and 35%, respectively (Busch and Kranefeld, 2013). Starting in grade 2, a monthly fee of 20 Euros is charged for participation, which increases to 35 Euros in grades 3 and 4. However, welfare recipients are exempt from these charges. At this level, music lesson fees within JeKi are lower than what is charged by most public music schools (see Section 5.4). Instruments are loaned to children free of charge and can be taken home.

5.3 Hypotheses: How does JeKi affect non-cognitive skill development?

We can illustrate the potential effects of participating in the JeKi program using a simple model of skill production (Cunha et al., 2010; Todd and Wolpin, 2003), in which JeKi (J) enters the skill production function in addition to the usual family and school inputs (F and S). Similar to Falck et al. (2015), each input can be of high (h) or low (l) productivity for skill development. The production of non-cognitive (and cognitive) skills Y is thus described by the following equation:

$$Y = f(S_p, F_p, J_p) \quad \text{with} \quad p \in \{h, l\} \tag{5.1}$$

As described above, JeKi takes place in addition to the regular school curriculum and – given the fixed total amount of time – partly replaces family (or leisure time) inputs F_p. If we abstract from the school and family inputs received irrespective of program participation, the change in skill production due to JeKi ($\Delta_J Y$) is:

$$\Delta_J Y = f(J_p, \cdot) - f(F'_p, \cdot) \tag{5.2}$$

where family input F'_p is crowded out by the program. The treatment effect can therefore be decomposed into the direct effect of the program ($\partial f / \partial J_p$) minus the skill production effect of the activity, which is crowded out ($\partial f/\partial F'_p$). I discuss each of these in the following.

The potential direct effect is activated by the program's two main components, which are playing a musical instrument and taking lessons in small groups. Learning to play a musical instrument is found beneficial for skill development in previous research. For example, the experimental study by Schellenberg (2004) shows that music training improves intelligence, at least in the short-run. Moreover, Hille and Schupp (2015) find that learning a musical instrument throughout childhood is associated with better school grades, as well as higher conscientiousness, openness and ambition. In their study, outcome differences between musically active and inactive adolescents remain positive and strongly significant even after controlling for a large number of parental background characteristics. Even compared to playing sports, music training is associated with better educational outcomes (Cabane et al., 2015). While there are many hypotheses according to which music affects further non-cognitive skills as well, no empirical study proves that this is the case (Winner et al., 2013).

Taking lessons in small groups might teach JeKi participants to closely interact with their peers and subordinate individual interests for the benefit of a common goal (Covay and Carbonaro, 2010). As a group activity, JeKi participation might therefore raise self-esteem and improve social skills (Lareau, 2011). More than in all other school subjects and many voluntary extracurricular activities, children participating in these small group lessons have to closely interact with their teacher, and might thereby improve their interpersonal skills. In addition, children might develop an improved sense of entitlement, which is the ability to speak up for something that one believes to deserve (Lareau, 2011).

As a third mechanism of program participation, we could consider the particular attention students receive from being part of a new and widely debated project. JeKi received substantial attention in the media and from politicians, with the hope that it would show "a positive transmission ... to the entire Federal Republic of Germany" (Kulturstiftung, 2015). Participating in a project, which generates such expectations, could positively or negatively affect students' attitude toward school.

According to the model described above, the average effect of JeKi participation also depends on how leisure time would have been spent in absence of the program. For example, Nonte and Schwippert (2014) compare JeKi participants to students attending schools with a reinforced sports curriculum and do not find any effects for the outcomes they study. JeKi could crowd out activities that are more or less (or just as) productive with respect to their effect on skill development. For example, non-participating schools might propose alternative programs involving music or sports. Alternatively, children might attend a sports club or music school on a voluntary basis, if they lack the possibility to take music lessons at their school within the JeKi program. On the other hand, the time spent with JeKi could have otherwise been spent by, for example, playing computer games or hanging out, with uncertain consequences for personality development.

The choice of leisure time activities strongly depends on socio-economic background. For example, Hille and Schupp (2015) find that, on average, children who learn a musical instrument outside of school have better educated and richer parents, who are more involved with their child's school success, who attend cultural events, and who carry out artistic activities themselves. Similar results are found for other activities (Cabane and Lechner, 2015). Therefore, I expect the average treatment effect of JeKi to be stronger for children from less favorable socio-economic backgrounds, because their counter-factual use of time is likely to be less productive for skill development than that of children from advantaged social backgrounds.

On the other hand, JeKi is voluntary after the first year. If children from less advantaged households are more likely to quit the program, effects are smaller for them compared to those who carry on through the end of grade 4. However, Busch and Kranefeld (2013) find that neither socio-economic characteristics nor migration background significantly predict whether children continue the program in grades 2 to 4.

In addition to the direct and indirect effects described so far, we might expect interactions between JeKi participation with school and family inputs productivity, in the sense of the dynamic complementarity described by Cunha and Heckman (2007). For example, JeKi might increase students' motivation for school and, thereby, the effectiveness of learning. Moreover, positive experience in the JeKi program might encourage students to continue practicing their musical instrument after grade 4. Depending on the productivity of the activity replaced, this could further contribute to the program's skill production effects.

Even though music education might affect various dimensions of cognitive and non-cognitive skills (Winner et al., 2013), the most relevant dimensions JeKi potentially improves are behavior, as well as the children's ability to get along with teachers and other children. The Strengths and Difficulties Questionnaire (SDQ) is particularly well-suited to assess such improvements (Goodman, 1997). I expect improvements in three of its five dimensions, which are related to social skills: a reduction in conduct problems, a reducation in peer problems, and increased pro-social behavior. The other two dimensions of the SDQ – hyperactivity and emotional symptoms – are probably less affected. Additionally, the JeKi program is likely to influence how children think about their school and whether they get along well with their teachers and peers.

These non-cognitive traits are important indicators of child development. A lack of social skills during childhood, especially with respect to adults, increases the likelihood of becoming a school dropout and of teenage pregnancy, while also negatively affecting adult labor market outcomes (Carneiro et al., 2007). Misbehavior in the classroom is also negatively related to educational attainment and earnings among young men (Segal, 2008, 2013), and is even considered as a possible source of the gender gap in academic achievement (Bertrand and Pan, 2013). Moreover, (classroom) misconduct correlates with lower conscientiousness (Ones et al., 2007; Roberts et al., 2003) and higher discount rates (Castillo et al., 2011), two important determinants of lower educational and labor market success.

5.4 Data

5.4.1 Description of the data

This study uses a unique data source, combining geo-coded information on primary schools and public music schools with individual-level data on 8- to 10-year-old children from the German Socio-Economic Panel (SOEP) and Families in Germany (FiD). By combining these data, I obtain information on each child's possibility to participate at the JeKi program or, alternatively, to attend music lessons at a nearby public music school. In the following, I describe the three data sets and how they are combined. For additional details, please refer to Appendix D.1.

First, I obtained a complete directory of North Rhine-Westphalia's primary schools for the year 2011, which is available from the federal state's statistical office (IT.NRW, 2011). It includes each school's address and total number of students. I created two additional variables indicating whether the school offers the JeKi program and, if applicable, the year it first participated. I received this information from the JeKi foundation, which coordinates the program.[25] A very similar project in the municipality of Monheim ("Monheimer Modell") is also included in the analyses of this paper. None of the schools withdrew from the program during the relevant survey years.

Second, I use detailed statistics on all German public music schools from 2008 and 2012, as published in the statistical yearbook of German municipalities (Deutscher Städtetag, ed, 2009, 2013).[26] These contain variables such as the number of teachers and students (actual and full-time equivalent), number of cooperations with general-education schools, fee revenues, revenues from state subsidies, as well as personnel expenditures.

The third, and main, data set contains individual-level data on 8- to 10-year-old children from two German household panel studies: the German Socio-Economic Panel (SOEP) and Families in Germany (FiD).[27] The SOEP was created in 1984 and, as of 2015, surveys a random sample of approximately 12,000 German households every year (Wagner et al., 2007). FiD data has been collected since 2010 as part of a large evaluation of family-policies in Germany. It is closely related to the SOEP, but concentrates on poor families, families with many children, and single parents; it also includes cohort samples of families with children born between 2007 and 2010. FiD surveys approximately 4,000 households (Schröder et al., 2013). Due to a large overlap in questionnaire items, these studies can be jointly analyzed.[28]

With a household questionnaire, as well as a variety of individual questionnaires, SOEP and FiD survey information from all household members, covering many dimensions of life, such as education and labor market participation, family life, as well as personality traits, opinions and values. In addition, parents receive specialized questionnaires every two or three years, in which they provide information about their child's development, as well as their strategies and attitudes toward child education. These specialized questionnaires are an excellent data source for this paper, as they contain all the variables necessary to test the hypotheses developed above.

[25] The list of JeKi schools can also be found on the program's website (www.jedemkind.de).

[26] The German Association of Public Music Schools (*Verein deutscher Musikschulen*) generously provided me these statistics in Excel files.

[27] Socio-Economic Panel (SOEP), data for years 1984-2013, version 30, SOEP, 2015, doi:10.5684/soep.v30.

[28] Starting from survey year 2014 (SOEP wave 31 and FiD wave 5), both data sets are distributed together under the label SOEP (v31).

Given each child's place of residence, I calculate the distance to and merge the characteristics of the nearest public music school and the five nearest primary schools, using the data sets described above. Using the geo-coordinates of the household's postal code, I measure the geodesic distance to the exact geo-coded addresses of primary and music schools.[29]

The combined data set allows me to construct all of the variables needed for the empirical part of this study, as described in the following. To define treatments, I use the information on whether each of the closest five primary schools participates at the JeKi program, in which year they first offered the program, as well as the number of students at each school. Another treatment variable is based on the fee revenues and number of students at the closest public music school. The exact treatment definitions are explained below.

This study examines two types of outcome variables. The first is related to JeKi's primary goal of providing children from disadvantaged backgrounds with access to music education. To assess whether this happens, I use two variables from the SOEP and FiD household questionnaire that ask parents whether their children attend a music club at school, and whether they play music outside of school. SOEP households have received these questions biannually since 2004, while FiD households receive them every year (2010 to 2013 in our data). As parents of JeKi participants might categorize their child's musical activities as being at or outside school, I created a combined variable indicating whether a child plays music at or outside of school. In one robustness check, I examined whether the definition of children's music activities matters for the results (see Section 5.6).

Further outcomes are taken from SOEP and FiD's parent-child questionnaires, which measure the two categories of non-cognitive skills that might be affected by the JeKi program. These include the Strengths and Difficulties Questionnaire (SDQ), as well as the child's attitude toward school as reported by parents in the year of their child's eighth and tenth birthday.[30] The SDQ describes whether children experience emotional symptoms, conduct problems, hyperactivity, peer problems, and to what extent they show pro-social

[29] A specially secure procedure in accordance with strict German data privacy legislation allows on site researchers at DIW Berlin to merge geo-information with SOEP households using their addresses at the street-block level (Goebel and Pauer, 2014). An application of this procedure is in Bauernschuster et al. (2014). However, due to a substantial amount of missing geo-coordinates, I decided to reduce precision and merge distances at the postal code level. As the aim is to select the primary and music schools individuals are most likely to attend, this degree of precision is probably sufficient.

[30] In the SOEP, the relevant variables are only available for age 10.

behavior (Goodman, 1997). Each of these five dimensions is assessed with two to five items, in which parents rate their approval to descriptions of the child's behavior on a 3-point Likert scale (Richter et al., 2013). Most studies take the sum of the difficulties (all variables except pro-social behavior) as a measure of SDQ. In addition, I also study each dimension separately.

The variables related to the attitude toward school each consist of one questionnaire item. On a 4-point Likert scale, parents were asked to judge whether their children (1) like to go to school; (2) like to study; (3) get along well with teachers; (4) get along well with classmates; and (5) follow lessons well. Table D.1 provides the exact wording of each item.

All dimensions of the SDQ and the attitude toward school were coded such that a higher value indicates a better outcome. Moreover, all variables were normalized to have mean 0 and a standard deviation of 1 within the control group. In addition, I created an aggregate index that is a weighted average of all non-cognitive outcome dimensions described above, following an approach suggested by Anderson (2008), with the aim of reducing false rejections of the null hypothesis due to multiple inferences. The aggregate index is an average, in which each normalized outcome is weighted by the inverse of its covariance matrix with respect to all other outcomes. With other words, variables receive a higher weight if they add more information to the aggregate index.

All outcome variables are measured when the child was 8 and 10 years old. As not everyone is part of the sample at both ages, and to increase the sample size in the baseline specification, I aggregated both age groups. To do so, I measured the outcome at age 10 if available, otherwise at age 8.[31] However, the results were also analyzed separately for the sample of 8- and 10-year-olds.

Covariates are mainly taken from the individual and household questionnaires of the child's parents. I observe gender, migration background and the age of school enrollment for the child. Moreover, covariates include the mother's age at birth, upper secondary school and university education of the parents, as well as their combined weekly work hours. In addition, I control for household income, household size, whether the child lives in a single-parent home, and whether the household lives in a rural or urban area. To take into account regional heterogeneities, I include characteristics of the county of residence,

[31] In this aggregated group, outcomes were measured at age 8 for about a third of the children.

which are available from the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR, 2015). These include GDP, household income, child poverty and youth unemployment rate, level of education, share of foreign students, as well as the share of children in daycare, all measured in 2012. Table D.1 presents the full list of control variables and how they are measured.

5.4.2 Estimation sample and treatment definitions

The sample used in this study consists of all children for whom all outcomes in at least one of the two outcome groups – music activities or non-cognitive skills – are observed. Therefore, to be in the sample, parents must have answered either the household or the parent-child questionnaires noted above. Being in the sample depends on birth year and questionnaire design, which are unrelated to any of the treatment or outcome variables.[32] Restricting the sample to North Rhine-Westphalia, the final data contains 1138 children born between 1997 and 2005. Among these, I have information on music activities for 997 and the non-cognitive skills for 689 children. Results are robust to only examining the 548 individuals with valid observations for all outcomes in both categories.

For 7% of the observations, I do not observe the parents' work hours. The information for the mother's age and the child's migration background are missing in 2% and 1%, respectively, of the cases. In addition, characteristics of the local public music school are missing for almost 3% of the sample.[33] Given the small sample size and the importance of these control variables, I recoded missing observations to 0 for binary and to the mean for continuous variables, and included four missing indicators in all estimations (one indicator for each of the first three variables, and a fourth indicator for at least one missing music school characteristic). However, all estimations are robust to restricting the sample to observations without any missing covariate. In this case, the sample sizes are 881 for music activities and 603 for non-cognitive skills.

The treatment group consists of children who had the possibility to participate in the JeKi universal music education program. To avoid bias resulting from selection into specific

[32] I lose 5% of the observations due to missing outcomes in the non-cognitive skills sample. Moreover, I lose 2% of the observations due to a missing treatment status. See Table D.2 for details on how the sample was constructed.

[33] To prevent missing values on the household income, I use the first plausible value of the imputed variable contained in the SOEP and FiD data distribution.

Table 5.2 – *Construction of treatment variables (main definitions[a])*

Treatment 1: JeKi share above 50%	
Steps carried out	*Details*
1. Identify potential primary schools (target schools)	→ select the 5 primary schools closest to household's place of residence → delete schools at a distance of more than 5km
2. Check whether JeKi was available in school enrollment year	→ do target schools have JeKi program? (if not: JeKi not available) → check whether JeKi was created in the year of school enrollment or before (JeKi available), or in a later year (JeKi not available)
3. Calculate JeKi share among target school students	→ add the total number of students at each of the target schools offering JeKi in the year of school enrollment → divide this number by the total number of students in all target schools
4. Transform into binary variable	→ 1 if JeKi share $\geq$ 50% 0 if JeKi share $<$ 50%
Treatment 2: Music school fees below 33rd percentile	
Steps carried out	*Details*
1. Calculate fees for each school	→ "total fee revenues" / "total number of students" for each music school (average fees, ignores potential income-based reductions)
2. Calculate 33rd percentile	→ among all music schools of North Rhine-Westphalia
3. Transform into binary variable	→ Identify the public music school closest to the child's home (target music school) → 1 if target music school fees $<$ 33rd percentile within federal state 0 if target music school fees $\geq$ 33rd percentile within federal state

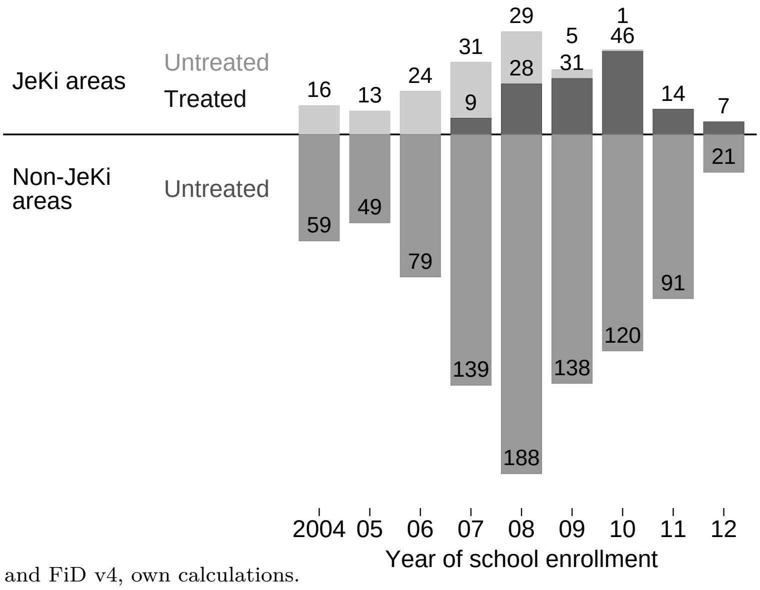
[a]Alternative treatment definitions were tested as robustness checks.

primary schools, I define treatment based on the probability to take part in the JeKi program. For that purpose, I assume that children attend one of the five closest primary schools within a radius of 5km.[34] Among these schools, I calculate the share of students having access to the JeKi program in the year of school enrollment. I consider children as treated if that share is at least 50%. Table 5.2 explains the exact steps of how the treatment was defined. In various robustness checks, I vary this definition and consider children as treated if the share of JeKi students is at least 5% or take the JeKi share as a continuous treatment variable. In further variations, I also consider schools farther than 5km away, ignore schools outside the municipality, and weight schools by distance to the child's home. Please refer to Section 5.6, as well as Appendix D.2 for more details.

Figure 5.2 describes the sample size by treatment status for each school enrollment year in the overall sample of individuals with observations for at least one outcome group.

[34] See Section 5.5 for a discussion of that assumption. I chose a cutoff of 5km, because no child in the sample lives more than 5km away from the closest primary school.

Figure 5.2 – *Number of observations in treatment and control group for each school enrollment year*



Source: SOEP v30 and FiD v4, own calculations.

The sample is split into JeKi and non-JeKi areas, given that the estimations rely on difference-in-differences (see Section 5.5). Non-treated individuals in JeKi areas attend schools that offer the JeKi program in later years, but not the year they enrolled, while treated individuals attend schools that already offer the program. Figure 5.2 shows that the first JeKi programs were introduced in 2007. In 2009, only 5 individuals in the sample attended schools that would start the program in a later year. Figure D.1 shows the same graph for each outcome sample separately.

As described in Section 5.3, the potential effects of JeKi depend on the counter-factual use of time. Therefore, I define an alternative treatment, representing one possible outside option for learning a musical instrument. I define children to receive the alternative treatment, if the public music school closest to their home charges fees in the lowest third of the public North Rhine-Westphalia music schools. This is a rather crude measure of music school fees, given that the German association of public music schools asks its members to request lower fees from low-income families (VDM, 2011). As described above, the data on public music schools only contain fee levels for 2008 and 2012. I assigned the earlier fee level to all children born before 2000, while those born in 2000 or later are assumed to pay 2012 fees. Given that the treatment variable is binary (fee levels below the 33rd percentile), the treatment status is not sensitive to such assignment decisions. Few public music schools changed their fee levels or their rank in the fee distribution between 2008 and 2012. Table 5.2 explains the details of how the music school fee treatment variable is constructed. As with JeKi participation, I propose various variations of this treatment, which are described in detail in Section 5.6, as well as Appendix D.2.

To model the trade-off between JeKi and music lessons at the local music school even more explicitly, I define a third treatment, combining both of the previously described variables. The combined treatment can take three values: (1) the JeKi share is above 50%, (2) the JeKi share is below 50% and music school fees are below the 33rd percentile of the fee distribution and (3) the JeKi share is below 50% and music school fees are above the 33rd percentile. Similar to the multiple treatment setting in Lechner (2001), I compare the differential effects by performing pairwise comparisons.

Average music school fees in North Rhine-Westphalia vary between 13 and 59 euros per month. Approximately 77% of the children who have access to JeKi live close to a music school that charges low fees (see Figure D.2). This results from the fact that public music schools are the primary school's cooperating partner for most of the JeKi projects. JeKi participants therefore count toward the total number of students at the music school. As JeKi lessons are partly financed through public subsidies, they contribute less to the music school's fee revenues, making their average fees decline. Nonetheless, many public music schools charge fees at a similarly low level, even though they do not act as a cooperating partner for JeKi. At the same time, 23% of the JeKi projects were implemented in areas, where the public music school charges relatively high fees.

Further summary statistics on all treatment, outcome and control variables, as well as detailed descriptions of how each variable was constructed, are found in Appendix D.1.

5.5 Empirical strategy

5.5.1 Identification of causal effects

This paper examines the causal effects of the JeKi program on music activities and non-cognitive skills. Moreover, I estimate how low music school fees affect participation in music lessons. The identification of causal effects is complicated by the fact that neither the schools that offer the JeKi program nor the program participants themselves are chosen randomly. The same is true for the alternative treatment of low public music school fees, which might be determined as a function of local demand. Such non-random selection biases the estimated treatment effects if unobserved factors affect the individual's assignment to the treatment group and simultaneously influence the outcomes of interest.

Endogeneity might occur both at the school and the individual level. On the one hand, both music school fees and JeKi participation might be driven by unobserved heterogeneity between primary or music schools. For example, we would overestimate the true effect of JeKi, if the program was systematically implemented by particularly motivated school principals. These school principals might be better able to run their schools. In addition to implementing JeKi, they might, for example, hire better quality teachers or provide better learning environments. Our estimates would wrongly count these additional quality inputs as causal effects of the JeKi program. Conversely, we might underestimate the true effect of JeKi, because non-participating schools might propose alternative activities involving music, sports, theater or any other activity. Rather than comparing schools with and without JeKi, we would compare JeKi schools to those with alternative programs, which might be beneficial as well.

Simultaneity could be a second school-related reason for biased estimates of the treatment effect. As the program is systematically implemented in poorer areas (see Section 5.6), we are likely to underestimate its effects due to systematic differences in the ability of the student body. The same is true for public music school fees, which are also more likely to be low in poorer areas.

On the other hand, selection bias could result from individual unobserved heterogeneity. For example, motivated parents could enroll their children at JeKi schools, or even move to JeKi areas. Similarly, parents could send their children to less expensive music schools farther away from home. Therefore, the estimated treatment effects would capture parental engagement, which benefits child development irrespective of music activities.

I take these potential sources of endogeneity into account by estimating intention-to-treat rather than treatment effects. Moreover, I exploit the regional variation of both the JeKi program and low music school fees. For the former, I can even take advantage of the temporal variation of program implementation and estimate difference-in-differences. Finally, I include a large number of individual- and county-level control variables, as well as county fixed effects. In the following, I describe how these approaches address the endogeneity problem, thus allowing for the identification of causal effects of both treatments.

To reduce the bias resulting from non-random selection into treatment by individuals, the results of this paper describe intention-to-treat effects of JeKi and low music school fees. Given that participation beyond the first year of the JeKi program is voluntary, actual enrollment would be an endogenous treatment, which could be related to individual-level characteristics that also affect the outcomes. I therefore consider children as treated, if they have the possibility to participate, irrespective of whether they actually sign up. As described in Section 5.4, individuals belong to the treatment group if the share of students with access to the JeKi program in the five primary schools closest to child's home (within a radius of 5km) in the year of school enrollment is greater than 50%. Similarly, living near a public music school that charges low fees can be considered as an intention-to-treat effect, because low fees only affect children in the sense that they provide a higher incentive to sign up for music classes.

The effect of having the possibility to learn a musical instrument – the "intention to treat" – is relevant for two reasons. First, it is the parameter policymakers can influence. Policymakers need to know whether creating a voluntary program has positive effects, and these depend on whether individuals actually sign up. Second, the intention-to-treat effect includes the potentially negative effects for non-participants and thereby describes a more complete picture of the program's overall effects. Some individuals might suffer from non-participation. They could receive a lower quality education if schools decide to concentrate their resources on program participants. Moreover, they could feel excluded and, consequently, reduce effort. The intention-to-treat effect takes such potential peer effects on non-participants into account by aggregating the effect of participation and non-participation among all individuals who have the possibility to participate.

While estimating intention-to-treat effects, I cannot rule out the possibility that parents from particular socio-economic backgrounds specifically select JeKi schools for their children in areas where only some of the five neighborhood schools offer the program.[35] If this is the case, my estimations capture the treatment effect of JeKi on those who choose to participate. With other words, individual selection into schools within neighborhoods

[35] In the sample, 15% of the children are surrounded by some schools with and some without the JeKi program; only 4% have access to the program in all of the five surrounding schools.

would lead me to estimate average treatment effects on the treated rather than average treatment effects.

In addition to focusing the estimates on the "intention to treat", I use regional variation to identify causal effects. As described in Section 5.2, JeKi exists in some areas of North Rhine-Westphalia and not in others. Similarly, average monthly fees at the federal state's public music schools range from about 12 to 60 Euros (see Section 5.4). Assignment to the treatment group therefore depends on the individual's place of residence.

Still, both treatments are not randomly distributed across regions (see Section 5.6). Therefore, I further refine the estimations by additionally taking advantage of temporal variation. For JeKi, this amounts to a difference-in-differences estimator. As described in Section 5.2, the program was gradually implemented starting in 2007 until it reached its full level of expansion in 2011. Given that I observe individuals in JeKi areas before and after the program was implemented, I can take into account time constant unobserved differences between areas with and without JeKi. To do so, I estimate the effect of having access to JeKi in the year of school enrollment and additionally control for time fixed effects, as well as a dummy for JeKi areas. This dummy is set to "1" if the "JeKi share" is higher than 50% in any year, even if JeKi is only implemented in a later year and the child thus not treated.

For the other treatment of low music school fees, a difference-in-differences estimator would theoretically be possible as well. I observe public music school fees in 2008 and 2012, and merge the information for the year that is more relevant for each child (see Section 5.4). By including music school fixed effects, I would identify treatment effects from the variation that arises if certain schools change their fees. However, music school fees were relatively stable between both years. Given that the treatment is defined as a binary variable indicating whether the fees were below average, only very few schools changed treatment status. To account for potential music school-specific unobserved heterogeneity, I therefore included various music school characteristics as control variables. These contain information on the number of students and school cooperations, total revenues and the share of federal state subsidies, as well as the share of students enrolled in basic classes, instrument and ensemble lessons. Finally, to capture any remaining unobserved heterogeneity, I include county fixed effects, as well as covariates describing characteristics of the child, the parents and the county of residence. Please refer to Section 5.4, as well as Appendix D.1 for further details and summary statistics of all control variables.

With the identification strategy described so far, I identify causal effects under two assumptions. First, I need to assume a common trend, according to which JeKi and non-JeKi areas would have evolved similarly with respect to the outcomes, had the JeKi program not existed. Second, I assume that families did not move in order to live closer to a JeKi school. North Rhine-Westphalia abolished primary school catchment areas in 2008, allowing parents to freely choose their child's primary school. However, numerous exceptions were granted even before 2008, *inter alia* to facilitate leisure time activities for children (Riedel et al., 2010). Moreover, the student composition with respect to migration background of the state's primary schools has not changed due to the reform (Makles and Schneider, 2015). Given the free choice of primary schools, as well as the numerous possibilities to learn a musical instrument in all regions of Germany, I argue that families are unlikely to move for the purpose of having access to the JeKi program or living closer to a cheap public music school. Whether these assumptions are plausible is discussed in Section 5.6, along with some robustness tests.

5.5.2 Estimation

In comparison to semi-parametric methods, such as propensity score matching, OLS can lead to biased coefficients due to the necessity to impose functional form restrictions. This is particularly true if the distribution of covariates differs strongly between treatment and control group (Imbens, 2015). However, to obtain correct standard errors for coefficients obtained by a propensity score matching estimator, one needs to take into account the two-step estimation procedure. Even though an analytical approximation exists, standard errors are usually estimated from bootstrap replications (Huber et al., 2015). Given the small sample size, and especially the small number of treated individuals (94 out of 689 in the smallest sample for non-cognitive skills), the bootstrap is not reliable in drawing samples with sufficient variation in the treatment status. Therefore, and in order to be in line with most other papers using difference-in-differences estimators, the main specifications were estimated using ordinary least squares (OLS). In any case, the distribution of covariates is rather similar between treatment and control group. However, the results are robust to changing the estimation procedure to propensity score matching or entropy balancing.

In the main specification, the effect of JeKi on outcome Y is estimated using the following estimation equation:

$$Y_i = \beta_1 T_i + \beta_2 G_i + \tau_{year} + \beta_3 X_i + \beta_4 X_{county} + \eta_{county} + \varepsilon_i \tag{5.3}$$

The first three terms on the right-hand side correspond to the standard difference-in-differences estimator. β_1 is the coefficient of interest, describing the effect of treatment T_i, indicating whether the share of children with access to the JeKi program in the nearest primary schools is above 50% in the year of school enrollment. With other words, it is an interaction between living in a treatment region and being actually treated. β_2 captures the time-constant unobserved heterogeneity between JeKi and non-JeKi areas, given that G_i is an indicator of whether the child is surrounded by schools, in which at some point in time the share of students having access to the JeKi program is or will be above 50%. τ_{year} are fixed effects for the year in which individual i enters primary school. They cover cohort effects common to treatment and control group. The remaining terms in the estimation equation contain the control variables described above: characteristics of child and parents (X_i), county-level covariates (X_{county}), as well as county fixed effects (η_{county}). In all estimations, standard errors are clustered at the county level.

5.6 Results

5.6.1 Effects of the JeKi program on access to music education

JeKi was created with the goal of providing access to music education, in particular, for children from less favorable social backgrounds. Table 5.3 shows that JeKi was indeed implemented in poorer regions. Households in areas with a JeKi coverage of at least 50% earn on average 170 Euros less every month than those situated elsewhere. The child poverty rate in JeKi areas is almost 10 percentage points and the share of foreign pupils almost 4 percentage points higher. On the individual level, these differences are similar. Only 23% of the children in JeKi areas have parents with a university degree, compared

Table 5.3 – *Individual and county-level background characteristics by treatment group status*

	JeKi share > 50%				Music school fees in lowest third			
	No	*Yes*	*Diff.*	*p*	*No*	*Yes*	*Diff.*	*p*
County-level characteristics								
Av. monthly HH income/person (Euros)	1739	1568	-171	0	1724	1651	-73	2
Yearly GDP/person (1000 euros)	34.0	30.9	-3.0	26	33.5	32.9	-0.5	79
Share students with Abitur (%)	36.3	36.6	0.3	78	36.7	35.6	-1.1	22
Share students without degree (%)	4.8	6.0	1.2	0	4.9	5.4	0.5	5
Youth unemployment rate (%)	19.6	18.7	-0.9	9	19.6	18.9	-0.7	5
Share foreign pupils (%)	7.9	11.7	3.8	0	8.2	9.9	1.7	4
Child poverty rate (%)	15.9	25.4	9.6	0	16.7	20.8	4.1	2
Individual-level characteristics								
Monthly net HH income (Euros)	3181	2947	-234	13	3183	3012	-171	25
Parent with Abitur (%)	33.4	34.3	0.9	85	35.2	30.1	-5.1	18
Parent with university degree (%)	29.4	23.2	-6.2	12	30.2	23.4	-6.8	6
Migration background (child, %)	35.6	33.1	-2.6	51	34.5	36.2	1.7	67
Only one parent at home (%)	14.8	16.9	2.1	40	14.6	16.7	2.1	34
Nb of children under 16 in household	2.6	2.6	0.0	94	2.5	2.7	0.2	6

Source: SOEP v30 and FiD v4, own calculations. T-test of differences between selected covariates in treatment and control groups for both treatments. Standard errors are clustered at the county level.

to 29% of the children living elsewhere.[36] The last four columns of Table 5.3 show that music school fees are lower in poorer areas as well. However, these differences are only about half the size as those for JeKi.

Given that participation in the JeKi program is voluntary after the first year, merely implementing JeKi in poorer areas is not sufficient to ensure that children from disadvantaged backgrounds actually start learning a musical instrument. To further investigate this hypothesis, Table 5.4 shows the causal effects of access to the JeKi program using the identification strategy described in the previous section. The first part of the table investigates the effects of JeKi on music activities and compares them with those of the alternative policy of low music school fees. Column (1) displays the effects for the full sample, while columns (2) to (6) show subgroup analyses, distinguishing children by their socio-economic status and gender. Each cell represents a distinct regression. Coefficients display the effect of access to the JeKi program (except for the second line, where the treatment is low music school fees) on the outcome indicated in the first column. Sam-

[36] Differences are less significant for individual than for county characteristics, simply due to the fact that the variation is at the individual level, while in the upper part of Table 5.3 the characteristics only vary across the 55 counties of North Rhine-Westphalia.

ple sizes by treatment and control group, as well as the R-square for each regression are presented in Tables D.10 and D.11.

The first part of Table 5.4 shows that JeKi can be considered successful with respect to its primary objective. On average, having access to the JeKi program increases the likelihood to play a musical instrument at or outside of school at the age of 9 by 14 percentage points (first row of column 1). With an average of 36% being musically active at that age, this amounts to an increase of almost 40%. The effect is considerably stronger for girls than for boys (columns 2 and 3). Music schools that charge fees in the lowest third of the fee distribution increase the likelihood of playing a musical instrument to a lower extent, by 10 percentage points (row 2).

Examining the heterogeneity of these effects by socio-economic status illustrates that the JeKi program is much more successful than low music school fees in reaching disadvantaged children. While both affect children from poorer families (column 4), the effect of JeKi is particularly strong, reaching 30 percentage points. This amounts to an increase of 115%, if we consider that on average only 26% of the children from households with a below-median income are musically active. Table 5.4 distinguishes according to household income, but the effects are similar (albeit somewhat smaller) when defining socio-economic status by parental education or migration background (see Table D.13). This is consistent with the hypothesis that children from richer households might engage in music activities even in absence of the JeKi program. Children from poor families might be less affected by variations in music school fees, as they are supposed to receive fee reductions in any case, as stated in the governing rules of the German association of public music schools (VDM, 2011).

5.6.2 Effects of the JeKi program on non-cognitive skills

As stated in the introduction, many policymakers claim that JeKi is an important policy not only for providing access to music education, but also for the personality development and educational success of its participants. According to the characteristics of the program (small group lessons in addition to the regular school curriculum), JeKi is particularly likely to affect social skills and school attitude, given that children learn to interact with peers and teachers in groups that are much small than regular school classes (see Section 5.3). The second part of Table 5.4 shows the effects of the JeKi program on these and

Table 5.4 – *Effects of access to the JeKi program (OLS estimates)*

	Full sample		by gender				by household income			
			Male		Female		< median		> median	
	(1)		(2)		(3)		(4)		(5)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*
MUSIC ACTIVITIES										
Plays music	0.14	2	0.06	52	0.20	3	0.30	0	0.04	61
Alternative treatment: Music school fees in lowest third										
Plays music	0.10	5	0.17	1	0.03	68	0.16	5	0.06	35
Sample size	997		502		495		443		554	
NON-COGNITIVE SKILLS										
Strengths and Difficulties Questionnaire										
No emotional symptoms	-0.04	84	0.02	91	-0.21	49	-0.03	92	-0.19	51
No conduct problems	0.40	3	0.38	14	0.59	9	0.26	17	0.59	8
Not hyperactive	0.15	37	0.28	33	0.32	29	-0.06	78	0.34	26
No peer problems	-0.03	86	-0.23	49	-0.03	88	-0.10	73	0.07	85
Pro-social behavior	0.19	33	0.03	92	0.18	62	0.21	51	0.28	44
Attitude towards school										
Likes to go to school	0.25	14	0.53	5	-0.06	84	0.23	35	0.21	52
Likes to study	0.34	12	0.94	0	0.02	95	-0.03	95	0.65	1
Follows lessons well	0.04	86	0.23	63	0.09	75	-0.59	17	0.71	0
Gets along well...										
...with classmates	0.14	46	0.00	99	0.09	76	0.26	34	0.16	54
...with teacher	0.27	2	0.52	7	-0.02	95	0.03	93	0.50	5
Combined non-cognitive skills scores										
Non-cog. skills (aggregate)	0.17	4	0.29	8	0.10	44	-0.00	97	0.34	2
Difficulties (from SDQ)	0.18	41	0.19	51	0.26	42	0.03	90	0.31	41
Sample size	689		341		348		327		362	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. All regressions are estimated using difference-in-differences, and additionally control for individual and regional characteristics, as well as county fixed effects. Effects of the alternative treatment (music school fees in lowest third) are estimated with standard OLS (no difference-in-differences), additionally including music school characteristics as control variables. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression are displayed in Tables D.10 and D.11. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

other dimensions of the Strengths and Difficulties Questionnaire, as well as the attitude toward school.

Overall, JeKi seems to positively affect most items from the Strengths and Difficulties Questionnaire and the attitude toward school, even though standard errors are often too large to yield significant coefficients. For the full sample, two effects are highly significant: Children with access to the JeKi program have fewer conduct problems and get along better with their teacher. In both cases, improvements are very large between one third and one half of a standard deviation. Effects of this size seem plausible, given that despite of measuring intention-to-treat effects, most children living near JeKi schools actually follow the JeKi program. Nonetheless, the size of these coefficients should be interpreted with care, given that these estimates represent intention-to-treat effects and the actual participation rate among individuals in the sample is unknown.

The strong improvement of the relation with the teacher due to JeKi could actually reflect the parents' appreciation of the program rather than the attitude of the child itself. Given that the information is provided by parents, they might judge their child's relation with the teacher better, because they appreciate the existence of the JeKi program at school. At the same time, one could argue that parents who answer questions about their child's relation with the teacher primarily focus on the main teacher, given that she or he spends much more time with the children at school than the JeKi teacher, who visits from the local music school once or twice per week.

The effects of low music school fees on non-cognitive skills are not presented here. The main reason is that the treatment is not very well defined. Even if low music school fees lead children to sign up at a local music school, it is not obvious which treatment they receive. Music schools propose a diverse course offer, ranging from basic introductory classes presenting various musical instrument, to instrument lessons for individuals or small groups to ensemble lessons such as orchestra or harmony. On average, only 62% of the students attending public music schools receive traditional musical instrument lessons (see Table D.5). While any of these courses might affect non-cognitive skills, they do so through multiple channels. Results would therefore be difficult to interpret or to compare to the effects of the JeKi program, where the treatment is similar across participants. Nonetheless, the effects of low music school fees on non-cognitive skills were estimated and did not yield any interesting or significant results (Table D.12).

The effects of JeKi on socio-emotional skills are larger for boys and among children from richer households (columns 2 to 5). The former seem to considerably improve their appreciation of school and studying as a consequence of the JeKi program, both to a large extent of more than half of a standard deviation. The attitude toward school is also affected among children from households with an above-median income. The strong overall positive effect on conduct, by contrast, is mainly driven by girls and children from richer households. Potentially due to the small sample size, the pattern of heterogeneities is rather unstable when socio-economic background is defined using parental education or the child's migration background (see Table D.13). When differentiating by parental education, it seems that the improvements in conduct and relations with the teacher are rather driven by children with less highly educated parents.

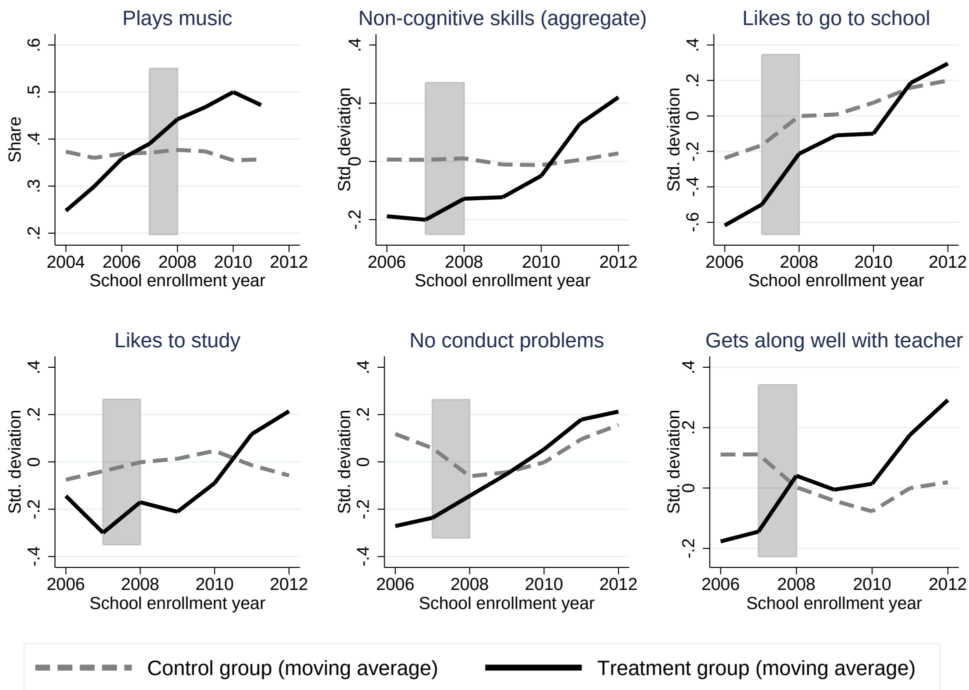
Anderson (2008) points out that studies examining multiple hypotheses typically find some positive effects simply due to false rejections of the null hypothesis. He stresses this as an issue that does not receive sufficient attention in economic research. Given that the sample size is rather small in the present study and that effect sizes seem to vary a lot, random rejections of the null hypothesis might explain some or all of the significant coefficients discussed above.

A simple way to reduce the number of hypotheses tested is to aggregate outcomes into a joint index (Anderson, 2008). The bottom part of Table 5.4 shows the effect of the JeKi program on two such aggregated outcome indices. The first is a weighted average of all ten non-cognitive outcomes presented in the upper part of the table. In the aggregation, each normalized outcome is weighted by the inverse of its covariance matrix with respect to all other non-cognitive outcomes. With other words, outcomes receive more weight in the aggregated index to the extent that they add new information to the index (Anderson, 2008). The second simply takes the average of the four difficulties from the Strengths and Difficulties Questionnaire: emotional symptoms, conduct problems, hyperactivity and peer problems. The coefficients of both aggregate outcomes are similar in size, but only statistically significant in the index, which summarizes all non-cognitive skills. The previous finding is confirmed, according to which effects are larger for boys and children from richer households.

5.6.3 Graphical evidence on the effects on music and socio-emotional skills

Figure 5.3 graphically illustrates some of the results for the JeKi treatment. The graph shows mean values of the outcomes for each cohort of students. The x-axis indicates the year of school enrollment. For each outcome, the black line represents the cohorts of children living in JeKi areas, both before and after the program was implemented. The dotted gray lines provide the mean outcome values for children in areas where JeKi never existed. Most of the JeKi programs were created in 2007 or 2008, as illustrated by the gray bar, but some were implemented as late as 2011. All averages are unconditional, without taking control variables into account.

Figure 5.3 – *Unconditional effects of the JeKi program*

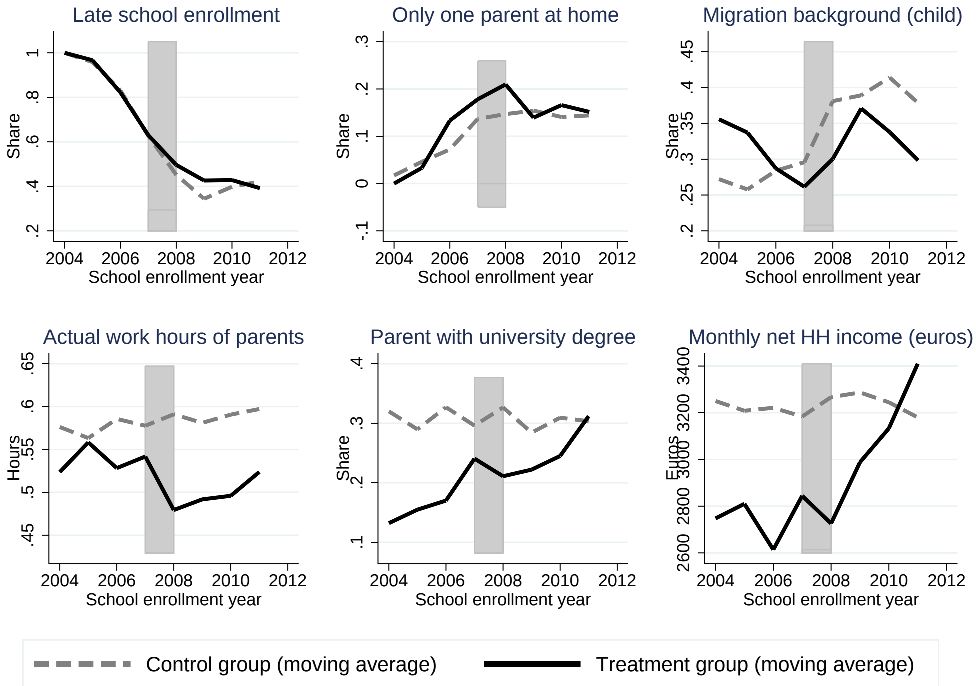


Source: SOEP v30 and FiD v4, own calculations. The black line shows yearly averages of outcome variables for students living in JeKi regions (both before and after the implementation of the program). The dotted gray line shows these averages for children in non-JeKi regions. All averages are moving averages of a 3-year window. The gray bar illustrates the years in which most of the JeKi projects were introduced.

The graphs confirm the regression results from Table 5.4. Before JeKi was created, music activities were lower in treatment than control regions. With the start of the first programs in 2007/2008, JeKi areas showed an increase in childhood music activities, while participation in other areas remained constant. The graphs also show that, with respect to socio-emotional skills, children in JeKi areas were generally doing less well before the program was implemented. Their appreciation of school and studying were approximately one fifth of a standard deviation lower, they had more conduct problems and got along less well with their teacher. When more and more children in these areas received the opportunity to take part in the JeKi program, these gaps gradually closed or even reversed.

Difference-in-differences estimators identify causal effects under the assumption of a common trend. With other words, in absence of treatment, outcomes in areas with and without JeKi are assumed to evolve similarly. Unfortunately, "playing music" is the only outcome variable for which we have several years of data before the first JeKi project was implemented. Figure 5.3 shows that the positive trend in music activities in treatment

Figure 5.4 – *Evolution of covariates in treatment and control group*



Source: SOEP v30 and FiD v4, own calculations. The graph shows yearly averages of some of the control variables used in this study, separately for students living in JeKi regions (black line, both before and after the implementation of the program) and for children in non-JeKi regions (dotted line). All averages are moving averages of a 3-year window. The gray bar illustrates the years in which most of the JeKi projects were introduced.

areas already existed prior to the implementation of JeKi. One explanation could be that some schools propose similar activities in advance, knowing that JeKi would be offered in the following year. The graph depicts 3-year moving averages, which could partly explain the pre-treatment positive trend as well.

To further investigate the plausibility of the common trend assumption, I plotted the same graph once again (Figure 5.4), this time describing the evolution of some control variables for the treatment and control group. These indicators of socio-economic status should not be affected by the existence of JeKi schools and thereby serve as a test for the common trend assumption.

The graph illustrates that some of the covariates indeed move in parallel between treatment and control group. These are – among many others not drawn in the graph – the year of school enrollment, whether the child lives in a single parent home, and whether she or he has a migration background. The lower part of Figure 5.4, however, shows that some covariates clearly do not move similarly in both groups. These are very relevant variables, which have been shown to affect participation in non-formal activities and skill development: the mother's education and work hours, as well as household income. These variables are taken into account in the main estimations of this paper, however, further unobserved characteristics might also not follow a parallel trend between treatment and control group. This is problematic, at least to the extent that these variables are uncorrelated with the control variables included in this study.

5.6.4 Comparing JeKi to varying local public music school fee levels

As described in Section 5.3, the effects of the JeKi program partly depend on activities that would have been carried out in absence of the program. Whether JeKi provides an advantageous opportunity to learn a musical instrument depends, among others, on the cost of taking music lessons outside the program. To explicitly model this trade-off, Table 5.5 shows the effects of JeKi separately in comparison to having access to music schools charging below or above average fees. A third comparison in the table examines the effects of below-average compared to above-average music school fees in regions that are not covered by the JeKi program.

The estimations presented so far give no hint as to whether the effects of the JeKi program can be attributed to learning to play a musical instrument or whether they result from other program characteristics, such as playing in a group or spending more time at school. If the former is true, we would expect the effects to be smaller when comparing JeKi to areas with low rather than high public music school fees. Children in non-JeKi areas with low music school fees are more likely to play music as well than those who have to pay a lot for musical instrument lessons. If JeKi has smaller effects in comparison to areas with low music school fees (where children are more likely to play music), the estimated effects are at least partly due to the activity of playing music. By contrast, if the effects are similar, no matter whether JeKi is compared to areas with high or low music school fees, these effects are likely to result from other program characteristics.

The results in the first row of Table 5.5 perfectly illustrate the cost differences for taking music lessons according to the level of music school fees. JeKi increases the probability to play a musical instrument by an insignificant 8 percentage points in areas with music school fees in the lowest third of the fee distribution. In areas, where music schools are more expensive and therefore the opportunity created by JeKi potentially matters more,

Table 5.5 – *Effect of a combined treatment (OLS estimates)*

	JeKi vs. Low music school fees (1)		JeKi vs. High music school fees (2)		Low music school fees vs. High music school fees (3)	
	Diff.	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*
Music activities						
Plays music	0.08	44	0.22	1	0.08	12
Sample size	347		779		868	
Non-cognitive skills						
Strengths and Difficulties Questionnaire						
No emotional symptoms	-0.16	59	-0.09	69	-0.19	15
No conduct problems	0.10	71	0.72	1	-0.20	10
Not hyperactive	0.03	93	0.19	38	0.07	44
No peer problems	-0.03	92	-0.05	83	-0.12	32
Pro-social behavior	-0.09	76	0.23	45	0.11	45
Attitude towards school						
Likes to go to school	0.13	55	0.34	27	0.17	21
Likes to study	0.45	13	0.08	76	0.07	60
Follows lessons well	0.18	48	-0.18	67	0.16	9
Gets along well...						
...with classmates	0.24	31	-0.02	93	0.03	79
...with teacher	0.33	5	0.26	11	-0.09	59
Combined non-cognitive skills scores						
Non-cog. skills (aggregate)	0.11	31	0.15	27	0.01	93
Difficulties (from SDQ)	-0.02	96	0.30	8	-0.15	19
Sample size	243		540		595	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression. Three columns present pairwise comparisons of three possible states of a treatment variable, which takes on the following values (the three groups are mutually exclusive): (1) Among the five closest primary schools, the share of students in a school having the JeKi program in the year of school enrollment is larger than 50% (same definition as in Table 5.4); (2) No JeKi program, but fees at the closest public music school are below the 33rd percentile; (3) No JeKi program and fees at the closest public music school are above the 33rd percentile. Please refer to Section 5.4 for more details on the treatment definition. All regressions control for individual and regional characteristics, as well as time and county fixed effects. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

musical activities increase by 22 percentage points thanks to the existence of the program. The comparison between non-JeKi regions with high and low music school fees reveals the expected pattern with music participation being 8 percentage points higher in regions with cheaper music schools. However, this difference is not statistically significant.

The lower part of Table 5.5 reveals that the positive effect of JeKi on conduct discussed above is driven by the comparison with areas where music schools charge high fees. Moreover, in this comparison, JeKi reduces children's general level of difficulties by almost a third of a standard deviation. By contrast, the improved relation with the teacher, which is caused by JeKi, is somewhat stronger in comparison to regions with low music school fees. An interpretation, which is in line with the hypotheses discussed in section 5.3, would be that playing music improves children's behavior, while small group activities positively affect relations with teachers. However, given the small sample size, we should be careful with such conclusions, as these positive and significant coefficients might appear by chance. Moreover, the aggregated non-cognitive skills score does not show significant coefficients in any of the three comparisons from Table 5.5.

5.6.5 Robustness

To examine the robustness of the results presented above, I performed various tests assessing the sensitivity of the main estimates with respect to measurement error, estimation procedures as well as the identifying assumptions. All tables showing these tests are presented in Appendix D.2.

As the construction of both treatment variables involved various steps (see Section 5.4), I extensively test whether the results change for alternative treatment definitions. For each treatment (JeKi and below-average music school fees), I test six alternative definitions and find that they all yield qualitatively similar results (Tables D.14 and D.15).

As described in Section 5.4, all outcome variables were measured at age 8 and 10. To increase the sample size, the main specification measures the outcomes at age 10 if available, and pools these with the additional observations for which outcomes are only available at age 8. Outcomes are only available at age 8 for one third of the sample. For another third, outcomes are only available at age 10. A third of the sample therefore reported all outcomes both at age 8 and 10. In a robustness check, I estimated the treatment effects separately for all individuals at age 8 and at age 10 (Table D.16). All effects are rather similar for both ages, which might seem surprising given that 8-year-old JeKi participants only attended the program for one or two years. An explanation could be that JeKi is mandatory in the first, but voluntary in the second year and thereafter. The share of individuals in the sample who actually take part in the program is higher at age 8 than age 10. Given that I estimate intention-to-treat effects, which average over the effects for participants and non-participants, longer participation (and thereby potentially larger effects) might cancel out with a lower number of participants.

The outcome variable describing music activities is composed of the answers to two questionnaire items, referring to music activities at or outside school. To test the robustness of the results with respect to the decision to pool these two questionnaire items, I separately examined the results for music activities at and outside school (Table D.17). Besides being less precisely estimated, the JeKi program has a much stronger effect on within-school activities when measured at age 8, while at age 10 some parents might consider the program to be an out-of-school activity. That seems plausible, given that in the first year, the program is carried out by the regular school teacher, while subsequent years are taught by teachers from a local music school.

In addition to examining the robustness of the results with respect to variable definitions, I re-estimated the effects using different estimation methods and samples. First, I re-estimated all effects using propensity score matching and entropy balancing rather than OLS (Table D.18, columns 1 and 2). The results are very similar. Second, I excluded all observations with at least one missing value on any of the covariates, which reduces the sample size by about 12% (Table D.18, column 3). This also does not change the results. In any case, the treatment has no relevant effects on whether a covariate is missing (Table D.17). Next, I estimated the effects without using a difference-in-differences estimator, which leaves them similar, and without including any covariates, which slightly increases them (Table D.19, columns 2 and 3). The results remain also similar when including panel weights, which are designed to make the data representative of the German population (Table D.19, column 1). As one (surprising) exception, including these survey weights eliminates the effect of JeKi on music activities.

Finally and most importantly, I assess the sensitivity of the results with respect to the identifying assumptions. The results of this paper have a causal interpretation if the following two statements are valid. First, in absence of the program, outcomes would have evolved similarly in regions (or schools) with and without regions (the common trend assumption). Second, families do not move to JeKi areas, in order to be able to benefit from the program.

In addition to the graphical evidence presented in Figure 5.3, I assess the plausibility of the common trend assumption by evaluating a placebo treatment. For the placebo treatment, I assume that each JeKi project started one year earlier than it actually did. Effects are estimated for the placebo start year, controlling for the actual JeKi start. If the placebo treatment shows positive effects, it is likely that JeKi schools already evolved differently even before the program was implemented. For the music school fee treatment, low music schools are drawn randomly. The results show that the common trend assumption might be violated with respect to the child's relations with the teacher, which also strongly improve thanks to the placebo treatment, even though the coefficients are not significant (Table D.20, column 1).

A second test of the common trend assumption examines how the JeKi program affects pre-treatment outcomes. According to Imbens and Wooldridge (2009), such tests using placebo outcomes are the more powerful, the more the placebo outcomes are closely related to the outcomes of interest. For this test, I benefit from the fact that some of the outcomes (music activities and the Strengths and Difficulties Questionnaire) were also measured when children were 6 years old. As JeKi only starts in the first grade of primary school, which children enter at age 6 or 7, these outcomes should not be affected by the program. Indeed, while some of the coefficients are large in magnitude, they are negative rather than positive, and (mostly) not statistically significant (Table D.20, column 2). Low music school fees also have an effect on music activities at age 6, because these almost do not vary over time and thus children were also treated at that age.

A third test is related to the assumption that families do not move to live closer to a JeKi school. To test this, I carried out three additional estimations (Table D.21). First, I assumed that children only attend schools within their municipality. Second and third, I restricted the estimation sample to those for whom I know that they did not move between entering the data/age 5 and when the outcomes are measured (age 8 or 10). As the decision about a new school entering the JeKi program was solely taken in the previous school year, I assume that parents could not have known about the program two years before their child entered primary school. The results are very similar to the main estimation results, which are therefore not driven by those who move. We should note, however, that this test has limited power, given that the subsample of non-movers is likely to be selective as well. It is therefore not obvious who is compared to whom in this test.

5.7 Conclusion

This paper examines if the universal music education program JeKi (*Jedem Kind ein Instrument*, an instrument for every child) affects its stated goal of providing access to musical instrument lessons to children from disadvantaged social backgrounds. I find that the program increases music participation among children from families with a household income below the median by 30 percentage points. Given that only 25% of these children played a musical instrument previously, this amounts to an increase of 115%. Despite being voluntary after the first year, the program thus successfully reaches those who did not attend musical instrument lessons before. The study by Busch and Kranefeld (2013) also finds that the choice to pursue the JeKi program beyond the mandatory first year depends neither on socio-economic status nor migration background.

However, JeKi does not affect music participation among children from households with an income above the median. This might be due to the fact that I use data measuring the incidence rather than the amount of music activities. Children from richer households are likely to learn a musical instrument, irrespective of whether they receive specific incentives at school. Given the data used, I am unable to identify whether participation in the JeKi program comes in addition to or instead of regular music school classes.

JeKi increases music participation more effectively than the alternative policy of low public music school fees. If the music school closest to a child's home charges fees within the lowest third of federal state's music school fee distribution, children from low-income households increase music participation as well. At 16 percentage points, this effect is lower than that of the JeKi program. Public music schools in Germany are obliged to charge lower fees from low-income households. The average fee level measured in the data might therefore not apply to the poorest households, which could partly explain the lower effect size.

Furthermore, this paper investigates whether JeKi affects the development of social and emotional skills. I find that the program improves conduct problems and relations with the teacher by about one half of a standard deviation. However, these effect sizes should be interpreted with care, given that they estimate intention-to-treat effects and the exact share of participating students is unknown. Effects are larger among boys and children from families with a household income above the median, who also like school and studying more thanks to the program.

Thorough robustness checks verify the sensitivity of these results with respect to multiple hypothesis testing, measurement error and the identifying assumptions. JeKi improves an aggregate index of socio-emotional skills by one sixth of a standard deviation. This confirms that the above-mentioned effects are not solely due to random and false rejections of the null hypothesis. All results are also robust to defining treatments, outcomes and socio-economic status differently, measuring outcomes at various ages, changing the estimation method, including survey weights, as well as restricting the sample to children with non-missing covariates.

The findings of this paper can be interpreted as causal, if areas with and without the JeKi program would have evolved similarly in absence of the JeKi program and if families did not move in order to live closer to a JeKi school. Given the large number of individual and regional control variables included in all estimations, as well as the broad availability of voluntary extracurricular music activities in all regions of Germany, these assumptions seem plausible. Moreover, outcomes are not affected by a placebo treatment happening one year earlier, and the JeKi program does not affect pre-treatment outcomes. In addition, results remain identical within the subsample of children who did not move between age 6 and 10.

To discuss the enormous expectations of policymakers, according to which universal music education is essential for children's personality development, we should ask whether the above-mentioned positive effects are a result of playing music. The JeKi program involves learning a musical instrument by interacting with teachers in small groups, and additionally joining an orchestra in grades 3 and 4. This study is unable to identify mechanisms that distinguish between these components. However, improvements in social skills and student-teacher relations are likely to result from the group-related component of the JeKi program. Nonte and Schwippert (2014) are unable to find positive effects of JeKi when comparing it to schools with an intensified sports curriculum. Hence, a program proposing sports or theater lessons in small groups might result in similar outcomes. Moreover, critics of the JeKi program state that it does not match regular music school courses with respect to quality, because classes are large and participation not entirely voluntary (Bossen, 2009). This could be another reason for a lack of music-specific effects like those found by Hille and Schupp (2015).

The primary goal of JeKi – to give children from less favorable socio-economic backgrounds access to music education – appears to be fulfilled. However, additional research is needed to find conclusive evidence on whether these programs affect cognitive and non-cognitive skills, as well as other educational outcomes. Further studies should use more detailed data and distinguish the mechanisms at work by comparing similar policies that differ in terms of content and implementation. A cost-benefit or cost-effectiveness analysis comparing different policies could determine to what extent such potential additional skill-related effects provide a further justification for public policy intervention.

6 Conclusion

The first part of this doctoral thesis investigated whether learning a musical instrument during adolescence affects the development of cognitive and non-cognitive skills. Chapter 2 came to the conclusion that even after taking into account numerous individual and parental background characteristics, adolescents with music training scored one sixth of a standard deviation above those not taking musical instrument lessons and were ten percentage points more likely to aim at attaining an upper secondary school or university degree. Moreover, these adolescents were about one fourth of a standard deviation more conscientious and open. These effects were larger among children from less favorable socio-economic backgrounds.

Do other non-formal educational activites yield the same benefits? Even when comparing musically active adolescents to those who play sports instead, chapter 3 found music training to benefit educational achievements, especially among children from more favorable social backgrounds. By contrast and as expected, playing sports improved (subjective) health.

Chapters 2 and 3 also confirmed that access to non-formal education, and in particular music training, strongly depends on socio-economic background. Given the benefits of non-formal education for skill development, unequal access might be a source of rising educational inequalities. Therefore, chapters 4 and 5 studied two German policies that aim to provide more equal access to music or sports activities.

Chapter 4 examined whether a subsidy for music school or sports club membership fees, addressed at poor households, increased participation in non-formal educational activities among eligible children. I found that the subsidy had no overall effect, but increased music and sports participation by 10 percentage points among eligible children from relatively richer households, as well as among children living with both of their parents. The subsidy thus only seemed to benefit welfare recipients, who are relatively more advantaged.

The subsidy might be ineffective on average, because it is too low to cover the costs of music or sports lessons. Another explanation might be that non-participation is a result of education and preferences rather than financial constraints. Therefore, chapter 5 studied another policy, which offers free or cheap musical instrument lessons in some primary schools in the German federal state of North Rhine-Westphalia: the *Jedem Kind ein Instrument* (an instrument for every child) program, known under the abbreviation JeKi.

I found that the JeKi program successfully increased music participation among children from less favorable socio-economic backgrounds. It did so more effectively than the alternative policy of reducing fees at the local public music school. In addition to increasing music participation, the JeKi program appeared to improve children's social skills, as well as their attitude towards school.

As all research projects, the results of these papers have to be interpreted in the light of some methodological weaknesses. Most importantly, all four papers claim to identify causal effects, but cannot entirely disregard doubts with respect to the identifying assumptions. Ideally, chapters 2 and 3 would have compared adolescents, who are identical along all observable and unobservable characteristics, except for their participation in extracurricular music activities. Unfortunately, I am not aware of any exogenous variation in music participation that could be used for identification within a representative and observational data set. Therefore, I chose selection-on-observables as a (second-best) strategy to take selection into account.

Despite the very rich set of covariates available in the German Socio-Economic Panel, I cannot exclude the possibility that unobserved heterogeneity drives the results presented in chapters 2 and 3. This is especially true, given that most of the covariates included in the estimations describe characteristics of the parents rather than of the adolescents themselves.

Instrumental variables might provide an alternative source of exogenous variation in music participation. Possible candidates that were considered as instrumental variables include the distance to or the average costs charged by the closest public music school. I managed to obtain these characteristics for all public music schools in Germany. However, variations in these measures were not sufficiently powerful to explain variations in music participation, as instrumental variables they would have been too weak. Even measures of preference for music rather than sports, as considered in chapter 3, were not powerful enough in explaining the choice between both activities. Moreover, we had difficulties convincing readers and seminar participants of the exogeneity of this measure.

Alternatively, the two policies studied in chapters 4 and 5 might have been used as instrumental variables that generate exogenous variation in music participation. Unfortunately, take-up of the Educational package is very low (see chapter 4). Therefore, the group of compliers would have been too small for the instrumental variable to be relevant.

Neither could the JeKi program (see chapter 5) be used as a measure of exogenous variation in music participation. North Rhine-Westphalia and Hamburg are the two only federal states in which an important share of the student population actually took part in the program. However, even though North Rhine-Westphalia is the largest German federal state, the sample that could be constructed using SOEP data was too small for instrumental variables estimations. Other data sets with a higher sample size do not provide geographical identifiers, which are a prerequisite to merge the data with information on JeKi schools. The federal state of Hamburg plans to share administrative school data with the research community. Although too late for my doctoral dissertation, new research opportunities with respect to the JeKi program will open up as soon as this announcement becomes reality.

Potential biases related to unobserved heterogeneity are somewhat less problematic in chapters 4 and 5. Still, other changes might have happened simultaneously with the introduction of the Educational package or the JeKi program, which I am not aware of and which affected the outcomes I studied. More importantly, the data used in both studies did not properly allow me to test the common trend assumption. Even in absence of both policy interventions, outcomes (music participation or non-cognitive skills) might have evolved more positively in the treatment than in the control group.

The other important weakness in all studies of this dissertation is related to the measurement of playing music and sports. Unlike most other studies on this topic in the fields of educational sciences, I use data that was not produced for the purpose of studying music participation. By contrast, other data sets that focus on music activities lack sufficient details on socio-economic background.

I measured music and sports participation using binary variables. In chapters 2 and 3, these were varied to some extent by restricting the group of musically active adolescents to those who took lessons outside of school or who have been playing for a certain number of years. Still, it would have been interesting to measure the extent to which children or adolescents actually engaged with music: the number of hours played per day or per week, the quality of the lessons taken, as well as a more detailed description of past musical activities, including those that have stopped before the questionnaire was answered.

Irrespective of identification and measurement issues, this dissertation almost entirely lacks a discussion of efficiency. Once the effects of music training – whether voluntary or encouraged through policy interventions – are identified, it is important to quantify them and to contrast them with the underlying costs. For example, it would have been interesting to compare the cost-efficiency between the Educational package, JeKi and other policies that stimulate skill development. However, such an analyses requires not only data on the costs associated with different policy interventions. It also assumes that benefits can be quantified or at least contrasted across policies.

To conclude, there is still much room for further research on the effects of music and other non-formal educational activities, as well as on potential policy interventions that encourage them. The discussion will hopefully continue and contribute to shaping new and innovative educational and social policies.

Bibliography

Almlund, Mathilde, Angela L. Duckworth, James J. Heckman, and Tim Kautz, "Personality Psychology and Economics," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 4 Elsevier Amsterdam 2011, pp. 1–181.

Anderson, Deborah, "'If You Let Me Play': The Effects of Participation in High School Athletics on Students' Educational and Labor Market Success," Technical Report, SSRN 2001.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Apel, Helmut and Dietrich Engels, "Bildung und Teilhabe von Kindern und Jugendlichen im unteren Einkommensbereich. Untersuchung der Implementationsphase des *Bildungs- und Teilhabepakets* im Auftrag des Bundesministeriums für Arbeit und Soziales," Federal Ministry of Labour and Social Affairs 2012.

Banks, James and Fabrizio Mazzonna, "The Effect of Education on Old Age Cognitive Abilities: Evidence from a Regression Discontinuity Design," *The Economic Journal*, 2012, *122* (560), 418–448.

Bargain, Olivier and Olivier Donni, "Targeting and child poverty," *Social Choice and Welfare*, 2011, *39* (4), 783–808.

Barnett, W. S., "Effectiveness of Early Educational Intervention," *Science*, 2011, *333* (6045), 975–978.

Barron, John M., Bradley T. Ewing, and Glen R. Waddell, "The Effects of High School Athletic Participation on Education and Labor Market Outcomes," *The Review of Economics and Statistics*, 2000, *82* (3), 409–421.

Bauernschuster, Stefan, Oliver Falck, and Ludger Woessmann, "Surfing alone? The internet and social capital: Evidence from an unforeseeable technological mistake," *Journal of Public Economics*, 2014, *117*, 73 – 89.

BBSR, "Indikatoren und Karten zur Raum- und Stadtentwicklung (INKAR)," Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR), www.inkar.de (accessed April 27th, 2015) 2015.

BDPM, "Über uns," Bundesverband deutscher Privatmusikschulen, http://www.bdpm-musikschulverband.de/aufgaben/ueber-uns/ (accessed April 22nd, 2015) 2015.

Becker, Gary S., “A Theory of the Allocation of Time,” *The Economic Journal*, 1965, *75* (299), pp. 493–517.

Bennett, Pamela R., Amy C. Lutz, and Lakshmi Jayaram, “Beyond the Schoolyard: The Role of Parenting Logics, Financial Resources, and Social Institutions in the Social Class Gap in Structured Activity Participation,” *Sociology of Education*, 2012, *85* (2), 131–157.

Bertrand, Marianne and Jessica Pan, “The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior,” *American Economic Journal: Applied Economics*, 2013, *5* (1), 32–64.

Besson, M., D. Schon, S. Moreno, A. Santos, and C. Magne, “Influence of musical expertise and musical training on pitch processing in music and language,” *Restorative Neurology and Neuroscience*, 2007, *25* (3-4), 399–410.

Bialystok, Ellen and Anne-Marie DePape, “Musical expertise, bilingualism, and executive functioning,” *Journal of Experimental Psychology, Human Perception and Performance*, 2009, *35* (2), 565–574.

Bilhartz, Terry D, Rick A Bruhn, and Judith E Olson, “The Effect of Early Music Training on Child Cognitive Development,” *Journal of Applied Developmental Psychology*, 1999, *20* (4), 615–636.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education,” *The Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

BMAS, “Evaluation der bundesweiten Inanspruchnahme und Umsetzung der Leistungen für Bildung und Teilhabe. Erster Zwischenbericht,” Federal Ministry of Labour and Social Affairs 2014.

_ , “The educational package: A new start for taking part,” Federal Ministry of Labour and Social Affairs 2015.

_ , “Evaluation der bundesweiten Inanspruchnahme und Umsetzung der Leistungen für Bildung und Teilhabe. Zweiter Zwischenbericht,” Federal Ministry of Labour and Social Affairs 2015.

BMBF, “Kultur macht Stark. BMBF fördert ausserschulische Bildungs- und Kulturangebote für benachteiligte Kinder und Jugendliche,” Press Statement 058/2012, Bundesministerium für Bildung und Forschung, http://www.buendnisse-fuer-bildung.de/de/kultur-macht-stark-175.php (accessed June 16th, 2014) 2012.

_ , “Die Ziele von Kultur macht stark: Bildungsarmut bekämpfen,” Bundesministerium für Bildung und Forschung, http://www.buendnisse-fuer-bildung.de/de/politischer-hintergrund-und-ziele.php (accessed June 16th, 2014) 2013.

Boca, Daniela Del, Chiara Monfardini, and Cheti Nicoletti, “Children’s and Parents’ Time-Use Choice and Cognitive Development during Adolescence,” Human Capital and Economic Opportunity Working Paper Series 2012-006, University of Chicago 2012.

Bossen, Anja, "JeKi: Die Zweifel sind übermächtig," *Neue Musikzeitung*, 2009, *58* (10).

Bourdieu, Pierre, "The Forms of Capital," in John G. Richardson, ed., *Handbook of Theory of Research for the Sociology of Education*, Greenwood Press 1986, pp. 241–258.

_ **and Jean-Claude Passeron**, *Reproduction in education, society and culture*, Vol. 4, SAGE Publications Limited, 1990.

Breuer, Christoph and Svenja Feiler, "Sportvereine in Deutschland – ein Überblick," in Christoph Breuer, ed., *Sportentwicklungsbericht 2013/2014. Analyse zur Situation der Sportvereine in Deutschland*, Köln: Sportverlag Strauß, 2015.

BSB, "Jedem Kind ein Instrument: Auftaktveranstaltung in der Fridjof-Nansen-Schule mit Senatorin Christa Goetsch," Pressemitteilung 18. Sep 2009, Behörde für Schule und Berufsbildung, http://bildungsserver.hamburg.de/contentblob/3873800/data/download-pdf-18-09-2009-pressemitteilung-zur-veranstaltung-mit-senatorin-goetsch-zum-jeki-start.pdf (accessed January 30th, 2015) 2009.

Busch, Thomas and Ulrike Kranefeld, "Wer nimmt an JeKi teil und warum? Programmteilnahme und musikalische Selbstkonzepte," in Koordinierungsstelle, ed., *Broschüre: JeKi-Forschungsschwerpunkt*, Bielefeld, 2013, pp. 46 – 49.

BVerfG, "Regelleistungen nach SGB II ('Hartz IV-Gesetz') nicht verfassungsgemäß," Pressemitteilung Nr. 5/2010 vom 9. Februar 2010, Bundesverfassungsgericht 2010.

Cabane, Charlotte, Adrian Hille, and Michael Lechner, "Mozart or Pelé? The effects of teenagers' participation in music and sports," Economics Working Paper Series 1509, University of St. Gallen, School of Economics and Political Science 2015.

_ **and Andrew Clark**, "Childhood Sporting Activities and Adult Labour-Market Outcomes," *Annals of Economics and Statistics*, 2015, *forthcoming.*

_ **and Michael Lechner**, "Physical activity of adults: A survey of correlates, determinants, and effects," *Jahrbücher für Nationalökonomie und Statistik*, 2015, *235* (4-5), 367–402.

Carlsson, Magnus, Gordon B. Dahl, Björn Öckert, and Dan-Olof Rooth, "The Effect of Schooling on Cognitive Skills," *Review of Economics and Statistics*, 2014, *97* (3), 533–547.

Carneiro, Pedro, Claire Crawford, and Alissa Goodman, "The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes," CEE Discussion Paper 92, Centre for the Economics of Education, LSE 2007.

Cascio, Elizabeth U. and Ethan G. Lewis, "Schooling and the Armed Forces Qualifying Test: Evidence from School-Entry Laws," *Journal of Human Resources*, 2006, *XLI* (2), 294–318.

Castillo, Marco, Paul J. Ferraro, Jeffrey L. Jordan, and Ragan Petrie, "The today and tomorrow of kids: Time preferences and educational outcomes of children," *Journal of Public Economics*, 2011, *95* (11-12), 1377–1385.

CDU, "Deutschlands Zukunft gestalten: Koalitionsvertrag zwischen CDU, CSU und SPD," https://www.cdu.de/sites/default/files/media/dokumente/koalitionsvertrag.pdf, accessed March 12th, 2015 2013.

Cobb-Clark, Deborah A and Stefanie Schurer, "The stability of big-five personality traits," *Economics Letters*, 2012, *115* (1), 11–15.

Covay, Elizabeth and William Carbonaro, "After the Bell: Participation in Extracurricular Activities, Classroom Behavior, and Academic Achievement," *Sociology of Education*, 2010, *83* (1), 20–45.

Craggs, Christopher, Kirsten Corder, Esther M.F. van Sluijs, and Simon J. Griffin, "Determinants of Change in Physical Activity in Children and Adolescents: A Systematic Review," *American Journal of Preventive Medicine*, 2011, *40* (6), 645 – 658.

Creech, Andrea, Patricia Gonzalez-Moreno, Lisa Lorenzino, and Grace Waitman, eds, *El Sistema and Sistema-inspired programmes: A literature review of research, evaluation, and critical debates*, Sistema Global, 2013.

Cunha, Flavio and James J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

_ and _ , "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 2008, *43* (4), 738–782.

_ , _ , and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

Currie, Janet, "Early Childhood Education Programs," *The Journal of Economic Perspectives*, 2001, *15* (2), 213–238.

_ and Firouz Gahvari, "Transfers in Cash and In-Kind: Theory Meets the Data," *Journal of Economic Literature*, 2008, *46* (2), 333–383.

Dahmann, Sarah and Silke Anger, "The Impact of Education on Personality: Evidence from a German High School Reform," IZA Discussion Papers 8139, Institute for the Study of Labor (IZA) 2014.

Damelang, Andreas and Georgi Kloß, "Poverty and the social participation of young people – an analysis of poverty-related withdrawal mechanisms," *Journal for Labour Market Research*, 2013, *46* (4), 321–333.

Deutsche Sportjugend, "Deutsche Sportjugend," http://www.dsj.de/deutsche-sportjugend/, accessed Dec 15th, 2015. 2014.

Deutscher Bundestag, "Entwurf eines Gesetzes zur Ermittlung von Regelbedarfen und zur Änderung des Zweiten und Zwölften Buches Sozialgesetzbuch," Drucksache 17/3404, Deutscher Bundestag 2010.

Deutscher Städtetag, ed., *Statistisches Jahrbuch deutscher Gemeinden* 2009.

_ , ed., *Statistisches Jahrbuch deutscher Gemeinden* 2013.

DiMaggio, Paul, "Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students," *American Sociological Review*, 1982, *47* (2), 189–201.

Donnellan, M. Brent and Richard E. Lucas, "Age differences in the big five across the life span: Evidence from two national samples," *Psychology and Aging*, 2008, *23* (3), 558–566.

Eccles, Jacquelynne S., Bonnie L. Barber, Margaret Stone, and James Hunt, "Extracurricular Activities and Adolescent Development," *Journal of Social Issues*, 2003, *59* (4), 865–889.

Eide, Eric R. and Nick Ronan, "Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond," *Economics of Education Review*, 2001, *20* (5), 431–442.

Emanuel, Ezekiel J., "Online education: MOOCs taken by educated few," *Nature*, 2013, *503* (7476), 342–342.

European Commission, "European Cultural Values, Special Eurobarometer 278," http://ec.europa.eu/public_opinion/archives/ebs/ebs_278_en.pdf, accessed December 15th, 2015. 2007.

_ , "European Cultural Values, Special Eurobarometer 412," http://ec.europa.eu/health/nutrition_physical_activity/docs/ebs_412_en.pdf, accessed December 15th, 2015. 2014.

Eurostat, "Youth in Europe. A statistical portrait," http://pjp-eu.coe.int/documents/1017981/1668203/YouthinEurope.pdf/40f42295-65e4-407b-8673-95e97026da4a, accessed December 15th, 2015 2009.

Ewing, Bradley T., "The Labor Market Effects of High School Athletic Participation: Evidence From Wage and Fringe Benefit Differentials," *Journal of Sports Economics*, 2007, *8* (3), 255–265.

Falck, Oliver, Constantin Mang, and Ludger Woessmann, "Virtually No Effect? Different Types of Computer Use and the Effect of Classroom Computers on Student Achievement," Working Paper, Ifo Institute at the University of Munich 2015.

Farb, Amy Feldman and Jennifer L. Matjasko, "Recent advances in research on school-based extracurricular activities and adolescent development," *Developmental Review*, 2012, *32* (1), 1 – 48.

Farrell, Lisa and Michael A. Shields, "Investigating the Economic and Demographic Determinants of Sporting Participation in England," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 2002, *165* (2), 335–348.

Felfe, Christina, Michael Lechner, and Andreas Steinmayr, "Sports and Child Development," IZA Discussion Paper 6105, Institute for the Study of Labor (IZA) 2011.

FMSB, "El Sistema," Official Site of the Simon Bolivar Music Foundation: Fundación Musical Simón Bolívar (FMSB), http://fundamusical.org.ve/ (accessed June 16th, 2014) 2013.

Frölich, Markus, "Nonparametric IV estimation of local average treatment effects with covariates," *Journal of Econometrics*, 2007, *139* (1), 35 – 75.

Garboua, Louis-Lévy and Claude Montmarquette, "A microeconometric study of theatre demand," *Journal of Cultural Economics*, 1996, *20*, 25–50.

Goebel, Jan and Bernd Pauer, "Datenschutzkonzept zur Nutzung von SOEPgeo im Forschungsdatenzentrum SOEP am DIW Berlin," *Zeitschrift für amtliche Statistik Berlin Brandenburg*, 2014, *3*, 42–47.

Goodman, Robert, "The Strengths and Difficulties Questionnaire: A Research Note," *Journal of Child Psychology and Psychiatry*, 1997, *38* (5), 581–586.

Grgic, Mariana and Ivo Züchner, *Medien, Kultur und Sport. Was Kinder und Jugendliche machen und ihnen wichtig ist. Die MediKuS-Studie.*, Beltz Juventa, 2013.

Gustafson, Sabrina L. and Ryan E. Rhodes, "Parental Correlates of Physical Activity in Children and Early Adolescents," *Sports Medicine*, 2006, *36* (1), 79–97.

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, 2012, *20* (1), 25–46.

_ and Yiqing Xu, "ebalance: A Stata Package for Entropy Balancing," *Journal of Statistical Software*, 2013, *54* (7).

Hansen, John D. and Justin Reich, "Democratizing education? Examining access and usage patterns in massive open online courses," *Science*, 2015, *350* (6265), 1245–1248.

Heckman, James and Rodrigo Pinto, "Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs," IZA Discussion Paper 7552, Institute for the Study of Labor (IZA) 2013.

Heckman, James J and Dimitriy V Masterov, "The productivity argument for investing in young children," *Applied Economic Perspectives and Policy*, 2007, *29* (3), 446–493.

Heckman, James J. and Tim Kautz, "Hard evidence on soft skills," *Labour Economics*, 2012, *19*, 451–464.

_ and _ , "Fostering and Measuring Skills: Interventions that Improve Character and Cognition," in James J. Heckman, James E. Humphries, and Tim Kautz, eds., *The Myth of Achievement Tests: The GED and the Role of Character in American Life*, Chicago, IL: University of Chicago Press, 2014, pp. 293–317.

_ , Jora Stixrud, and Sergio Urzua, "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 2013, *103* (6), 2052–86.

Heineck, Guido and Silke Anger, "The Returns to Cognitive Abilities and Personality Traits in Germany," *Labour Economics*, 2010, *17* (3), 535–46.

Hille, Adrian, *Etude d'impact des Orchestres à l'Ecole*, Institut Montaigne, 2010.

_ **and Jürgen Schupp**, "How learning a musical instrument affects the development of skills," *Economics of Education Review*, 2015, *44*, 56–82.

_ **, Annegret Arnold, and Jürgen Schupp**, "Leisure behavior of young people: Education-oriented activities becoming increasingly prevalent," *DIW Economic Bulletin*, 2014, *1*, 26–36.

_ **, Mila Staneva, and C. Katharina Spieß**, "More and more students, especially those from middle-income households, are using private tutoring," *DIW Economic Bulletin*, 2016, *6*, 63–71.

Ho, Yim-Chi, Mei-Chun Cheung, and Agnes S. Chan, "Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children," *Neuropsychology*, 2003, *17* (3), 439–450.

Horst, Klazine Van Der, Marijke J Chin A Paw, Jos W R Twisk, and Willem Van Mechelen, "A brief review on correlates of physical activity and sedentariness in youth," *Medicine and science in sports and exercise*, 2007, *39* (8), 1241–1250.

Huber, Martin, Michael Lechner, and Andreas Steinmayr, "Radius matching on the propensity score with bias adjustment: Finite sample behaviour, tuning parameters and software implementation," *Empirical Economics*, 2015, *49* (1), 1–31.

_ **,** _ **, and Conny Wunsch**, "The performance of estimators based on the propensity score," *Journal of Econometrics*, 2013, *175* (1), 1–21.

Ichino, Andrea, Fabrizia Mealli, and Tommaso Nannicini, "From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity?," *Journal of Applied Econometrics*, 2008, *23* (3), 305–327.

Imai, Kosuke, Luke Keele, and Dustin Tingley, "A general approach to causal mediation analysis," *Psychological Methods*, 2010, *15* (4), 309–334.

Imbens, Guido W., "The role of the propensity score in estimating dose-response functions," *Biometrika*, 2000, *87* (3), 706–710.

_ , "Matching Methods in Practice: Three Examples," *Journal of Human Resources*, 2015, *50* (2), 373–419.

_ **and Jeffrey M. Wooldridge**, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 2009, *47* (1), 5–86.

IT.NRW, "Verzeichnis der allgemeinbildenden Schulen in Nordrhein-Westfalen 2011," Information und Technik Nordrhein-Westfalen, Geschäftsbereich Statistik. 2011.

JeKi-Stiftung, "'Jedem Kind ein Instrument' startet mit Erfolgsquote ins zweite Schuljahr," Pressemitteilung 28. Mai 2008, JeKi-Stiftung, https://www.jedemkind.de/programm/presse/eingebundene_dokumente/080528_pi_zwischenbilanz.pdf (accessed January 27th, 2015) 2008.

_ , "Jedem Kind ein Instrument: Programmstandards," http://www.jekits.de/app/uploads/2014/10/120326_Programmstandards_2011_2012.pdf (accessed April 28th, 2015) 2011.

JeKi-Stiftung, "Jedem Kind ein Instrument: Zahlen und Fakten," http://www.mfkjks.nrw.de/web/ media_get.php?mediaid=29707 &fileid=98596 &sprachid=1 (accessed April 28th, 2015) 2013.

JeKi-Stiftung, "Jedem Kind ein Instrument: Programmbeschreibung," http://www.mfkjks.nrw.de/web/media_get.php?mediaid=29706&fileid=98595&sprachid=1 (accessed April 28th, 2015) 2014.

_ , "Informationen für Musikschulen," https://www.jedemkind.de/programm/musikschulen/informationen.php (accessed January 27th, 2015) 2015.

Jürges, Hendrik, Kerstin Schneider, and Felix Büchel, "The effect of central exit examinations on student achievement: quasi-experimental evidence from TIMSS Germany," *Journal of the European Economic Association*, 2005, *3* (5), 1134–1155.

Kulturstiftung, "Jedem Kind ein Instrument: Programm zur Förderung der kulturellen Bildung – der Beitrag der Kulturstiftung des Bundes zur europäischen Kulturhauptstadt 2010," Kulturstiftung des Bundes, http://www.kulturstiftung-des-bundes.de/cms/de/programme/kunst_der_vermittlung/archiv/ jedem_kind_ein_instrument_2929_162.html (accessed July 24th, 2015) 2015.

Lang, Frieder R., Dennis John, Oliver Lüdtke, Jürgen Schupp, and Gert G. Wagner, "Short assessment of the Big Five: Robust across survey methods except telephone interviewing," *Behavior Research Methods*, 2011, *43* (2), 548–567.

Lareau, Annette, *Unequal Childhoods: Class, Race, and Family Life*, 2nd ed., University of California Press, 2011.

_ **and Elliot B. Weininger**, "Cultural Capital in Educational Research: A Critical Assessment," *Theory and Society*, 2003, *32* (5/6), 567–606.

Lechner, Michael, "Identification and estimation of causal effects of multiple treatments under the conditional independence assumption," in Michael Lechner and Friedhelm Pfeiffer, eds., *Econometric Evaluation of Labour Market Policies*, Vol. 13 of ZEW Economic Studies, Physica-Verlag HD, 2001, pp. 43–58.

_ , "Long-run labour market and health effects of individual sports activities," *Journal of Health Economics*, 2009, *28* (4), 839 – 854.

_ **and Conny Wunsch**, "Are Training Programs More Effective When Unemployment Is High?," *Journal of Labor Economics*, 2009, *27* (4), 653–692.

_ **and Nazmi Sari**, "Labor market effects of sports and exercise: Evidence from Canadian panel data," *Labour Economics*, 2015, *35*, 1 – 15.

_ **and Paul Downward**, "Heterogeneous sports participation and labour market outcomes in England," Economics Working Paper Series 1325, University of St. Gallen, School of Economics and Political Science 2013.

Light, Richard L., “Children’s Social and Personal Development Through Sport: A Case Study of an Australian Swimming Club,” *Journal of Sport & Social Issues*, 2010, *34* (4), 379–395.

Lindenberg, Siegwart, “Social Production Functions, Deficits, and Social Revolutions: Prerevolutionary France and Russia,” *Rationality and Society*, 1989, *1* (1), 51–77.

Lipscomb, Stephen, “Secondary school extracurricular involvement and academic achievement: a fixed effects approach,” *Economics of Education Review*, 2007, *26* (4), 463 – 472.

Long, James E and Steven B Caudill, “The Impact of Participation in Intercollegiate Athletics on Income and Graduation,” *The Review of Economics and Statistics*, 1991, *73* (3), 525–31.

Loui, Psyche, Kenneth Kroog, Jennifer Zuk, Ellen Winner, and Gottfried Schlaug, “Relating Pitch Awareness to Phonemic Awareness in Children: Implications for Tone-Deafness and Dyslexia,” *Frontiers in Psychology*, 2011, *2* (111), 1–5.

Lunn, Pete and Elish Kelly, “Accounting for Taste: An Examination of Socioeconomic Gradients in Attendance at Arts Events,” ESRI Working Paper 283, Economic and Social Research Institute (ESRI), Dublin/Ireland 2009.

Makles, Anna and Kerstin Schneider, “Much Ado about Nothing? The Role of Primary School Catchment Areas For Ethnic School Segregation: Evidence From a Policy Reform,” *German Economic Review*, 2015, *16* (2), 203–225.

Marcus, Jan, Janina Nemitz, and C. Katharina Spieß, “Veränderungen in der gruppenspezifischen Nutzung von ganztägigen Schulangeboten – Längsschnittanalysen für den Primarbereich,” *Zeitschrift für Erziehungswissenschaft*, 2015, *forthcoming.*

McCrae, Robert R. and Paul T. Costa, “The Five Factor Theory of Personality,” in Oliver P. John, Richard W. Robins, and Lawrence A. Pervin, eds., *Handbook of personality: Theory and research 2*, Guilford, New York 1999, pp. 139–153.

Menninghaus, Winfried, *Wozu Kunst? Ästhetik nach Darwin*, Suhrkamp Berlin, 2011.

MIZ, “Kooperationen von Musikschulen im VdM mit allgemein bildenden Schulen und anderen Partnern,” Deutsches Musikinformationszentrum (MIZ), http://www.miz.org/intern/uploads/statistik127.pdf (accessed April 22nd, 2015) 2014.

_ , “Einnahmen und Ausgaben der Musikschulen im VdM,” Deutsches Musikinformationszentrum (MIZ), http://www.miz.org/intern/uploads/statistik125.pdf, accessed January 7th 15th, 2015. 2015.

_ , “Schülerinnen und Schüler in veschiedenen Fächern der VdM-Musikschulen,” Deutsches Musikinformationszentrum (MIZ), http://www.miz.org/downloads/statistik/3/statistik3.pdf, accessed January 7th, 2015. 2015.

Nering, M. E., “The effect of piano and music instruction on intelligence of monozygotic twins,” *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 2002, *63* (3A), 812.

Neville, H., "Effects of music training on brain and cognitive development in underprivileged 3- to 5-year-old children: Preliminary results," in B. Rich and C. Asbury, eds., *Learning, Arts, and the Brain: The Dana Consortium Report on Arts and Cognition*, The Dana Foundation, New York/Washington, DC 2008, pp. 105–106.

Nonte, Sonja and Knut Schwippert, "Transfereffekte von Instrumentalunterricht. Der Einfluss des Instrumentalunterrichts auf die soziale und motivationale Entwicklung von Schülerinnen und Schülern im schulischen Kontext," in Andreas Lehmann-Wermser, Veronika Busch, Knut Schwippert, and Sonja Nonte, eds., *Mit Mikrofon und Fragebogen in die Grundschule*, Waxmann, Münster, 2014, pp. 31–62.

Ones, Deniz S., Stephan Dilchert, Chockalingam Viswesvaran, and Timothy A. Judge, "In Support of Personality Assessment in Organizational Settings," *Personnel Psychology*, 2007, *60* (4), 995–1027.

Ormel, Johan, Siegwart Lindenberg, Nardi Steverink, and Lois M. Verbrugge, "Subjective Well-Being and Social Production Functions," *Social Indicators Research*, 1999, *46*, 61–90.

Pervin, Lawrence A., Daniel Cervone, and Oliver P. John, *Personality: Theory and Research*, 9th ed., John Wiley & Sons, Inc., 2005.

Pfeifer, Christian and Thomas Cornelissen, "The impact of participation in sports on educational attainment: New evidence from Germany," *Economics of Education Review*, 2010, *29* (1), 94–103.

Racine, Jeffrey S. and James G. MacKinnon, "Inference via kernel smoothing of bootstrap values," *Computational Statistics & Data Analysis*, 2007, *51* (12), 5949–5957.

Raudsepp, Lennart, "The relationship between socio-economic status, parental support and adolescent physical activity," *Acta Pædiatrica*, 2006, *95* (1), 93–98.

Rees, Daniel I. and Joseph J. Sabia, "Sports participation and academic performance: Evidence from the National Longitudinal Study of Adolescent Health," *Economics of Education Review*, 2010, *29* (5), 751 – 759.

Richter, David, Maria Metzing, Michael Weinhardt, and Jürgen Schupp, "SOEP Scales Manual," SOEP Survey Papers 138 2013.

Riedel, Andrea, Kerstin Schneider, Claudia Schuchart, and Horst Weishaupt, "School Choice in German Primary Schools: How Binding are School Districts?," *Journal for Educational Research Online / Journal für Bildungsforschung Online*, 2010, *2* (1), 94–120.

Ritchie, Laura and Aaron Williamon, "Measuring distinct types of musical self-efficacy," *Psychology of Music*, 2011, *39* (3), 328–344.

Roberts, Brent, Richard Robins, Kali Trzesniewski, and Avshalom Caspi, "Personality Trait Development in Adulthood," in Jeylan Mortimer and Michael Shanahan, eds., *Handbook of the Life Course*, Handbooks of Sociology and Social Research, Springer US, 2003, pp. 579–595.

Roden, Ingo, Dietmar Grube, Stephan Bongard, and Gunter Kreutz, "Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study," *Psychology of Music*, 2014, *42* (2), 284–298.

_ **, Gunter Kreutz, and Stephan Bongard**, "Effects of a school-based instrumental music program on verbal and visual memory in primary school children: a longitudinal study," *Frontiers in Psychology*, 2012, *3* (572).

Sallis, J. F., J. J. Prochaska, and W. C. Taylor, "A review of correlates of physical activity of children and adolescents," *Medicine & Science in Sports & Exercise*, 2000, *32* (5), 963–75.

Schellenberg, E. Glenn, "Music Lessons Enhance IQ," *Psychological Science*, 2004, *15* (8), 511–514.

_ , "Long-Term Positive Associations Between Music Lessons and IQ," *Journal of Educational Psychology*, 2006, *98* (2), 457–468.

_ , "Examining the association between music lessons and intelligence," *British Journal of Psychology*, 2011, *102* (3), 283–302.

_ , "Music training and nonmusical abilities," in S. Hallam, I. Cross, and M. Thaut, eds., *The Oxford Handbook of Music Psychology*, Oxford University Press, 2016, pp. 415–429.

_ **and Sylvain Moreno**, "Music lessons, pitch processing, and g," *Psychology of Music*, 2010, *38* (2), 209–221.

Schober, Pia and C. Katharina Spieß, "Frühe Förderung und Betreuung von Kindern: Bedeutende Unterschiede bei der Inanspruchnahme besonders in den ersten Lebensjahren," *DIW Wochenbericht*, 2012, *79* (43), 17–28.

Schröder, Mathis, Rainer Siegers, and C. Katharina Spieß, "'Familien in Deutschland' – FiD," *Schmollers Jahrbuch*, 2013, *133* (4), 595–606.

Schroeder, Carsten, C. Katharina Spieß, and Johanna Storck, "Private Spending on Children's Education: Low-Income Families Pay Relatively More," *DIW Economic Bulletin*, 2015, *5* (8), 113–123.

Schumacher, Ralph, "Pauken mit Trompeten. Lassen sich Lernstrategien, Lernmotivation und soziale Kompetenzen durch Musikunterricht fördern?," Bildungsforschung Band 32, Bundesministerium für Bildung und Forschung (BMBF) 2009.

Schupp, Jürgen and Sabrina Herrmann, "Kognitionspotenziale Jugendlicher. Ergänzungen zum Jugendfragebogen der Längsschnittstudie Sozio-oekonomisches Panel (SOEP)," DIW-Data Documentation No. 43, DIW Berlin 2009.

Seaman, Bruce A., "Chapter 14: Empirical Studies of Demand for the Performing Arts," in Victor A. Ginsburg and David Throsby, eds., *Handbook of the Economics of Art and Culture*, Vol. 1, Elsevier, 2006, pp. 415 – 472.

Segal, Carmit, "Classroom Behavior," *Journal of Human Resources*, 2008, *43* (4), 783–814.

_ , "Misbehavior, Education, and Labor Market Outcomes," *Journal of the European Economic Association*, 2013, *11* (4), 743 – 779.

Shahin, Antoine J., Larry E. Roberts, Wilkin Chau, Laurel J. Trainor, and Lee M. Miller, "Music training leads to the development of timbre-specific gamma band activity," *NeuroImage*, 2008, *41* (1), 113 – 122.

Singh, A, L Uijtdewilligen, JR Twisk, W van Mechelen, and Chinapaw MM, "Physical activity and performance at school: A systematic review of the literature including a methodological quality assessment," *Archives of Pediatrics & Adolescent Medicine*, 2012, *166* (1), 49–55.

Southgate, Darby E. and Vincent J. Roscigno, "The Impact of Music on Childhood and Adolescent Achievement," *Social Science Quarterly*, 2009, *90* (1), 4–21.

Specht, Jule, Boris Egloff, and Stefan C. Schmukle, "Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five," *Journal of Personality and Social Psychology*, 2011, *101* (4), 862–882.

_ , _ , and _ , "Everything Under Control? The Effects of Age, Gender, and Education on Trajectories of Perceived Control in a Nationally Representative German Sample," *Developmental Psychology*, 2013, *49* (2), 353–364.

Spychiger, Maria, Lucia Gruber, and Franziska Olbertz, "Musical Self-Concept: Presentation of a Multi-Dimensional Model and Its Empirical Analyses," in Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, and Päivi-Sisko Eerola, eds., *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, 2009.

Statistisches Bundesamt, "Schulen und Klassen nach Schularten: Allgemeinbildende und berufliche Schulen," Statistisches Bundesamt, https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Schulen/Tabellen/AllgemeinBildendeBeruflicheSchulenSchularten.html (accessed July 24th, 2015) 2015.

Stevenson, Betsey, "Beyond the Classroom: Using Title IX to Measure the Return to High School Sports," *Review of Economics and Statistics*, 2010, *92* (2), 284–301.

Stuart, Elizabeth A., "Matching Methods for Causal Inference: A Review and a Look Forward," *Statistical Science*, 2010, *25* (1), 1–21.

Switlick, Bettina and Claudia Bullerjahn, "Ursachen und Konsequenzen des Abbruchs von Instrumentalunterricht. Eine quantitative und qualitative Umfrage bei Studierenden der Universität Hildesheim.," in Niels Knolle, ed., *Musikpädagogik vor neuen Forschungsaufgaben*, Essen 1999, pp. 167–195.

Todd, Petra E. and Kenneth I. Wolpin, "On the Specification and Estimation of the Production Function for Cognitive Achievement," *The Economic Journal*, 2003, *113* (485), F3–F33.

_ **and** _ , "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps," *Journal of Human Capital*, 2007, *1* (1), 91–136.

Vaughn, Kathryn and Ellen Winner, "SAT Scores of Students Who Study the Arts: What We Can and Cannot Conclude about the Association," *Journal of Aesthetic Education*, 2000, *34* (3/4), 77–89.

VDM, "Richtlinien für die Mitgliedschaft im Verband deutscher Musikschulen e.V. (VdM)," Verband deutscher Musikschulen, http://www.musikschulen.de/medien/doks/vdm/richtlinien-des-vdm-2011_logo.pdf (accessed June 16th, 2014) 2011.

_ , "Mitgliedschulen im VdM," Verband deutscher Musikschulen, http://www.musikschulen.de/ musikschulen/fakten/vdm-musikschulen/index.html (accessed April 2nd, 2015) 2014.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp, "The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements," *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 2007, *127* (1), 139–169.

Weinhardt, Michael and Jürgen Schupp, "Multi-Itemskalen im SOEP Jugendfragebogen," Data Documentation 60, DIW Berlin, German Institute for Economic Research 2011.

Weininger, Elliot B., Annette Lareau, and Dalton Conley, "What Money Doesn't Buy: Class Resources and Children's Participation in Organized Extracurricular Activities," *Social Forces*, 2015.

Winner, Ellen, Thalia Goldstein, and Stéphan Vincent-Lancrin, *Art for Art's Sake?*, OECD Publishing, 2013.

Yaish, Meir and Tally Katz-Gerro, "Disentangling 'Cultural Capital': The Consequences of Cultural and Economic Resources for Taste and Participation," *European Sociological Review*, 2012, *28* (2), 169–185.

Appendix to "How learning a musical instrument affects the development of skills"

A.1 Data

This study uses data from the 29th distribution of the German Socio-Economic Panel (SOEP), which covers all SOEP survey years from 1984 to 2012 (Wagner et al., 2007). The SOEP is a longitudinal household study for Germany, which surveys a random sample of more than 12,000 households in Germany on an annual basis. Since 2000, SOEP household members are asked to answer the SOEP Youth Questionnaire in the year they turn 17, which contains numerous questions on the respondent's childhood and youth. The sample examined in this study contains all adolescents who have answered the SOEP Youth Questionnaire between 2001 and 2012 (in the first year, questions on musical activities during childhood and youth had not been asked). From 2001 to 2012, the cohort size decreases annually from approximately 350 to 250 observations per year.

Table A.1 describes how we obtain our final samples from the original data. In order to allow comparisons between the different outcomes used in this study, we create three samples, each of which will be used to study one group of outcome variables. The largest sample consists of 3,488 individuals and includes valid observations for all outcome variables which were available for each survey year. These include school grades, perceived control, as well as all variables of the categories time use and ambition. For the Big Five personality traits, as well as for cognitive skills, the sample sizes are considerably smaller. This is due to the fact that questions assessing the Big Five personality traits were only introduced in the SOEP Youth Questionnaire in 2006. The questions examining adolescents' cognitive skills were answered by survey respondents from the year 2004 onwards. However, as cognitive skills are assessed in interactive tests, these are only available for

Table A.1 – *Sample construction*

Sample description	**Remaining observations**
All respondents to the SOEP Youth Questionnaire 2001-2012	3,958
Drop individuals with missing answers for questions 16: Do you play a musical instrument? 19: How old were you when you started to play music? 20: Music lessons outside of school	3,942
Drop individuals with missing information on federal state	3,941
Sample for estimating the propensity score	**3,941**
I. Outcomes available for every survey year	
Drop individuals with missing school grades	3,764
Drop individuals with missing perceived control	3,623
Drop individuals with missing estimation of future job success	3,572
Drop individuals with missing educational aspirations	3,508
Drop individuals with missing time use information	3,488
Sample for outcomes available in every survey year	**3,488**
II. Cognitive skills: measured since 2004, only questionnaires answered in the presence of interviewer	
Drop individuals surveyed between 2001 and 2003	2,609
Drop individuals who did not answer the cognitive skills test	1,878
Drop individuals who answered the cognitive skills test only partly	1,847
Sample for cognitive skills	**1,847**
III. Big 5 Personality traits: Only measured since 2006	
Drop individuals surveyed between 2001 to 2005	1,872
Drop individuals with missing for at least one personality dimension	1,815
Sample for Big 5 personality traits	**1,815**

survey respondents who answered the SOEP Youth Questionnaire in the presence of an interviewer.

Table A.2 provides a detailed description of all outcomes considered in this study, how and in which years they were measured. Moreover, the table shows the units in which the original variables were assessed, and how they were transformed for this study. All outcome variables were taken from answers to the SOEP Youth Questionnaire, which is provided in the BIOAGE17 file of the SOEP data.

Table A.3 provides further details on the way cognitive skills and personality were assessed. Cognitive skills were measured with three small tests, in which survey respondents have to interactively answer questions assessing their level of intelligence. Personality was measured with a self-assessment, in which adolescents had to provide their degree of ap-

Table A.2 – *List of outcome variables*

Variable	**How assessed**	**Available when**	**Measurement and units**	
			Original	*In this study*
Cognitive skills				
Figures	See Table A.3	2004-2012	0-20 points	Normalized
Analogies	See Table A.3	2004-2012		(within control group)
Maths operators	See Table A.3	2004-2012		Mean: 0
Average cog. skills	Mean of the above	2004-2012		SD: 1
School grades				
German	Self-reported	2001-2012	1 (very good)	Normalized
Mathematics	Self-reported	2001-2012	to 6 (fail)	(within control group
First foreign language	Self-reported	2001-2012		and school track)
Average grade	Mean of the above	2001-2012		Mean: 0 | SD: 1
Personality				
Conscientiousness	See Table A.3	2006-2012	7-point	Normalized
Openness	See Table A.3	2006-2012	Likert-scale	(within control group)
Agreeableness	See Table A.3	2006-2012		Mean: 0
Extraversion	See Table A.3	2006-2012		SD: 1
Neuroticism	See Table A.3	2006-2012		
Perceived control	See Table A.3	2001-2012		
Time use				
Watch TV daily	Direct question	2001-2012	0 (no)/1 (yes)	0 (no)/1 (yes)
Read books daily	Direct question	2001-2012	0 (no)/1 (yes)	0 (no)/1 (yes)
Ambition				
Aim upper secondary school degree	Direct question	2001-2012	0 (no)/1 (yes)	0 (no)/1 (yes)
Aim university	Direct question	2001-2012	0 (no)/1 (yes)	0 (no)/1 (yes)
Desired profession	How likely to find a job in your field?	2001-2012	0% to 100%	0% to 100%
Job success likely	How likely to be successful and get ahead?	2001-2012	0% to 100%	0% to 100%

proval on a variety of items. For each personality dimension, the final score is simply the mean of the respective items.

Table A.4 describes all control variables used in the estimations of this study. Moreover, the table indicates each variable's source file in the SOEP data, as well as the year in which the variable was measured. Missing values were coded to 0 for binary and to the mean for continuous variables. Additional indicators are included in all estimations, which are coded to one if at least one variable in a group of outcomes is missing. Such indicators were created for the following variable groups:

- *Parents' personality:* "1" if one of the following variables is missing: conscientiousness, extraversion, neuroticism, agreeableness, openness, or appreciation of the arts

Table A.3 – *Items measuring cognitive skills and personality*

Variable	**Items**
Cognitive skills	
Figures	Identify the figure needed to complete a row
Analogies	Identify word pairs, e.g. meadow-grass vs. forest-? [trees]
Maths operators	Insert mathematics operators into small calculus problems
Personality	I see myself as someone who...
Conscientiousness	...does a thorough job
	...does things effectively and efficiently
	...tends to be lazy *(reversed)*
Openness	...is original, comes up with new ideas
	...values artistic experiences
	...has an active imagination
	...is eager for knowledge
Agreeableness	...is sometimes somewhat rude to others *(reversed)*
	...has a forgiving nature
	...is considerate and kind to others
Extraversion	...is communicative, talkative
	...is outgoing, sociable
	...is reserved *(reversed)*
Neuroticism	...worries a lot
	...gets nervous easily
	...is relaxed, handles stress well *(reversed)*
Perceived control	How my life goes, depends on myself
	Compared to others, I have not achieved what I deserved *(reversed)*
	What one achieves is mainly a question of luck and fate *(reversed)*
	I often have the experience that others make decisions regarding my life *(reversed)*
	When I encounter difficulties I have doubts about my abilities *(reversed)*
	Opportunities in life are determined by social conditions *(reversed)*
	I have little control over the things that happen in my life *(reversed)*

- *Household income:* “1” if household income is missing
- *Cultural consumption:* “1” if one of the following variables is missing: attending cultural events and being artistically active
- *Number of books at home:* “1” if the number of books at home is missing
- *Secondary school recommendation:* “1” if the secondary school recommendation is missing
- *Other covariates:* “1” if any of the other covariates are missing

In addition to these missing indicators as well as the variables described in Table A.4, all estimations include an indicator stating whether the individual entered the SOEP sample

Table A.4 – *List of control variables*

Variable	SOEP file	Measured at age...
Motivation for music enrollment[1]		
Mother has no degree	PGEN (generated individ. variables)	age 5 or entry[2]
Mother has completed only 9 years of school	PGEN	age 5 or entry[2]
Mother has vocational degree	PGEN	age 5 or entry[2]
Mother has university degree	PGEN	age 5 or entry[2]
Mother has migration background	PPFAD (individual core information)	time constant
Mother's age at birth	PPFAD	time constant
Parents care about school achievement	BIOAGE17 (youth questionnaire)	retrospectively at 17
Parents do not support learning	BIOAGE17	retrospectively at 17
Conflict with parents due to school results	BIOAGE17	retrospectively at 17
Parents attend parent-teacher meetings	BIOAGE17	retrospectively at 17
Parents attend teacher's consultation hours	BIOAGE17	retrospectively at 17
Parents actively contact school teachers	BIOAGE17	retrospectively at 17
Parents engage as parent representatives	BIOAGE17	retrospectively at 17
Parents do not engage with child's school	BIOAGE17	retrospectively at 17
Conscientiousness (mother)[3]	P (individual questionnaire)	2005 or 2009[4]
Extraversion (mother)[3]	P	2005 or 2009[4]
Neuroticism (mother)[3]	P	2005 or 2009[4]
Agreeableness (mother)[3]	P	2005 or 2009[4]
Openness (mother)[3]	P	2005 or 2009[4]
Mother attends cultural events every month	P	age 5 or entry[2]
Mother never attends cultural events	P	age 5 or entry[2]
Mother is artistically active every month	P	age 5 or entry[2]
Mother is never artistically active	P	age 5 or entry[2]
Mother's appreciation for the arts	P	2005 or 2009[4]
More than 200 books at home	H (household questionnaire)	2001, 2006 or 2011[4]
Fewer than 50 books at home	H	2001, 2006 or 2011[4]
Constraints for music enrolment		
Monthly household log net income (simple, squared)	HGEN (generated household variables)	age 5 or entry[2]
Rooms per person at home	HGEN	age 5 or entry[2]
Number of siblings in the SOEP	BIOSIB	time constant
Child is the firstborn	BIOSIB	time constant
Federal state (15 dummies)	HBRUTTO (fieldwork information)	age 5 or entry[2]
Household lives in a rural area	HBRUTTO	age 5 or entry[2]
Adolescent's motivation to continue playing music		
Recommendation for upper secondary school	BIOAGE17	retrospectively at 17
Recommendation for lower secondary school	BIOAGE17	retrospectively at 17
Adolescent attends upper secondary school	BIOAGE17	retrospectively at 17
Gender	PPFAD	time constant
Birth year (10 dummies)	PPFAD	time constant

[1] Mother's value if available, otherwise the father's value is used.
[2] If the household was not in the SOEP when the child was aged 5, these variables were measured in the year the household entered the SOEP. About 60 percent of the sample entered the SOEP later than age 5. On average, individuals enter the sample at age 8.2.
[3] For items and measurement, please refer to Table A.3.
[4] The earliest year with non-missing observation.

after age five. In this case, covariates are likely to be measured after the start of the treatment. The age at which an individual (and thereby her or his parents) enters the SOEP is determined exogenously. In order to increase the total sample size, but also to compensate for panel attrition, new samples of the German population are regularly recruited. Recruitment always takes place at the household level. Thus, we can first observe the parental characteristics of an adolescent, who will answer the SOEP Youth Questionnaire at age 17, in the year her or his family was sampled for participating in the SOEP.

Finally, all estimations control for the SOEP sample the individual is part of. The SOEP consists of several samples that were added over time in order to increase the overall sample size of the SOEP. In some of these samples, parts of the population were overrepresented (foreigners, families with many children or high-income families). Including these sample fixed effects allows us to account for between-group differences which are inherent to the survey design.

Tables A.5, A.6 and A.7 show the summary statistics for all control and outcome variables used in this study.

Table A.5 – *Summary statistics of all outcome variables*

	Mean	Standard deviation	Minimum	Maximum	Sample size
Cognitive skills					
Average cognitive skills	0.06	1.00	-3.00	2.53	1,847
Analogies	0.07	1.01	-2.27	3.17	1,847
Figures	0.05	1.00	-2.79	2.67	1,847
Maths operators	0.02	0.99	-2.51	1.42	1,847
School grades[1]					
Average mark	-0.03	1.00	-3.39	4.35	3,488
German mark	-0.03	1.00	-2.82	3.99	3,488
Language mark	-0.03	1.00	-2.45	3.07	3,488
Maths mark	-0.02	1.01	-2.25	3.11	3,488
Personality					
Conscientiousness	0.03	1.00	-3.15	1.95	1,815
Openness	0.06	1.01	-3.36	2.20	1,815
Agreeableness	0.02	1.00	-4.75	1.82	1,815
Extraversion	0.01	1.00	-3.13	1.69	1,815
Neuroticism	0.01	1.01	-2.51	2.72	1,815
Perceived control	0.02	0.99	-3.84	2.52	3,488
Time use					
Watch TV daily	0.77	0.42	0.00	1.00	3,488
Read daily	0.26	0.44	0.00	1.00	3,488
Ambition					
Student aims Abitur	0.46	0.50	0.00	1.00	3,488
Student aims university	0.32	0.47	0.00	1.00	3,488
Job success likely	0.71	0.19	0.00	1.00	3,488
Desired job likely	0.70	0.21	0.00	1.00	3,488

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012 pooled), own calculations. Summary statistics of all outcome variables examined in this paper. The mean and standard deviation of normalized variables differs from zero because they were normalized within the control group.

Table A.6 – *Summary statistics of all control variables (part 1 of 2)*

	Mean	Standard deviation	Minimum	Maximum	Sample size
Motivation to play music					
Mother has no degree	0.16	0.37	0.00	1.00	3,941
Mother has completed only 9 years of school	0.28	0.45	0.00	1.00	3,941
Mother has vocational degree	0.67	0.47	0.00	1.00	3,941
Mother has university degree	0.30	0.46	0.00	1.00	3,941
Mother has migration background	0.18	0.38	0.00	1.00	3,941
Mother's age at birth	27.51	4.98	14.00	49.00	3,941
Parents care about school achievement	0.25	0.43	0.00	1.00	3,941
Parents do not support learning	0.23	0.42	0.00	1.00	3,941
Conflict with parents due to school results	0.54	0.50	0.00	1.00	3,941
Parents attend parent-teacher meeting	0.74	0.44	0.00	1.00	3,941
Parents attend teacher's consultation hours	0.56	0.50	0.00	1.00	3,941
Parents actively contact school teachers	0.22	0.42	0.00	1.00	3,941
Parents engage as parent representatives	0.17	0.37	0.00	1.00	3,941
Parents do not engage with the child's school	0.10	0.30	0.00	1.00	3,941
Conscientiousness (mother)	0.87	0.11	0.33	1.00	3,941
Extraversion (mother)	0.72	0.15	0.14	1.00	3,941
Agreeableness (mother)	0.81	0.12	0.29	1.00	3,941
Openness (mother)	0.66	0.16	0.14	1.00	3,941
Neuroticism (mother)	0.59	0.16	0.14	1.00	3,941
Mother never attends cultural events	0.36	0.48	0.00	1.00	3,941
Mother attends cultural events every month	0.11	0.31	0.00	1.00	3,941
Mother is never artistically active	0.51	0.50	0.00	1.00	3,941
Mother is artistically active every month	0.19	0.39	0.00	1.00	3,941
Mother's appreciation for the arts	0.63	0.24	0.14	1.00	3,941
More than 200 books at home	0.31	0.46	0.00	1.00	3,941
Fewer than 50 books at home	0.26	0.44	0.00	1.00	3,941
Constraints for playing music					
Log monthly net household income	7.70	0.52	5.39	9.90	3,941
Log monthly net household income (squared)	59.64	8.06	29.09	98.08	3,941
Number of siblings in the SOEP	1.45	1.19	0.00	11.00	3,941
Child is the firstborn	0.51	0.50	0.00	1.00	3,941
Rooms per person at home	1.14	0.44	0.23	6.00	3,941
Household lives in a rural area	0.25	0.43	0.00	1.00	3,941
Schleswig-Holstein	0.03	0.18	0.00	1.00	3,941
Hamburg	0.01	0.10	0.00	1.00	3,941
Lower Saxony	0.09	0.29	0.00	1.00	3,941
Bremen	0.01	0.08	0.00	1.00	3,941
North Rhine-Westphalia	0.21	0.41	0.00	1.00	3,941
Hesse	0.06	0.25	0.00	1.00	3,941
Rhineland-Palatinate	0.06	0.23	0.00	1.00	3,941

Source: SOEP v29 (2001-2012 pooled), own calculations. Summary statistics of all control variables included in the estimations of this paper.

Table A.7 – *Summary statistics of all control variables (part 2 of 2)*

	Mean	**Standard deviation**	**Minimum**	**Maximum**	**Sample size**
Constraints for playing music (continued)					
Baden-Württemberg	0.13	0.34	0.00	1.00	3,941
Bavaria	0.14	0.35	0.00	1.00	3,941
Saarland	0.01	0.09	0.00	1.00	3,941
Berlin	0.04	0.19	0.00	1.00	3,941
Brandenburg	0.04	0.19	0.00	1.00	3,941
Mecklenburg-Vorpommern	0.02	0.15	0.00	1.00	3,941
Saxony	0.07	0.25	0.00	1.00	3,941
Saxony-Anhalt	0.04	0.20	0.00	1.00	3,941
Adolescent's motivation to continue playing music					
Recommendation for upper secondary school	0.39	0.49	0.00	1.00	3,941
Recommendation for lower secondary school	0.15	0.36	0.00	1.00	3,941
Student attends upper secondary school	0.32	0.47	0.00	1.00	3,941
Girl	0.49	0.50	0.00	1.00	3,941
Born in 1984	0.08	0.28	0.00	1.00	3,941
Born in 1985	0.09	0.28	0.00	1.00	3,941
Born in 1986	0.09	0.29	0.00	1.00	3,941
Born in 1987	0.10	0.30	0.00	1.00	3,941
Born in 1988	0.09	0.29	0.00	1.00	3,941
Born in 1989	0.08	0.27	0.00	1.00	3,941
Born in 1990	0.09	0.28	0.00	1.00	3,941
Born in 1991	0.07	0.25	0.00	1.00	3,941
Born in 1992	0.06	0.24	0.00	1.00	3,941
Born in 1993	0.05	0.23	0.00	1.00	3,941
Survey-specific control variables					
Sample A	0.20	0.40	0.00	1.00	3,941
Sample B	0.06	0.25	0.00	1.00	3,941
Sample C	0.13	0.33	0.00	1.00	3,941
Sample D	0.04	0.19	0.00	1.00	3,941
Sample E	0.05	0.21	0.00	1.00	3,941
Sample F	0.38	0.49	0.00	1.00	3,941
Sample G	0.07	0.25	0.00	1.00	3,941
In survey at age 5	0.38	0.49	0.00	1.00	3,941
Missing indicators					
Missing: parents' personality	0.12	0.32	0.00	1.00	3,941
Missing: Household income	0.07	0.25	0.00	1.00	3,941
Missing: Cultural consumption parents	0.06	0.25	0.00	1.00	3,941
Missing: Number of books at home	0.03	0.18	0.00	1.00	3,941
Missing: Secondary school recommendation	0.02	0.15	0.00	1.00	3,941
Missing: mother's age, education, migration	0.03	0.17	0.00	1.00	3,941
Missing: other convariates	0.02	0.14	0.00	1.00	3,941

Source: SOEP v29 (2001-2012 pooled), own calculations. Summary statistics of all control variables included in the estimations of this paper.

A.2 Estimation of the propensity score and matching

This section describes in detail how the present study takes control variables into account using propensity score matching.

1. **Estimation of the propensity score.** The propensity score describes the probability to be treated given the covariates. It is estimated using a probit model. For a list of all covariates, please refer to Table A.4. The propensity score was estimated with the largest available sample of 3,941 observations, which excludes only those for whom information on the treatment status is not available (please refer to Table A.1 for details on how the sample was constructed).

   Table A.8 shows the marginal effects for all covariates of the probit model estimating the propensity score. The treatment definition is given in the table header. Standard errors are clustered at the family level and are given in parentheses. All coefficients and standard errors are indicated in percent. Many coefficients are not significant, which is partly due to the fact that most covariates are strongly correlated with each other. We still include these insignificant coefficients in the estimation of the propensity score, given that the aim is to balance all covariates as well as possible (Stuart, 2010). Efron's R-Square is 0.178 percent, which indicates that the model is rather well able to explain adolescents' participation in music training.

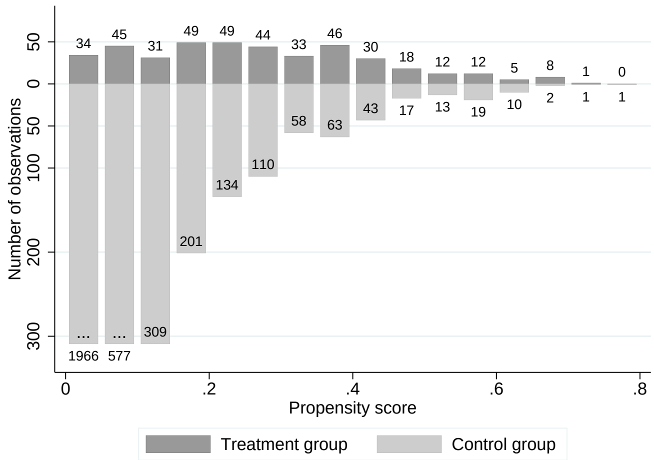
   All coefficients point in the direction we would expect, according to which adolescents from more favorable socio-economic backgrounds are more likely to play and learn a musical instrument during their childhood. Some variables are particularly important in explaining the decision to take up music lessons. Everything else equal, an adolescent whose mother has a university degree is 3.3 percent more likely to make music according to our treatment definition. The parents' taste for the arts is also decisive. An appreciation for artistic experiences as well as regular attendance at cultural events are factors which are characteristic for parents with musically active children. Moreover, the parents' cultural capital, measured by the number of books in the household, is strongly correlated with musical activities. Having more than 200 books increases the likelihood that the adolescent plays music by 2.6 percent. Having fewer than 50 books decreases it to the same extent. As argued in the main

Table A.8 – *Estimation of the propensity score (probit model, marginal effects, in percent)*

Play music at age 17, started at age 8 or before, take music lessons		
Motivation to play music		
Mother has no degree	-0.2	(1.8)
Mother has completed only 9 years of school	-0.4	(1.0)
Mother has vocational degree	0.8	(1.0)
Mother has university degree	3.3***	(0.9)
Mother has migration background	0.6	(1.3)
Mother's age at birth	0.0	(0.1)
Parents care about school achievement	0.7	(0.8)
Parents do not support learning	-0.5	(0.8)
Conflict with parents due to school results	0.1	(0.7)
Parents attend parent-teacher meeting	1.0	(1.0)
Parents attend teacher's consultation hours	-0.6	(0.7)
Parents actively contact school teachers	0.1	(0.8)
Parents engage as parent representatives	2.7***	(0.9)
Parents do not engage with the child's school	-1.1	(1.6)
Conscientiousness (mother)	-0.4	(3.4)
Extraversion (mother)	-1.0	(2.7)
Agreeableness (mother)	-1.5	(3.0)
Openness (mother)	-3.7	(4.0)
Neuroticism (mother)	2.9	(2.2)
Mother never attends cultural events	-1.9^{+}	(1.0)
Mother attends cultural events every month	2.6***	(1.0)
Mother is never artistically active	-2.3**	(0.9)
Mother is artistically active every month	-0.1	(0.9)
Mother's appreciation for the arts	7.0***	(2.6)
More than 200 books at home	2.6***	(0.9)
Fewer than 50 books at home	-2.6**	(1.3)
Constraints for playing music		
Log monthly net household income	-7.9	(12.1)
Log monthly net household income (squared)	0.5	(0.8)
Number of siblings in the SOEP	0.3	(0.4)
Child is the firstborn	-0.6	(0.7)
Rooms per person at home	1.7**	(0.8)
Household lives in a rural area	-0.9	(1.1)
Adolescent's motivation to continue playing music		
Recommendation for upper secondary school	4.9***	(0.9)
Recommendation for lower secondary school	-0.7	(1.5)
Student attends upper secondary school	1.9**	(0.9)
Girl	4.6***	(0.7)
Missing indicators, birth and region fixed effects	Yes	
Number of observations	3,941	
Efron's R-Square	0.178	

Source: SOEP v29 (2001-2012 pooled), own calculations. Probit model estimating the probability to be treated. Treatment definition: Play music at age 17, started at age 8 or before, take music lessons. In addition, the estimation controls for an indicator whether the individual entered the SOEP after age 5, as well as SOEP sample indicators (see Appendix A for details). Standard errors in parentheses are clustered at the family level. Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

Figure A.1 – *Common support*



text of this study, financial resources of the household play an insignificant role, once the other characteristics are taken into account. In Table A.8, the marginal effect of household income is even negative, albeit insignificant.

2. **Check common support.** Comparing the distribution of the propensity score between treatment and control group indicates whether we can find comparable control observations for each treated individual. Figure A.1 provides graphical evidence that this common support assumption holds. The graph shows the number of observations in the treatment (dark grey) and control group (light grey) along the distribution of the propensity score. The majority of individuals in the control group have very low predicted probabilities to be part of the treatment group. Still, due to the high number of observations in the control group, we observe a sufficient number of control observations with high propensity scores as well.

3. **Radius matching.** We construct a control group, which is comparable to the treatment group by using radius matching with replacement and a caliper of one percent. This allows us to benefit from the large number of control observations (only about ten percent of the observations in our sample belong to the treatment group). For each treated observation, we match all individuals of the control group with the same propensity score plus or minus one percent. Each of these matched observations receives a weight of 1 divided by the number of matched control observations for that

treated individual. Thus, the weights of all control observations, which are matched to a treated individual, sum to one.

4. **Check balancing of covariates.** Table A.9 shows the mean differences between treatment and control groups for each covariate before and after matching. We can see that after matching there remain virtually no differences between both groups. Matching thus successfully allowed us to construct a control group, which is comparable to the treatment group in terms of observable characteristics.

5. **Calculate outcome difference.** The estimation results presented in all outcome tables of this paper are obtained by calculating the mean outcome difference between the treatment and matched control group. Algebraically, we estimate:

$$A\hat{T}T = \frac{1}{N_T}\sum_{i=1}^{N_T} T_i y_i - \frac{1}{N_C}\sum_{i=1}^{N_C}(1 - T_i)\hat{w}_i y_i \tag{A.1}$$

where N_T and N_C are the number of treated and control observations, T_i is the treatment indicator, and y_i the outcome for individual i. Control observations are weighted with weight $\hat{w}_i$, which is obtained from matching as described above. Standard errors are clustered at the family level and estimated by bootstrap with 1999 replications.

Table A.9 – *Balancing of covariates after propensity score matching*

	Before matching		**After matching**	
	Difference	t-value	Difference	t-value
Motivation to play music				
Mother has no degree	-0.12***	-6.09	0.01	0.49
Mother has completed only 9 years of school	-0.18***	-7.66	0.01	0.42
Mother has vocational degree	-0.05**	-2.17	0.02	0.52
Mother has university degree	0.35***	15.35	-0.01	-0.16
Mother has migration background	-0.09***	-4.44	-0.01	-0.36
Mother's age at birth	2.10***	8.22	-0.16	-0.49
Parents care about school achievement	0.02	0.78	0.01	0.17
Parents do not support learning	-0.04^{+}	-1.74	0.00	0.15
Conflict with parents due to school results	-0.07***	-2.64	-0.03	-0.84
Parents attend parent-teacher meeting	0.10***	4.17	0.01	0.29
Parents attend teacher's consultation hours	-0.00	-0.03	0.00	0.07
Parents actively contact school teachers	0.03	1.49	-0.00	-0.16
Parents engage as parent representatives	0.16***	8.29	-0.01	-0.16
Parents do not engage with the child's school	-0.05***	-3.06	-0.01	-0.54
Conscientiousness (mother)	-0.01**	-2.17	-0.00	-0.04
Extraversion (mother)	0.00	0.44	-0.01	-0.78
Agreeableness (mother)	-0.01	-1.24	-0.00	-0.20
Openness (mother)	0.04***	4.47	-0.01	-0.51
Neuroticism (mother)	-0.01	-0.98	0.00	0.19
Mother never attends cultural events	-0.26***	-10.56	0.00	0.13
Mother attends cultural events every month	0.18***	11.04	0.01	0.32
Mother is never artistically active	-0.26***	-10.08	-0.01	-0.18
Mother is artistically active every month	0.19***	9.19	0.03	0.89
Mother's appreciation for the arts	0.09***	7.31	-0.00	-0.13
More than 200 books at home	0.34***	14.49	-0.00	-0.00
Fewer than 50 books at home	-0.22***	-9.96	0.00	0.05
Constraints for playing music				
Log monthly net household income	0.32***	11.95	-0.00	-0.13
Log monthly net household income (squared)	4.99***	12.18	-0.07	-0.12
Number of siblings in the SOEP	0.01	0.18	0.02	0.22
Child is the firstborn	-0.04	-1.36	0.02	0.54
Rooms per person at home	0.17***	7.35	0.01	0.40
Household lives in a rural area	-0.06**	-2.46	0.00	0.03
Adolescent's motivation to continue playing music				
Recommendation for upper secondary school	0.40***	16.27	-0.01	-0.25
Recommendation for lower secondary school	-0.12***	-6.43	0.00	0.19
Student attends upper secondary school	0.35***	14.80	-0.01	-0.24
Girl	0.16***	6.25	0.01	0.21

Source: SOEP v29 (2001-2012 pooled), own calculations. Differences between adolescents with and without music training before and after matching. Further variables used for matching, but not displayed in this table are: missing indicators, birth and region fixed effects, an indicator whether the individual entered the SOEP after age 5, SOEP sample indicators. Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

A.3 Further results

Robustness I: Heterogeneous effects

Tables A.11 and A.12 show the heterogeneity of outcome differences between musically active and inactive adolescents by examining different subgroups. Estimations were separately carried out by gender, living area, household income, parental education, cultural capital as well as the type of secondary school the adolescent attends. All estimations use the same treatment definition, according to which an adolescent plays music at age 17, has started no later than age 8 and takes or has taken music lessons outside of school. The sample size of the treatment and control group for each subgroup is given in Table A.10.

A comparison of the share of musically active individuals in the different subgroups reveals an expected pattern. Music is more prevalent in families with higher socio-economic status. The most pronounced differences are observed if we distinguish by the household's cultural capital and by the type of secondary school attended by the adolescent. In both

Table A.10 – *Sample sizes for heterogeneous results*

Sample definition (Column in Table A.11 / A.12)	**Control group**	**Treatment group**	**% treated**
Full sample from baseline specification	3,524	417	10.6
Subgroups by gender:			
Male (1)	1,844	151	7.6
Female (2)	1,680	266	13.7
Subgroups by living area:			
Urban area (3)	2,620	333	11.3
Rural area (4)	904	84	8.5
Subgroups by household income:			
Household income below median (5)	1,738	114	6.2
Household income above median (6)	1,786	303	14.5
Subgroups by parental education:			
Mother without Abitur/university degree (7)	2,068	115	5.3
Mother with Abitur/university degree (8)	1,371	298	17.9
Subgroups by cultural capital:			
Fewer than 200 books at home (9)	2,560	162	6.0
More than 200 books at home (10)	964	255	20.9
Subgroup by type of secondary school attended:			
Adolescent does not attend upper secondary school (11)	2,534	155	5.8
Adolescent attends upper secondary school (12)	990	262	20.9

Source: SOEP v29 (2001-2012 pooled), own calculations. Number of treatment and control observations for the subgroup estimations presented in Tables A.11 and A.12. Sample sizes are given for the full sample with which the propensity score is estimated. Sample sizes for the three outcome groups (general, cognitive skills, Big Five) can be provided by the authors on request.

cases, 21 percent of the more advantaged group play and learn a musical instrument and have done so since age eight or earlier. Only 6 percent of the adolescents from vocational school tracks and with fewer books at home are engaged with music. Stark differences in attendance at extracurricular music lessons can also be observed if we distinguish by parental education and income. 15 to 18 percent of the children from richer and more educated parents play music according to our treatment definition. Only around 5 percent of those from less highly educated and poorer families do so. Unsurprisingly, girls are more likely to play music than boys (14 versus 8 percent). Probably related to opportunities in the neighborhood of the household, adolescents in urban areas are more likely to engage with music than those living in rural areas (11 versus 9 percent).

Heterogeneities in outcome differences between musically active and inactive adolescents can be summarized as follows. The association between musical activities and better school grades is considerably stronger among adolescents from less educated family backgrounds. Coefficients are higher in the subgroups with less highly educated parents, with fewer books at home and attending medium or lower secondary school. On average, school grades among the musically active in these groups are one fifth of a standard deviation above those of their musically inactive peers. Among adolescents from more highly educated backgrounds, these differences are at most half as large and insignificant. Surprisingly, the distinction by household income provides conflicting results. Outcome differences with respect to school grades between those who play music and those who do not can only be observed among those with greater household income. In addition, musically active adolescents who attend upper secondary school have better cognitive skills than their non-musical peers, a difference which amounts to one fifth of a standard deviation.

If we distinguish by gender and living area (rural or urban), we observe almost no heterogeneities. However, outcome differences in conscientiousness are larger for girls and in urban areas. Furthermore, musically active boys, as well as adolescents from families with few books at home, have a higher perception of control than their peers. Other than that, only small and insignificant heterogeneities can be observed with respect to personality in all subgroups. Insignificant coefficients, which are due to the small number of treated individuals in some of the subgroups, make more detailed comparisons impossible.

Table A.11 – *Outcome differences between adolescents with and without music training (heterogeneous effects, part 1 of 2)*

	Gender				**Living area**				**Household income**			
	male		*female*		*not rural*		*rural*		*<median*		*>median*	
	(1)		(2)		(3)		(4)		(5)		(6)	
Cognitive skills (in std. dev.)												
Average cognitive skills	0.02	(0.16)	0.05	(0.11)	0.09	(0.09)	0.11	(0.25)	-0.07	(0.19)	0.10	(0.10)
Analogies	0.07	(0.16)	0.14	(0.12)	0.13	(0.10)	0.24	(0.31)	0.10	(0.21)	0.07	(0.11)
Figures	0.09	(0.17)	0.03	(0.11)	0.11	(0.10)	0.03	(0.31)	0.04	(0.21)	0.11	(0.11)
Maths operators	-0.08	(0.15)	-0.04	(0.12)	0.01	(0.10)	0.00	(0.29)	-0.23	(0.19)	0.06	(0.10)
School grades[1] *(in std. dev.)*												
Average school grade	-0.15	(0.13)	-0.17**	(0.09)	-0.17**	(0.07)	-0.15	(0.25)	0.01	(0.17)	-0.20**	(0.08)
German grade	-0.18	(0.13)	-0.15^{+}	(0.08)	-0.15**	(0.07)	-0.11	(0.24)	-0.08	(0.17)	-0.16^{+}	(0.08)
Foreign language grade	-0.12	(0.14)	-0.12	(0.09)	-0.14^{+}	(0.07)	-0.17	(0.26)	0.07	(0.17)	-0.20**	(0.08)
Mathematics grade	-0.06	(0.13)	-0.12	(0.10)	-0.11	(0.08)	-0.08	(0.26)	0.02	(0.17)	-0.11	(0.09)
Personality (in std. dev.)												
Conscientiousness	0.17	(0.16)	0.29**	(0.14)	0.24**	(0.11)	0.14	(0.34)	0.36	(0.23)	0.16	(0.12)
Openness	0.45***	(0.17)	0.11	(0.13)	0.19^{+}	(0.11)	0.42	(0.34)	0.33^{+}	(0.20)	0.22^{+}	(0.12)
Agreeableness	0.15	(0.17)	0.09	(0.13)	0.07	(0.10)	0.22	(0.34)	0.31	(0.21)	0.05	(0.11)
Extraversion	-0.04	(0.17)	-0.07	(0.13)	-0.04	(0.10)	-0.12	(0.34)	-0.12	(0.23)	0.02	(0.11)
Neuroticism	0.01	(0.16)	0.07	(0.13)	-0.01	(0.11)	0.22	(0.34)	0.17	(0.23)	-0.01	(0.12)
Perceived control	0.18	(0.12)	-0.07	(0.09)	0.01	(0.07)	-0.09	(0.25)	0.09	(0.16)	0.02	(0.08)
Time use (in %)												
Watch TV daily	-8.73	(6.39)	-9.41**	(4.74)	-9.38**	(3.85)	-24.71**	(11.49)	-16.02^{+}	(8.23)	-9.38**	(4.33)
Read books daily	3.36	(6.30)	7.52	(4.77)	10.87***	(3.95)	5.61	(12.21)	4.07	(8.08)	6.31	(4.42)
Ambition (in %)												
Aim Abitur	1.63	(4.44)	4.38	(3.08)	3.55	(2.32)	3.99	(10.35)	8.99	(7.63)	2.64	(2.72)
Aim university	11.21**	(5.37)	7.12^{+}	(4.19)	11.24***	(3.36)	4.72	(12.06)	14.72^{+}	(8.27)	6.61^{+}	(3.65)
Job success likely	1.81	(2.20)	0.69	(1.53)	-0.54	(1.29)	6.17	(3.79)	4.81^{+}	(2.87)	-0.74	(1.47)
Desired profession likely	3.41	(2.46)	0.94	(1.78)	0.12	(1.44)	7.30	(4.64)	6.96**	(3.27)	0.00	(1.69)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. The table shows the effects of music for different subgroups of the population. (1) Boys. (2) Girls. (3) Not living in rural areas. (4) Living in rural areas. (5) Household income below median. (6) Household income above median. Propensity score matching is used to account for control variables (radius matching with caliper 0.01). Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}\ p < 0.1 **\ p < 0.05 ***\ p < 0.01$

Table A.12 – *Outcome differences between adolescents with and without music training (heterogeneous effects, part 2 of 2)*

	Parents' education				**Cultural capital**				**Secondary school**			
	low		*high*		*low*		*high*		*lower*		*upper*	
	(7)		(8)		(9)		(10)		(11)		(12)	
Cognitive skills (in std. dev.)												
Average cognitive skills	0.08	(0.16)	0.09	(0.10)	0.06	(0.13)	0.05	(0.12)	0.14	(0.12)	0.22^{+}	(0.11)
Analogies	0.14	(0.18)	0.12	(0.11)	0.17	(0.14)	-0.01	(0.13)	0.23	(0.14)	0.21^{+}	(0.12)
Figures	0.05	(0.17)	0.07	(0.11)	0.14	(0.14)	0.08	(0.13)	0.15	(0.14)	0.21^{+}	(0.12)
Maths operators	0.01	(0.17)	0.03	(0.10)	-0.12	(0.13)	0.04	(0.13)	-0.00	(0.14)	0.11	(0.12)
School grades[1] (in std. dev.)												
Average school grade	-0.21	(0.14)	-0.12	(0.08)	-0.23**	(0.11)	-0.03	(0.10)	-0.23^{+}	(0.12)	-0.08	(0.09)
German grade	-0.18	(0.14)	-0.13	(0.08)	-0.14	(0.11)	-0.11	(0.10)	-0.31**	(0.12)	-0.04	(0.09)
Foreign language grade	-0.09	(0.15)	-0.10	(0.08)	-0.09	(0.11)	-0.05	(0.10)	-0.13	(0.12)	-0.08	(0.09)
Mathematics grade	-0.19	(0.15)	-0.06	(0.09)	-0.28**	(0.13)	0.08	(0.10)	-0.11	(0.12)	-0.05	(0.10)
Personality (in std. dev.)												
Conscientiousness	0.25	(0.19)	0.17	(0.13)	0.23	(0.14)	0.13	(0.15)	0.20	(0.14)	0.16	(0.15)
Openness	0.35^{+}	(0.18)	0.22^{+}	(0.12)	0.43***	(0.16)	0.12	(0.15)	0.29^{+}	(0.15)	0.21	(0.13)
Agreeableness	0.16	(0.17)	0.10	(0.12)	0.23	(0.15)	0.09	(0.14)	0.16	(0.16)	0.07	(0.13)
Extraversion	0.07	(0.18)	0.01	(0.12)	-0.04	(0.15)	0.00	(0.14)	-0.07	(0.17)	0.04	(0.12)
Neuroticism	-0.13	(0.20)	0.08	(0.12)	0.02	(0.16)	-0.00	(0.14)	0.12	(0.16)	-0.00	(0.14)
Perceived control	0.15	(0.14)	-0.01	(0.08)	0.27**	(0.11)	-0.10	(0.09)	-0.00	(0.11)	0.06	(0.09)
Time use (in %)												
Watch TV daily	-7.14	(6.55)	-12.17***	(4.38)	-12.99**	(5.37)	-10.13**	(5.12)	-14.38**	(5.62)	-9.55**	(4.75)
Read books daily	3.81	(7.05)	7.67^{+}	(4.66)	14.14***	(5.44)	1.11	(5.37)	9.12	(5.65)	4.50	(4.85)
Ambition (in %)												
Aim Abitur	4.60	(6.33)	3.47	(2.79)	6.04	(4.65)	2.90	(3.32)	6.79	(5.68)	1.99	(2.29)
Aim university	7.43	(6.76)	8.57**	(3.95)	8.81	(5.49)	7.55	(4.59)	6.84	(5.84)	6.80^{+}	(4.02)
Job success likely	1.51	(2.45)	0.64	(1.53)	3.70^{+}	(2.06)	-0.06	(1.70)	0.74	(2.14)	1.13	(1.62)
Desired profession likely	5.47^{+}	(2.95)	0.76	(1.73)	5.49**	(2.24)	-0.79	(2.06)	1.80	(2.32)	1.51	(1.87)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. The table shows the effects of music for different subgroups of the population. (7) Mother does not have upper secondary school or university degree. (8) Mother has upper secondary school or university degree. (9) Fewer than 200 books at home. (10) More than 200 books at home. (11) Adolescent attends lower or medium secondary school. (12) Adolescent attends upper secondary school (Gymnasium). Propensity score matching is used to account for control variables (radius matching with caliper 0.01). Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}\ p < 0.1 **\ p < 0.05 ***\ p < 0.01$

Robustness II: Varying treatment definitions

Table A.14 shows outcome differences between musically active and inactive adolescents for different treatment definitions. The sample sizes for each treatment and control group are given in Table A.13. All alternative definitions are variants of the one used in the baseline specification of this paper, according to which an individual plays a musical instrument at age 17, has started to do so at age 8 or before and has taken music lessons outside of school. In columns (1) to (3) of Table A.14, one or both of the latter two restrictions are left out progressively. Column (4) examines adolescents who play music and take music lessons, but have only started to play at age 9 or after. This allows distinguishing whether the focus on playing music in childhood drives our results. The sample size is smaller in column (4), given that adolescents who started to play music before age 9 were dropped. Finally, we extend the main definition by additionally requiring individuals to play classical music (column 5) or to play music at least on a weekly basis (column 6). For the last two definitions, the sample size is slightly smaller as well, due to missing answers on the variables required to define them.

Table A.13 – *Sample sizes for treatment definitions*

Treatment definition (Column in Table A.14)	**Control group**	**Treatment group**	**% treated**
Main definition: Play music at age 17, started at age 8 or before, take music lessons	3,524	417	10.6
Main definition without music lessons (1): Play music at 17, started at age 8 or before	3,476	465	11.8
Main definition without age restriction (2): Play music at age 17, take music lessons	3,232	709	18.0
Main definition without age restriction and lessons (3): Play music at age 17	2,978	963	24.4
Main definition with different age restriction (4): Play music at 17, started at age 9 or later, take music lessons	3,267	257	7.3
Main definition, classical music only (5): Music at 17, started at 8 or before, have lessons, play classical music	3,649	154	4.0
Main definition, play at least weekly (6): Music at 17, started at 8 or before, have lessons, play weekly	3,528	360	9.3

Source: SOEP v29 (2001-2012 pooled), own calculations. Number of treatment and control observations for the treatment definitions presented in Table A.14. Sample sizes are given for the full sample with which the propensity score is estimated. Sample sizes for the three outcome groups (general, cognitive skills, Big Five) can be provided by the authors on request.

Comparing the results obtained from various treatment definitions reveals the following pattern. First, if we drop the requirement that the individual takes music lessons outside of school (columns 1 and 3), all effects on school grades become much smaller. This could either be due to the fact that music lessons are somewhat similar to school lessons, which explains the positive effect on the latter. Alternatively, adolescents with better school grades prior to playing music might be more inclined to take lessons rather than only playing music for themselves.

Second, if we drop the age restriction or look at adolescents who started to play music rather late (columns 2, 3 and 4), the effect on conscientiousness vanishes. Such a result supports the explanation that long-term exposure to music fosters non-cognitive skills such as conscientiousness. However, it contradicts the competing hypothesis that individuals who start to play music in their adolescence are those who are particularly conscientious.

Third, if we consider adolescents who started to play music later than age 8 (column 4), all effects are much smaller. This is not astonishing if we believe that the duration of exposure to music plays a role.

Fourth, in addition to outcome differences which we find for all treatment definitions, adolescents who play and learn classical music (column 5) score much better in the cognitive skills test. On average, they score one fifth of a standard deviation above those who play other types of music or do not play music at all. This confirms previous findings, according to which music improves cognitive skills (Schellenberg, 2004). It also raises the question whether only classical music affects intelligence or whether the more intelligent individuals play classical rather than other types of music.

Finally, the only effect which is stable both in significance and magnitude throughout all treatment definitions is the aim to attend university. This is not surprising if we consider that this outcome variable is only the affirmation of an intention, expressed at age 17. Current exposure to music is likely to be more relevant than past exposure in determining this intention. However, all treatment definitions involve being active in music at age 17.

Table A.14 – *Outcome differences between adolescents with and without music training (different treatment definitions)*

	≤ *age 8 & 17*		*age 17 lessons*		*age 17*		> *age 8 & 17 lessons*		≤ *age 8 & 17 lessons classical*		≤ *age 8 & 17 lessons weekly*	
	(1)		(2)		(3)		(4)		(5)		(6)	
Cognitive skills (in std. dev.)												
Average cognitive skills	0.06	(0.07)	0.05	(0.06)	0.00	(0.06)	0.05	(0.09)	0.22	(0.14)	0.04	(0.08)
Analogies	0.11	(0.08)	0.09	(0.07)	0.06	(0.06)	0.07	(0.10)	0.21	(0.15)	0.08	(0.09)
Figures	0.09	(0.08)	0.10	(0.07)	0.09	(0.06)	0.12	(0.10)	0.21	(0.15)	0.12	(0.09)
Maths operators	-0.03	(0.08)	-0.05	(0.07)	-0.11^{+}	(0.06)	-0.04	(0.10)	0.12	(0.15)	-0.07	(0.09)
School grades[1] *(in std. dev.)*												
Average school grade	-0.13**	(0.06)	-0.18***	(0.05)	-0.13***	(0.04)	-0.12	(0.08)	-0.21^{+}	(0.12)	-0.11^{+}	(0.07)
German grade	-0.13**	(0.06)	-0.16***	(0.05)	-0.14***	(0.05)	-0.14^{+}	(0.08)	-0.16	(0.11)	-0.12^{+}	(0.07)
Foreign language grade	-0.12^{+}	(0.06)	-0.06	(0.05)	-0.06	(0.05)	0.02	(0.08)	-0.17	(0.11)	-0.11	(0.07)
Mathematics grade	-0.06	(0.06)	-0.17***	(0.06)	-0.10**	(0.05)	-0.16^{+}	(0.08)	-0.14	(0.13)	-0.05	(0.07)
Personality (in std. dev.)												
Conscientiousness	0.20**	(0.09)	0.14^{+}	(0.08)	0.08	(0.07)	0.06	(0.11)	0.16	(0.16)	0.23**	(0.09)
Openness	0.29***	(0.09)	0.33***	(0.08)	0.32***	(0.07)	0.34***	(0.11)	0.10	(0.18)	0.31***	(0.10)
Agreeableness	0.12	(0.08)	0.14^{+}	(0.07)	0.14**	(0.07)	0.17	(0.11)	-0.06	(0.16)	0.11	(0.09)
Extraversion	-0.08	(0.09)	0.02	(0.08)	0.04	(0.07)	0.08	(0.11)	-0.13	(0.16)	-0.04	(0.09)
Neuroticism	0.12	(0.09)	0.06	(0.07)	0.08	(0.07)	0.05	(0.11)	0.02	(0.18)	0.02	(0.09)
Perceived control	-0.00	(0.06)	0.04	(0.05)	0.04	(0.04)	0.05	(0.08)	-0.02	(0.11)	0.07	(0.06)
Time use (in %)												
Watch TV daily	-10.08***	(3.02)	-7.64***	(2.63)	-5.04**	(2.18)	-2.67	(3.51)	-11.40**	(5.67)	-12.48***	(3.57)
Read books daily	5.69^{+}	(3.20)	6.02**	(2.78)	7.67***	(2.31)	5.73	(3.71)	9.00	(6.08)	7.72**	(3.68)
Ambition (in %)												
Aim Abitur	3.61^{+}	(2.14)	6.08***	(1.84)	4.87***	(1.68)	7.26**	(3.02)	1.81	(3.59)	4.36^{+}	(2.23)
Aim university	8.10***	(2.73)	9.49***	(2.36)	9.61***	(2.11)	8.79**	(3.66)	9.87**	(4.89)	7.96**	(3.10)
Job success likely	0.35	(1.04)	0.49	(0.94)	-0.26	(0.83)	0.41	(1.50)	-1.54	(2.09)	1.14	(1.15)
Desired profession likely	1.37	(1.17)	1.65	(1.05)	0.61	(0.95)	2.54	(1.64)	-0.36	(2.34)	2.32^{+}	(1.34)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. The table shows the effects of music for different treatment definitions. (1) Play music at 17, started at age 8 or before. (2) Play music at age 17, take music lessons. (3) Play music at age 17. (4) Play music at 17, started at age 9 or later, take music lessons. (5) Music at 17, started at 8 or before, have lessons, play classical music. (6) Music at 17, started at 8 or before, have lessons, play weekly. See Table A.13 for the number of treated and control observations of each treatment definition. Propensity score matching is used to account for control variables (radius matching with caliper 0.01). Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

Robustness III: Varying specifications and subsamples

In order to further examine the robustness of our estimations, we carry out a variety of robustness checks, the results of which are presented in Tables A.16 and A.17. Due to variations in sample composition and treatment definitions, sample sizes are not necessarily comparable between all robustness estimations. Therefore, Table A.15 presents the sample sizes of treatment and control group for each robustness specification.

Table A.15 – *Sample sizes for robustness checks*

Estimation details (Column in Table A.16 and A.17)	**Control group**	**Treatment group**	**% treated**	**Treatment definition**
OLS estimation (1)	3,524	417	10.6	Main
Propensity score matching: radius matching (2), *caliper 0.02 instead of 0.01 in main specification*	3,524	417	10.6	Main
Propensity score matching: Kernel matching (3), *Epanechnikov kernel*	3,524	417	10.6	Main
Subsample: complete information all outcomes (4), *one sample for all outcomes*	1,074	165	13.3	Main
Subsample: complete information all covariates (5), *no missing indicators*	2,607	327	11.1	Main
Subsample: individuals in sample before treatment started (6), *other treated individuals dropped*	3,524	171	4.6	Main
Subsample: Individuals in sample at age five (7), *when covariates were measured*	1,345	151	10.1	Main
Controls: add number of siblings X household income and firstborn X household income (8), *keep original control variables*	3,524	417	10.6	Play music
Controls: take father's characteristics instead of mother's (use mother's if father's not available) (9), *replace original control variables*	3,524	417	10.6	Play music
Treatment: play music at age 17 (10), *original treatment group dropped*	2,978	546	15.5	Play music
Treatment: play music at least sometimes (11), *original treatment group dropped*	2,067	1,405	40.5	Some music
Treatment: music age 12, gave up before 17 (12), *subsample with complete music history*	223	55	19.8	Give up

Source: SOEP v29 (2001-2012 pooled), own calculations. Number of treatment and control observations for the robustness checks presented in Tables A.16 and A.17. Treatment definitions (right column): *Main definition:* Play music at age 17, started at age 8 or before, take music lessons. *Play music:* Answer yes to the question Do you play a musical instrument or pursue singing seriously? *Some music:* Answer daily, weekly, monthly or less often (but not never!) to the question assessing the frequency of playing music. *Give up:* Play music at age 12, but give up before age 17. Sample sizes are given for the full sample with which the propensity score is estimated. Sample sizes for the three outcome groups (general, cognitive skills, Big Five) can be provided by the authors on request.

The pattern of sample sizes is as expected – approximately ten percent of the individuals are treated, if we consider the treatment definition used in the main specification of this paper. Unsurprisingly, if we drop parts of the treatment group, the share of treated individuals is lower (as in specification 6, where we drop treated individuals who entered our sample after starting to play music). A broader treatment definition (as in specifications 10 and 11, where we consider as treated those who had some contact with music) leads to a larger share of treated individuals. In specification (4), we restrict our estimations to those individuals for whom we have valid information on all outcome variables (rather than examining the results by group of outcomes). Here, the share of treated individuals is slightly higher, which might be due to non-random non-response in the outcomes. However, an estimation of the effect of music on a variable coded to "1" if at least one outcome is missing (and "0" if we have complete information about all outcomes for an individual) shows no significant coefficient. Thus, we have no evidence that playing music is correlated with missing outcome information.

Specifications (1) to (7) confirm the robustness of this paper's main results by applying different estimation methods and sample restrictions. Varying the estimation method by using ordinary least squares (OLS), by doubling the caliper for radius matching as well as by applying kernel rather than radius matching provides us with exactly the same results as found in the baseline specification (columns 1 to 3). Changing the sample we study yields similar results as well, with some outcomes being insignificantly different. Column (4) restricts the sample to observations for which we observe all outcome variables. Doing so considerably reduces the sample size, which is the reason why we created groups of outcome variables in the main specification. Column (5) restricts the sample to individuals for whom we observe all covariates, rather than recoding missing covariates and adding a missing indicator, as we do in the main specification. In column (6), we check whether our results are robust to the fact, that covariates were measured after the start of music classes for some individuals. We restrict the treatment group to those who we observe in the sample before they start to engage with music. This restriction does not alter our results, despite a small and negative effect for cognitive skills. Similarly, in column (7), we restrict the sample to those for whom we can measure the covariates at age 5. Measuring all covariates at the same age provides us with a more homogeneous and therefore comparable sample. Still, the results do not change, despite small negative effects in the test of mathematics operators.

Specifications (8) and (9) modify the control variables used in the estimations. In column (8), we maintain all covariates of the main specification, but additionally control for two interaction terms. Given that household income might influence music participation differently, depending on the number of siblings in the household, we control for a variable interacting the number of siblings with log monthly net household income. Moreover, we interact this income with an indicator of whether the individual is the firstborn child of his parents. In column (9), we construct our control variables with information on the father rather than on the mother, as in our main specification. Our results are robust to these modifications.

Specifications (10) to (12) address the issue of partly treated individuals in the control group. Due to our treatment definition, some individuals in the control group played music as well, but less intensely. This is true for those who play music at age 17, but started later than age 8 (our main treatment definition considers adolescents who started at age 8 or before). Partly treated are also those who play music, but did not receive music lessons outside of school. Finally, some adolescents state that they do not "play a musical instrument or pursue singing seriously" (question 16 of the SOEP Youth Questionnaire), but still indicate that they play music daily, weekly, monthly or less often, but not never (question 15 of the SOEP Youth Questionnaire).

Columns (10) and (11) of Table A.17 estimate the effect of music for these partly treated adolescents. In both columns, those who were treated according to our main specification (play music at age 17, have started at age 8 or before, have music lessons outside of school) are dropped. The respective group is compared to those who do not engage with music, examining how being partly treated is associated with cognitive and non-cognitive skills. Outcome differences of the opposite direction, compared to our main specification would indicate that partly treated individuals respond to music negatively. In this case, we which would overestimate the true effects of music, because we would add the positive effect of playing music according to our main definition to the negative effect of playing music less intensely, rather than subtracting these from each other. Columns (10) and (11) show that adolescents who play music at age 17, but have done so less intensely score either weekly better or similar to those who do not play music at 17.

A major shortcoming of our data is related to the fact that we cannot observe adolescents who played music when they were young, but gave up before age 17. These individuals are part of our control group, because the SOEP Youth Questionnaire does not ask them about past musical activities. Those who answer “No” to question 16 “Do you play a musical instrument or pursue singing seriously” skip the following four questions on the details of involvement with music. Around 60 percent of all musically active children give up in their early teenage years. The three most important reasons are a lack of motivation, critical life events, and dissatisfaction with the teacher (Switlick and Bullerjahn, 1999). Hence, ending music training is possibly related to weaker school performance. We can test this hypothesis using the SOEP household questionnaire, which has been asking parents about their child’s leisure time activities on a biannual basis since 2006. As the interview year depends on birth year, this allows us to construct a random subsample of 328 individuals, for whom we observe the complete history of musical activities since age 12. In this subsample, we compare outcome differences between those who never played a musical instrument after age 12 to those who played a musical instrument at age 12 and gave up before age 17. Similar to all other estimations, we apply propensity score matching to account for observable family background characteristics. However, due to the small sample size, we had to omit some covariates in order to avoid collinearity. Moreover, the sample size is too small to estimate standard errors by bootstrap. We present standard errors which do not take into account that the propensity score was estimated.

The effects of music for this subsample are presented in column (12) of Table A.17. Due to the small sample size, none of the outcome differences are significant. Still we see that the direction of the effect points in the same direction for most outcomes, even though magnitudes are greatly reduced. However, individuals who gave up music seem to have school grades below those who were never involved with music. Our main results might therefore overestimate the true effect of music on school grades, given that some of the partly treated individuals might actually suffer from music practice. However, this is compensated by the fact that other partly treated adolescents weakly benefit from music, as described above.

Table A.16 – *Outcome differences between adolescents with and without music training (Robustness checks, part 1 of 2)*

	OLS		*Caliper matching (r = 0.02)*		*Kernel matching*		*No missing outcomes*		*No missing covariates*		*In sample before treatment*	
	(1)		(2)		(3)		(4)		(5)		(6)	
Cognitive skills (in std. dev.)												
Average cognitive skills	0.08	(0.06)	0.07	(0.07)	0.08	(0.07)	0.14	(0.10)	0.03	(0.09)	-0.11	(0.11)
Analogies	0.14**	(0.07)	0.14^{+}	(0.08)	0.14^{+}	(0.08)	0.17	(0.11)	0.06	(0.09)	-0.01	(0.13)
Figures	0.07	(0.07)	0.09	(0.08)	0.10	(0.08)	0.15	(0.11)	0.10	(0.09)	-0.08	(0.13)
Maths operators	0.00	(0.07)	-0.04	(0.08)	-0.01	(0.08)	0.03	(0.10)	-0.06	(0.09)	-0.14	(0.12)
School grades[1] (in std. dev.)												
Average school grade	-0.15***	(0.05)	-0.16***	(0.06)	-0.16***	(0.06)	-0.19^{+}	(0.11)	-0.09	(0.07)	-0.21**	(0.11)
German grade	-0.14***	(0.05)	-0.15**	(0.06)	-0.15***	(0.06)	-0.23**	(0.11)	-0.09	(0.07)	-0.20^{+}	(0.11)
Foreign language grade	-0.09^{+}	(0.05)	-0.10^{+}	(0.06)	-0.12**	(0.06)	-0.11	(0.11)	-0.09	(0.07)	-0.17	(0.11)
Mathematics grade	-0.11**	(0.06)	-0.12^{+}	(0.06)	-0.10^{+}	(0.06)	-0.11	(0.12)	-0.03	(0.08)	-0.13	(0.11)
Personality (in std. dev.)												
Conscientiousness	0.25***	(0.07)	0.21**	(0.09)	0.20**	(0.09)	0.16	(0.12)	0.22**	(0.10)	0.29**	(0.13)
Openness	0.27***	(0.07)	0.27***	(0.09)	0.26***	(0.09)	0.30**	(0.12)	0.23**	(0.10)	0.26**	(0.13)
Agreeableness	0.13^{+}	(0.07)	0.14^{+}	(0.08)	0.11	(0.08)	0.05	(0.12)	0.09	(0.10)	0.11	(0.14)
Extraversion	-0.03	(0.07)	-0.01	(0.09)	-0.00	(0.08)	-0.08	(0.13)	-0.00	(0.10)	-0.11	(0.15)
Neuroticism	0.06	(0.07)	0.06	(0.09)	0.04	(0.09)	0.02	(0.13)	0.04	(0.10)	0.04	(0.15)
Perceived control	0.04	(0.05)	0.04	(0.06)	0.04	(0.06)	0.15	(0.11)	-0.00	(0.07)	0.08	(0.10)
Time use (in %)												
Watch TV daily	-11.77***	(2.83)	-10.83***	(3.28)	-11.14***	(3.14)	-7.34	(6.10)	-12.39***	(3.78)	-9.81^{+}	(5.52)
Read books daily	7.04**	(2.86)	5.23	(3.35)	6.18^{+}	(3.21)	9.58	(6.21)	2.69	(3.82)	4.96	(5.67)
Ambition (in %)												
Aim Abitur	3.83^{+}	(1.99)	4.09**	(1.97)	5.10**	(1.98)	4.26	(4.29)	4.88**	(2.44)	4.78	(4.35)
Aim university	9.78***	(2.44)	8.44***	(2.81)	9.48***	(2.75)	12.76**	(5.39)	11.38***	(3.21)	10.06^{+}	(5.24)
Job success likely	1.13	(0.99)	1.03	(1.09)	1.01	(1.05)	0.32	(2.10)	1.00	(1.31)	0.55	(1.80)
Desired profession likely	1.58	(1.12)	1.88	(1.22)	1.85	(1.18)	2.01	(2.28)	1.87	(1.53)	1.77	(1.95)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. The table shows the effects of music with different estimation methods, sample restrictions and specifications. (1) OLS estimation. (2) PSM: radius matching with caliper 0.02. (3) PSM: Kernel matching. (4) Complete info outcomes. (5) Complete info covariates. (6) In sample before treatment starts. See Table A.15 for the number of treated and control observations of each treatment definition. Propensity score matching is used to account for control variables (radius matching with caliper 0.01). Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}\ p < 0.1$ $**\ p < 0.05$ $***\ p < 0.01$

Table A.17 – *Outcome differences between adolescents with and without music training (Robustness checks, part 2 of 2)*

	In sample at age 5		*Interactions sibl. x HHinc*		*Father's characteristics*		*Music, but not treated*		*Some music (not never)*		*Give up before age 17*	
	(7)		(8)		(9)		(10)		(11)		(12)	
Cognitive skills (in std. dev.)												
Average cognitive skills	-0.15	(0.13)	0.08	(0.08)	0.08	(0.08)	0.02	(0.06)	-0.03	(0.06)	0.04	(0.27)
Analogies	0.07	(0.15)	0.13	(0.08)	0.12	(0.08)	0.06	(0.07)	-0.01	(0.06)	0.01	(0.26)
Figures	-0.12	(0.14)	0.08	(0.08)	0.11	(0.08)	0.12^{+}	(0.07)	0.02	(0.06)	0.08	(0.27)
Maths operators	-0.25^{+}	(0.15)	-0.01	(0.08)	-0.02	(0.08)	-0.09	(0.07)	-0.06	(0.06)	0.01	(0.26)
School grades[1] *(in std. dev.)*												
Average school grade	-0.24^{+}	(0.13)	-0.16***	(0.06)	-0.18***	(0.06)	-0.06	(0.05)	-0.08^{+}	(0.04)	0.20	(0.26)
German grade	-0.22^{+}	(0.12)	-0.14**	(0.06)	-0.18***	(0.06)	-0.08	(0.06)	-0.13***	(0.04)	-0.12	(0.26)
Foreign language grade	-0.17	(0.12)	-0.12^{+}	(0.06)	-0.13**	(0.06)	-0.01	(0.05)	-0.02	(0.04)	0.31	(0.25)
Mathematics grade	-0.16	(0.13)	-0.11^{+}	(0.07)	-0.12^{+}	(0.07)	-0.04	(0.06)	-0.03	(0.04)	0.20	(0.24)
Personality (in std. dev.)												
Conscientiousness	0.27^{+}	(0.15)	0.21**	(0.09)	0.23**	(0.09)	-0.01	(0.08)	0.12^{+}	(0.06)	0.10	(0.20)
Openness	0.25^{+}	(0.15)	0.25***	(0.09)	0.28***	(0.09)	0.39***	(0.08)	0.34***	(0.06)	0.07	(0.26)
Agreeableness	0.26^{+}	(0.15)	0.12	(0.09)	0.08	(0.08)	0.16**	(0.08)	-0.06	(0.06)	0.10	(0.20)
Extraversion	-0.15	(0.15)	-0.01	(0.09)	-0.05	(0.09)	0.12	(0.08)	0.17***	(0.06)	0.28	(0.23)
Neuroticism	0.06	(0.16)	0.06	(0.09)	-0.03	(0.09)	0.06	(0.08)	0.03	(0.06)	0.31	(0.22)
Perceived control	0.11	(0.11)	0.03	(0.06)	0.04	(0.06)	0.01	(0.05)	0.00	(0.04)	-0.05	(0.22)
Time use (in %)												
Watch TV daily	-10.06^{+}	(5.87)	-10.64***	(3.42)	-11.01***	(3.35)	-2.68	(2.43)	-3.08^{+}	(1.81)	-6.31	(11.26)
Read books daily	5.66	(6.34)	6.64^{+}	(3.42)	7.04**	(3.46)	9.68***	(2.57)	0.63	(1.91)	-16.54^{+}	(9.26)
Ambition (in %)												
Aim Abitur	7.19	(5.17)	3.91^{+}	(2.17)	4.77**	(2.15)	5.42**	(2.16)	2.80	(1.75)	4.30	(12.74)
Aim university	8.32	(6.30)	8.69***	(2.94)	9.27***	(2.84)	8.74***	(2.39)	3.07^{+}	(1.84)	3.00	(12.74)
Job success likely	1.56	(2.06)	1.32	(1.12)	0.94	(1.14)	-0.49	(1.07)	1.32^{+}	(0.77)	5.56	(4.14)
Desired profession likely	3.56	(2.36)	2.16^{+}	(1.27)	2.01	(1.29)	0.69	(1.17)	2.28***	(0.89)	2.47	(4.44)

[1] Note that in Germany, better performance is rewarded with a lower school grade.

Source: SOEP v29 (2001-2012), own calculations. The table shows the effects of music with different estimation methods, sample restrictions and specifications. (7) In sample at age 5. (8) Controls: add interactions nb of siblings and firstborn with HH income. (9) Controls: father's instead of mother's characteristics. (10) Treatment: music, but not original treatment (which is dropped). (11) Treatment: some music (original treatment dropped). (12) Treatment: give up music before 17 (subsample). See Table A.15 for the number of treated and control observations of each treatment definition. Propensity score matching is used to account for control variables (radius matching with caliper 0.01). Standard errors in parentheses are clustered at the household level and estimated by bootstrap (1999 replications). Significance levels: $^{+}$ $p < 0.1$ $**$ $p < 0.05$ $***$ $p < 0.01$

Appendix to “Mozart or Pelé? The effect of teenagers’ participation in music and sports”

B.1 Data

Activity definitions and selection of sample

Table B.1 – *Definitions of musical and athletic activities*

Activity	Definition
Music	- Play music at age 16 or 17 (answer “yes” to Question 16 of youth questionnaire) - Have started to play music at age 14 or earlier (answer “14” or less to Question 19 of youth questionnaire)
Sports (competitively and non-competitively)	- Play sports at age 16 or 17 (answer “yes” to Question 21 of youth questionnaire) - Have started to play sports at age 14 or earlier (answer “14” or less to Question 24 of youth questionnaire)
Sports (only competitively)	- Play sports at age 16 or 17 (answer “yes” to Question 21 of youth questionnaire) - Have started to play sports at age 14 or earlier (answer “14” or less to Question 24 of youth questionnaire) - Regularly take part in sports competitions (answer “yes” to Question 25 of youth questionnaire)

Note: SOEP v29, 2001-2012. Detailed questions of youth questionnaire are described in Table 3.1. The question numbers used here refer to that table.

Table B.2 – *Sample selection*

Step	**Remaining # of observations**
Pooled data of all individuals who have answered the SOEP youth questionnaire when they were 17 (complete youth data contains survey years 2000 to 2012)	4,190
Drop survey year 2000 (music-related questions were not asked)	3,958
Drop individuals with missing information on music or sports participation (missing answers to questions 16, 19, 21, 24 and 25, see Table 3.1)	3,835
Missing information on the outcomes	
Sample with no missing information on cognitive skills: Cognitive skills were measured since 2006, however, some individuals from survey years 2004 and 2005 were also given the opportunity to take the test. Moreover, the test can only be administered among survey participants with a personal interview (test is not possible for adolescents who answered the questionnaire alone for themselves on paper or on a computer). Therefore, for estimations using cognitive skills as outcome measures, the following individuals are dropped: - survey years 2001 to 2003 - individuals who did not have the opportunity to take the test (telephone interview or self-administered interview) - individuals who refused to take the test	1,806
Sample with no missing information on outcomes only measured since 2006: Big Five personality traits, risk aversion, trust and time preferences were only measured since 2006. Therefore, in all estimations using these variables as outcomes, the following individuals were dropped: - individuals who answered youth questionnaire between 2001 and 2005 - individuals who refused to provide answers to the questions	1,742
Sample with no missings on health measures at age 18. Individuals were dropped who were no longer part of the SOEP at age 18 or who refused to provide an answer on any of the outcomes.	2,166
Sample with no missings on standard outcome measures. For all other outcomes (school type attended, class repeated, school grades, opinions about means of success and perceived control), individuals were dropped, who did not provide an answer on at least one of the outcomes.	3,345

Descriptive statistics

Table B.3 – *Covariates by type of activity*

	Not active	One activity: *Sports (competitive and non-competitive)*	One activity: *Sports (competitive only)*	One activity: *Music*	**Both activities**
Female	0.55	0.41	0.35	0.59	0.56
Recommendation upper sec. school	0.25	0.38	0.44	0.55	0.68
Difference: birth year - 1981	6.96	7.61	7.65	8.07	8.44
Birth year 1982 or 1983	0.10	0.06	0.06	0.07	0.04
Birth order	1.68	1.72	1.70	1.65	1.72
Single parent household	0.26	0.22	0.22	0.17	0.17
Number of siblings in SOEP	2.26	2.27	2.28	2.32	2.41
Household lives in rural area	0.28	0.24	0.24	0.29	0.22
North West Germany	0.13	0.15	0.14	0.12	0.16
North Rhine Westphalia	0.20	0.23	0.22	0.17	0.18
Hesse	0.06	0.08	0.08	0.05	0.07
Berlin and Brandenburg	0.09	0.07	0.06	0.08	0.05
Saxony	0.07	0.06	0.06	0.08	0.07
South East Germany	0.09	0.08	0.08	0.09	0.04
City states: Bremen, Hamburg, Berlin	0.05	0.04	0.03	0.06	0.06
HH net overall wealth (100,000 Euros)	1.46	2.05	2.39	2.64	3.02
Log labor market inc. (both parents)	14.03	14.37	14.46	14.59	14.82
Age of mother at birth	26.81	27.40	27.52	28.38	29.19
Hours working (mother)	21.15	22.10	23.18	23.14	21.16
Hours working (father)	37.64	39.48	40.41	42.44	42.91
Mother working in services	0.52	0.58	0.61	0.65	0.63
Father working in services	0.29	0.34	0.35	0.44	0.45
Mother in manufacturing or agriculture	0.11	0.11	0.11	0.12	0.11
Father in manufacturing or agriculture	0.32	0.35	0.36	0.32	0.32
ISEI socio-economic status (highest among parents, scale from 0 to 90)	39.06	44.66	46.41	52.00	53.95
Mother's job requires no training	0.25	0.21	0.19	0.16	0.16
Father's job requires no training	0.18	0.13	0.13	0.08	0.07
Willingness to take risk (mother, 0 to 10)	4.01	4.19	4.29	4.06	4.04
Willingness to take risk (father, 0 to 10)	4.85	5.11	5.15	4.93	4.99
Missing: Willingn. to take risk (mother)	0.09	0.06	0.06	0.08	0.05
Missing: Willingn. to take risk (father)	0.26	0.21	0.21	0.16	0.15
Parent with university degree	0.19	0.31	0.35	0.47	0.51
Parent with vocational degree	0.77	0.75	0.73	0.67	0.68
Parent with upper sec. school degree	0.17	0.30	0.34	0.49	0.52
Parent with migration background	0.24	0.20	0.18	0.14	0.17
Agreeableness (mother, 0 to 1)	0.80	0.81	0.81	0.80	0.80
Conscientiousness (mother, 0 to 1)	0.87	0.87	0.87	0.86	0.85
Extraversion (mother, 0 to 1)	0.72	0.72	0.72	0.72	0.71
Neuroticism (mother, 0 to 1)	0.59	0.59	0.60	0.59	0.58
Openness (mother, 0 to 1)	0.65	0.65	0.65	0.68	0.67
Agreeableness (father, 0 to 1)	0.75	0.75	0.75	0.76	0.77
Conscientiousness (father, 0 to 1)	0.86	0.86	0.86	0.85	0.86
Extraversion (father, 0 to 1)	0.67	0.68	0.68	0.67	0.68
Neuroticism (father, 0 to 1)	0.55	0.54	0.54	0.54	0.53
Openness (father, 0 to 1)	0.61	0.62	0.62	0.65	0.65

Note: SOEP v29, 2001-2012, own calculations. Shares except if otherwise indicated. The exact definitions of each activity are given in Table B.1.

Table B.4 – *Outcomes by type of activity*

	Not active	One activity: *Sports (competitive and non-competitive)*	One activity: *Sports (competitive only)*	One activity: *Music*	Both activities
Cognitive skills and school achievements at age 17					
Cognitive skills - Average	-0.24	-0.00	0.13	0.19	0.36
Cognitive skills - Analogies	-0.23	-0.03	0.08	0.26	0.39
Cognitive skills - Figures	-0.25	0.02	0.07	0.17	0.35
Cognitive skills - Math	-0.13	0.02	0.17	0.08	0.15
School grades - Average	-0.15	-0.02	0.01	0.20	0.32
School grades - Maths	-0.12	0.03	0.08	0.02	0.22
School grades - German	-0.13	-0.03	-0.02	0.29	0.24
School grades - 1st for. lang.	-0.11	-0.03	-0.03	0.18	0.26
Attends upper sec. school	0.23	0.38	0.44	0.57	0.66
Aim to enrol at university	0.17	0.29	0.34	0.51	0.60
Attends university at age 20	0.13	0.20	0.21	0.30	0.36
Personality at age 17					
Big 5: Conscientiousness	-0.09	0.03	0.03	-0.05	0.13
Big 5: Extraversion	-0.11	0.04	0.11	-0.16	0.19
Big 5: Neuroticism	0.03	-0.03	-0.13	0.09	-0.03
Big 5: Openness	-0.15	-0.09	-0.11	0.34	0.34
Big 5: Agreeableness	-0.07	-0.03	-0.00	0.05	0.17
Willingness to take risk	-0.10	0.08	0.16	-0.05	0.00
Subjective health and life style at age 18					
Satisfaction with health	7.92	8.18	8.16	8.01	8.04
Current health situation	2.03	1.88	1.87	2.03	1.94
Currently smoking	0.32	0.26	0.24	0.18	0.14
Other leisure activities at age 17					
Watching TV daily	0.81	0.80	0.79	0.69	0.67
Playing computer games daily	0.26	0.30	0.31	0.20	0.21
Reading daily	0.22	0.22	0.21	0.35	0.39
Aggregate outcome indices (Anderson, 2008)					
Cognitive skills & school grades	0.06	0.02	0.04	-0.04	-0.09
Big Five	-0.09	-0.02	-0.01	0.07	0.19
Non-cognitive skills	-0.10	0.00	0.02	0.04	0.17
Outcomes that should not be affected					
Own room at home	0.86	0.91	0.93	0.94	0.94
Amount of weekly pocket money (Euros)	10.49	10.73	10.36	10.26	9.93
25% or more foreign students in school class	0.34	0.31	0.30	0.21	0.23
Attending church monthly	0.14	0.18	0.18	0.24	0.38

Note: SOEP v29, 2001-2012, own calculations. The exact definitions of each activity are given in Table B.1. The exact measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). Aggregate outcome indices are generated following Anderson (2008) and refer to groups from the standard outcomes in the upper part of the table. All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best"), current health situation (1 "Bad" to 5 "Very good"), and 'amount of weekly pocket money' (in Euros).

Table B.5 – *Characteristics of musical and athletic activities by gender*

	Neither music nor sports		Music only	Sports only (competitive and non-competitive)	Competitive sports only	Music + sports (competitive and non-competitive)	
	Music	*Sports*				*Music*	*Sports*
	(1)		(2)	(3)	(4)	(5)	
Boys							
Average starting age			8.8	9.1	8.6	8.6	9.1
Share (%)...							
...doing sports/music in team	3	12	60	73	86	56	65
...with lessons/instructions	1	21	76	73	93	77	74
...playing sports/music daily	8	9	50	32	39	44	39
Number of adolescents in sample	616		135	973	577	220	
Girls							
Average starting age			8.5	9.6	9.4	8.1	9.2
Share (%)...							
...doing sports/music in team	1	11	48	49	66	48	54
...with lessons/instructions	1	13	81	70	92	80	77
...playing sports/music daily	7	4	33	25	33	34	22
Number of adolescents in sample	758		198	667	307	281	

Note: SOEP v29, 2001-2012, own calculations. Column (4) is a subgroup of column (3), otherwise each column contains distinct individuals, who carry out either nothing (1), music (2), sports (3 with subgroup 4) or both (5). The exact definitions of sports and music participation are given in Table B.1.

B.2 Further estimation results

Results of the estimation of the propensity score

Table B.6 – *Average marginal effects of the estimation of the propensity score*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Marginal effects	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*
Constant	-0.743	0	-0.531	0	-0.729	0	-0.187	40
Female	0.106	0	0.250	0	0.096	0	-0.035	31
Recommendation for upper secondary school	0.044	2	0.089	0	0.148	0	0.141	0
Difference: birth year - 1981	0.007	1	0.006	14	0.006	3	0.000	94
Birth year 1982 or 1983	0.053	22	0.006	91	-0.053	17	-0.138	13
Birth order	-0.033	1	-0.043	2	-0.040	0	0.004	88
Single parent household	-0.007	78	0.005	89	0.018	50	0.021	71
Number of siblings in SOEP	0.023	1	0.030	2	0.032	0	0.011	50
Household lives in rural area	0.039	9	0.058	8	0.011	65	-0.055	23
North West Germany	-0.056	1	-0.105	0	-0.051	4	0.056	30
North Rhine Westphalia	-0.058	1	-0.100	0	-0.092	0	-0.020	72
Hesse	-0.058	4	-0.064	19	-0.060	5	0.041	55
Berlin and Brandenburg	-0.031	45	-0.044	51	-0.095	0	-0.094	37
Saxony	0.070	11	0.133	1	0.074	9	-0.047	51
South East Germany	0.008	82	-0.075	14	-0.089	1	-0.166	5
City states: Bremen, Hamburg, Berlin	0.083	17	0.239	0	0.141	2	-0.003	98
Household net overall wealth (in 100,000 Euros)	0.001	80	-0.002	47	0.001	73	0.000	100
Log of labour market income (both parents)	-0.001	89	0.012	28	0.005	59	0.017	35
Age of mother at birth	0.005	3	0.011	0	0.009	0	0.004	42
Hours working (mother)	-0.001	19	-0.002	5	-0.002	1	-0.001	60
Hours working (father)	0.001	13	0.002	12	0.002	4	0.001	70
Mother working in services	0.073	1	0.056	17	0.056	6	-0.051	38
Father working in services	-0.005	88	-0.028	58	-0.019	60	-0.044	52
Mother working in manufacturing or agriculture	0.120	1	0.138	2	0.077	9	-0.056	48
Father working in manufacturing or agriculture	-0.020	55	-0.072	15	-0.041	25	-0.061	37
ISEI socio-economic status (highest among parents, 0 to 90)	0.001	51	0.000	85	0.001	31	0.001	70

Table B.6 to be continued.

Table B.6 continued.

	Music only vs. sport only (1)		**Music only vs. sport (competitive only)** (2)		**Music + sport vs. sport only** (3)		**Music + sport vs. music only** (4)	
	Marginal effects	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*
Mother's job requires no training	-0.033	15	0.008	83	-0.008	76	0.036	47
Father's job requires no training	-0.014	62	-0.062	13	-0.042	17	-0.021	77
Willingness to take risks (mother, 0 to 10)	-0.008	7	-0.015	3	-0.008	8	0.005	61
Willingness to take risks (father, 0 to 10)	-0.011	1	-0.015	2	-0.011	1	0.002	80
Missing: Willingness to take risks (mother)	0.107	3	0.132	2	0.080	10	-0.046	58
Missing: Willingness to take risks (father)	-0.056	6	-0.118	1	-0.045	18	0.016	82
Parent with university degree	0.004	90	0.012	76	0.048	9	0.035	55
Parent with vocational degree	-0.021	38	-0.003	92	0.035	12	0.061	20
Parent with upper secondary school degree	0.039	14	0.050	18	0.033	21	-0.038	47
Parent with migration background	-0.023	34	0.015	68	0.027	31	0.078	12
Agreeableness (mother, 0 to 1)	-0.162	3	-0.288	1	-0.129	10	0.130	42
Conscientiousness (mother, 0 to 1)	0.010	91	-0.041	76	-0.092	31	-0.243	17
Extraversion (mother, 0 to 1)	-0.081	22	-0.152	12	-0.088	21	-0.041	77
Neuroticism (mother, 0 to 1)	-0.025	66	-0.106	21	-0.042	48	-0.058	61
Openness (mother, 0 to 1)	0.142	3	0.227	2	0.120	8	-0.065	63
Agreeableness (father, 0 to 1)	0.079	32	0.184	12	0.124	11	0.054	73
Conscientiousness (father, 0 to 1)	0.048	57	0.120	37	0.120	18	0.179	37
Extraversion (father, 0 to 1)	-0.072	30	-0.075	46	-0.077	29	0.059	68
Neuroticism (father, 0 to 1)	0.157	1	0.099	28	0.087	17	-0.099	45
Openness (father, 0 to 1)	0.141	5	0.243	2	0.172	2	-0.034	82
Efron's R Square	0.111		0.186		0.160		0.069	
Number of observations	1973		1463		2141		834	

Note: SOEP v29, 2001-2012, own calculations. All covariates are binary except if otherwise indicated. The exact definitions of each activity are given in Table 3.2. Inference based on 4999 bootstrap replications.

Table B.7 – *Average marginal effects of the estimation of the propensity score (comparisons to inactive adolescents)*

	Sports only vs. doing nothing (1)		**Music only vs. doing nothing** (2)		**Music + sport vs. doing nothing** (3)	
	Marginal effects	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*
Constant	-0.162	22	-0.794	0	-0.731	0
Female	-0.140	0	0.052	0	0.030	10
Recommendation for upper secondary school	0.102	0	0.113	0	0.219	0
Difference: birth year - 1981	0.005	9	0.012	0	0.011	0
Birth year 1982 or 1983	-0.103	1	0.010	79	-0.076	3
Birth order	0.018	18	-0.021	14	-0.028	4
Single parent household	-0.011	72	-0.012	69	-0.006	84
Number of siblings in SOEP	0.001	94	0.023	1	0.030	0
Household lives in rural area	-0.040	10	0.026	27	-0.008	76
North West Germany	0.044	13	-0.028	33	-0.013	63
North Rhine Westphalia	0.033	22	-0.055	3	-0.087	0
Hesse	0.064	8	-0.031	41	-0.017	66
Berlin and Brandenburg	-0.005	91	-0.055	14	-0.114	0
Saxony	-0.071	9	0.023	57	0.017	67
South East Germany	-0.023	56	-0.015	66	-0.090	1
City states: Bremen, Hamburg, Berlin	-0.065	20	0.034	51	0.065	19
Household net overall wealth (in 100,000 Euros)	0.003	42	0.002	42	0.002	36
Log labour market income (both parents)	-0.012	55	0.006	76	0.031	12
Age of mother at birth	-0.002	39	0.004	8	0.008	0
Hours working (mother)	-0.001	33	-0.001	18	-0.002	4
Hours working (father)	0.000	85	0.001	46	0.000	73
Mother working in services	0.040	20	0.097	0	0.078	1
Father working in services	-0.004	92	-0.010	77	-0.029	41
Mother working in manufacturing or agriculture	0.037	35	0.146	0	0.094	3
Father working in manufacturing or agriculture	0.029	45	-0.017	64	-0.054	13
ISEI socio-economic status (highest among parents 0 to 90)	0.000	94	0.001	16	0.001	8
Mother's job requires no training	-0.029	24	-0.047	5	-0.022	39
Father's job requires no training	-0.044	12	-0.034	22	-0.056	5
Willingness to take risk (mother, 0 to 10)	0.005	31	-0.004	44	-0.004	45
Willingness to take risk (mother, 0 to 10)	0.010	4	-0.003	53	-0.001	84
Missing: Willingness to take risk (mother)	-0.033	43	0.089	5	0.067	13
Missing: Willingness to take risk (father)	0.003	94	-0.051	13	-0.042	22
Parent with university degree	0.053	9	0.021	51	0.055	9
Parent with vocational degree	0.000	100	-0.033	21	0.033	14
Parent with upper sec. school degree	0.076	1	0.106	0	0.098	0
Parent with migration background	-0.042	11	-0.044	8	0.012	66
Agreeableness (mother, 0 to 1)	0.140	10	-0.029	72	0.030	71
Conscientiousness (mother, 0 to 1)	-0.087	37	-0.069	47	-0.219	2
Extraversion (mother, 0 to 1)	-0.005	95	-0.040	58	-0.024	74
Neuroticism (mother, 0 to 1)	0.054	36	0.035	57	0.007	91
Openness (mother, 0 to 1)	-0.027	69	0.103	13	0.075	26

Table B.7 to be continued.

Table B.7 continued.

	Sports only vs. doing nothing (1)		**Music only vs. doing nothing** (2)		**Music + sport vs. doing nothing** (3)	
	Marginal effects	*p-value in %*	*Marginal effects*	*p-value in %*	*Marginal effects*	*p-value in %*
Agreeableness (father, 0 to 1)	-0.061	48	0.039	67	0.149	8
Conscientiousness (father, 0 to 1)	0.064	52	0.061	53	0.152	13
Extraversion (father, 0 to 1)	0.226	0	0.050	56	0.087	27
Neuroticism (father, 0 to 1)	-0.039	57	0.178	1	0.123	7
Openness (father, 0 to 1)	-0.143	6	0.124	10	0.071	34
Efron's R Square	0.111		0.186		0.160	
Number of observations	1973		1463		2141	

Source: SOEP v29, 2001-2012, own calculations. All covariates are binary, except if otherwise indicated. The exact definitions of each activity are given in Table 3.2. Inference based on 4999 bootstrap replications.

Table B.8 – *Balancing of covariates after matching*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Std. bias in %	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*
Female	0.0	100	0.0	100	0.0	100	0.0	100
Recommendation for upper secondary school	-0.8	80	-2.5	40	-1.7	59	-1.7	59
Difference: birth year - 1981	-2.9	33	3.3	27	0.4	89	0.5	87
Birth year 1982 or 1983	-0.1	98	-1.0	75	-1.0	76	-1.0	76
Birth order	1.3	67	0.8	78	-1.0	74	-1.1	73
Single parent household	-0.2	96	1.1	71	0.9	77	0.9	77
Number of siblings in SOEP	1.5	62	0.5	87	-0.8	80	-0.8	80
Household lives in rural area	-0.6	83	1.7	58	0.9	76	1.0	75
North West Germany	-0.9	77	1.2	70	1.5	63	1.5	63
North Rhine Westphalia	-4.5	13	-1.4	65	-1.8	56	-1.8	56
Hesse	1.2	70	0.8	78	-0.4	91	-0.4	91
Berlin and Brandenburg	0.5	85	-0.7	80	0.0	99	0.0	99
Saxony	0.0	100	1.4	63	0.2	94	0.2	94
South East Germany	-0.1	96	-0.5	87	-2.2	49	-2.2	49
City states: Bremen, Hamburg, Berlin	1.4	65	1.1	72	-0.2	96	-0.2	96
Household net overall wealth (in 100,000 Euros)	1.2	69	0.2	96	-1.0	74	-1.1	73
Log labour market income (both parents)	-2.9	33	-2.0	50	-5.5	8	-5.5	8
Age of mother at birth	1.8	54	1.4	64	1.3	69	1.3	68
Hours working (mother)	-0.7	82	-1.2	69	-2.9	36	-2.8	37
Hours working (father)	0.3	91	0.9	77	-0.4	91	-0.4	90
Mother working in services	-1.0	74	-0.9	76	-0.5	87	-0.4	89
Father working in services	-1.6	59	-0.4	89	-1.8	57	-1.8	56
Mother working in manufacturing or agriculture	0.6	85	0.3	93	-1.8	56	-1.8	56
Father working in manufacturing or agriculture	1.6	60	0.6	83	-0.7	83	-0.6	84
ISEI socio-economic status (highest among parents, 0 to 90)	-1.4	63	-2.3	44	-4.4	16	-4.4	15
Mother's job requires no training	1.5	62	0.1	98	-4.7	13	-4.7	13
Father's job requires no training	-1.2	69	-1.6	59	-1.7	58	-1.7	58
Willingness to take risks (mother, 0 to 10)	-0.7	82	-0.8	78	-4.2	18	-4.1	19

Table B.8 to be continued.

Table B.8 continued.

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Std. bias in %	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*
Willingness to take risks (father, 0 to 10)	1.6	58	2.1	48	-1.6	61	-1.6	60
Missing: Willingness to take risks (mother)	0.7	82	1.6	60	-1.6	62	-1.6	61
Missing: Willingness to take risks (father)	-0.6	83	2.3	43	2.3	46	2.3	46
Parent with university degree	-1.0	73	-0.4	90	-1.3	68	-1.3	68
Parent with vocational degree	1.1	71	-0.8	79	-0.1	97	-0.2	95
Parent with upper secondary school degree	-2.0	51	-1.5	62	-1.9	55	-1.9	55
Parent with migration background	-2.1	49	0.1	97	2.5	43	2.5	43
Agreeableness (mother, 0 to 1)	-1.8	54	-1.8	54	-0.9	78	-0.9	77
Conscientiousness (mother, 0 to 1)	-1.4	64	-2.0	51	0.6	84	0.6	85
Extraversion (mother, 0 to 1)	-2.1	48	-2.3	43	-3.1	32	-3.1	33
Neuroticism (mother, 0 to 1)	0.2	94	1.6	59	0.8	79	0.8	79
Openness (mother, 0 to 1)	-1.5	63	-0.9	75	0.4	89	0.4	89
Agreeableness (father, 0 to 1)	-2.5	41	-1.1	71	-3.7	23	-3.8	23
Conscientiousness (father, 0 to 1)	-1.3	67	-2.4	42	-1.1	72	-1.1	73
Extraversion (father, 0 to 1)	0.5	86	-1.6	59	0.3	93	0.1	97
Neuroticism (father, 0 to 1)	-0.3	93	-1.3	67	1.1	73	1.1	72
Openness (father, 0 to 1)	-0.2	93	-1.1	72	-0.2	95	-0.2	94

Note: SOEP v29, 2001-2012, own calculations. The table indicates the standard bias (including the p-value) between treatment and control group after matching when estimating the Average Treatment Effect on the Treated (ATET). The biases for the estimation of the Average Treatment Effect on the Non-treatet (ATENT) are not significant for any of the covariates either and can be provided by the authors on request. All covariates are binary, except if otherwise indicated. The exact definitions of each activity are given in Table 3.2. Inference based on 4999 bootstrap replications.

Table B.9 – *Balancing of covariates after matching*

	Sports only vs. doing nothing (1)		**Music only vs. doing nothing** (2)		**Music + sport vs. doing nothing** (3)	
	Std. bias in %	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*
Female	0.0	100	0.0	100	0.0	100
Recommendation for upper secondary school	0.0	99	0.2	98	-2.2	75
Difference: birth year - 1981	-1.6	64	0.2	98	-1.7	80
Birth year 1982 or 1983	1.9	59	-0.4	96	2.6	70
Birth order	0.9	80	-2.4	76	-0.9	89
Single parent household	-2.1	54	6.5	42	-0.4	95
Number of siblings in SOEP	2.7	43	-0.5	95	-1.9	77
Household lives in rural area	0.2	95	2.9	71	-3.0	65
North West Germany	4.5	20	9.6	23	-4.2	53
North Rhine Westphalia	-0.8	83	-0.6	94	2.3	73
Hesse	-0.7	83	-2.8	72	-5.8	39
Berlin and Brandenburg	-1.0	77	-13.4	9	-4.4	51
Saxony	-2.5	48	-11.1	16	5.5	42
South East Germany	-0.6	85	10.2	20	-2.0	77
City states: Bremen, Hamburg, Berlin	0.7	84	-4.1	61	-2.4	72
Household net overall wealth (in 100,000 Euros)	6.4	7	-3.6	65	-10.0	14
Log labour market income (both parents)	-1.8	60	7.2	37	-2.6	70
Age of mother at birth	-1.0	78	-0.3	97	-1.9	78
Hours working (mother)	-2.0	56	-4.6	56	-2.7	69
Hours working (father)	-2.1	55	2.0	80	0.5	94
Mother working in services	-1.7	62	-1.4	86	-0.8	90
Father working in services	0.0	99	2.1	79	3.7	59
Mother working in manufacturing or agriculture	0.4	92	-0.6	94	0.9	90
Father working in manufacturing or agriculture	0.8	81	-9.2	25	-2.9	66
ISEI socio-economic status (highest among parents, 0 to 90)	-1.0	77	2.0	80	5.1	45
Mother's job requires no training	0.5	89	7.8	33	-1.6	81
Father's job requires no training	1.0	78	-2.6	74	5.7	40
Willingness to take risks (mother, 0 to 10)	-2.1	55	7.3	36	0.0	99
Willingness to take risks (father, 0 to 10)	0.8	82	5.4	50	-1.0	88
Missing: Willingness to take risks (mother)	1.6	64	4.3	58	-0.1	98
Missing: Willingness to take risks (father)	-0.6	87	9.0	26	0.0	100
Parent with university degree	-2.8	42	0.8	92	3.5	60
Parent with vocational degree	2.8	42	3.7	64	-6.8	31
Parent with upper secondary school degree	0.5	89	1.6	84	3.2	63
Parent with migration background	0.4	90	-7.4	35	-2.3	73
Agreeableness (mother, 0 to 1)	2.2	53	0.9	91	-4.7	49
Conscientiousness (mother, 0 to 1)	-2.9	41	-0.1	99	-5.4	42
Extraversion (mother, 0 to 1)	-0.2	95	0.8	92	-5.4	42
Neuroticism (mother, 0 to 1)	-3.1	37	-4.5	57	9.4	16
Openness (mother, 0 to 1)	-1.9	58	5.6	48	-3.1	64
Agreeableness (father, 0 to 1)	-1.7	62	9.6	23	6.0	38
Conscientiousness (father, 0 to 1)	-1.3	70	7.3	36	-3.1	64

Table B.9 to be continued.

Table B.9 continued.

	Sports only vs. doing nothing (1)		**Music only vs. doing nothing** (2)		**Music + sport vs. doing nothing** (3)	
	Std. bias in %	*p-value in %*	*Std. bias in %*	*p-value in %*	*Std. bias in %*	*p-value in %*
Extraversion (father, 0 to 1)	-2.0	56	-2.2	78	4.7	49
Neuroticism (father, 0 to 1)	1.5	66	-8.8	27	-6.3	35
Openness (father, 0 to 1)	-1.5	67	2.3	77	9.7	15

Note: SOEP v29, 2001-2012, own calculations. The table indicates the standard bias (including the p-value) between treatment and control group after matching when estimating the Average Treatment Effect on the Treated (ATET). The biases for the estimation of the Average Treatment Effect on the Non-treatet (ATENT) are not significant for any of the covariates either and can be provided by the authors on request. All covariates are binary, except if otherwise indicated. The exact definitions of each activity are given in Table 3.2. Inference based on 4999 bootstrap replications.

Full set of effects

Table B.10 – *Average effects of music vs. sports*

	Music only vs. sport only (1)		Music only vs. sport (competitive only) (2)		Music + sport vs. sport only (3)		Music + sport vs. music only (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.012	90	0.073	48	0.233	1	0.230	5
Cognitive skills - Analogies	0.054	58	0.143	17	0.216	1	0.169	15
Cognitive skills - Figures	0.053	59	0.108	26	0.276	0	0.229	5
Cognitive skills - Math	-0.045	63	-0.065	51	0.070	45	0.123	30
School grades - Average	0.034	66	0.091	19	0.151	3	0.123	20
School grades - Maths	-0.119	14	-0.075	30	0.113	12	0.233	2
School grades - German	0.172	2	0.196	0	0.113	11	-0.051	59
School grades - 1st for. lang.	0.063	42	0.130	6	0.103	13	0.046	63
Attends upper sec. school	0.081	4	0.083	2	0.135	0	0.063	18
Aim to enrol at university	0.106	1	0.115	0	0.187	0	0.087	6
Attends university at age 20	0.024	48	0.085	2	0.084	1	0.065	13
Personality at age 17								
Big 5: Conscientiousness	-0.114	21	-0.053	60	0.059	51	0.165	14
Big 5: Extraversion	-0.165	9	-0.171	8	0.128	13	0.294	1
Big 5: Neuroticism	0.103	31	0.104	26	0.013	88	-0.090	47
Big 5: Openness	0.407	0	0.348	0	0.349	0	-0.058	63
Big 5: Agreeableness	0.048	60	0.032	73	0.206	2	0.152	17
Willingness to take risk	-0.211	4	-0.239	1	-0.181	6	0.021	87
Subjective health and life style at age 18								
Satisfaction with health	-0.119	45	-0.341	2	-0.143	28	-0.050	80
Current health situation	-0.138	4	-0.195	0	-0.048	39	0.083	30
Currently smoking	-0.007	88	0.008	83	-0.110	0	-0.107	5
Other leisure activities at age 17								
Watching TV daily	-0.099	0	-0.080	1	-0.115	0	-0.016	71
Playing computer games daily	-0.075	2	-0.099	0	-0.044	14	0.029	47
Reading daily	0.067	5	0.134	0	0.127	0	0.062	14
Aggregated outcome indices (Anderson, 2008)								
Cognitive skills & school grades	-0.020	75	0.011	87	-0.084	16	-0.016	83
Big Five	0.050	39	0.086	11	0.148	0	0.130	6
Non-cognitive skills	0.015	76	0.041	42	0.107	1	0.116	5
Outcomes that should not be affected								
Own room at home	0.041	5	0.025	21	0.017	36	-0.014	55
Weekly pocket money (Euros)	-0.330	73	-0.278	85	-0.793	48	0.782	66
25% or more foreign students in school class	-0.048	20	-0.039	30	0.004	91	0.022	63
Attending church monthly	0.007	88	0.065	8	0.126	0	0.144	0

Note: Effects presented are average treatment effects for the target population. Inference is based on 4999 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). Aggregate outcome indices are generated following Anderson (2008) and refer to groups from the standard outcomes in the upper part of the table. All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”), current health situation (1 “Bad” to 5 “Very good”), and ‘weekly pocket money’ (in Euros).

Heterogeneity of the effects

Table B.11 – *Average effects of music vs. sports (female adolescents)*

	Music only vs. sport only		Music only vs. sport (competitive only)		Music + sport vs. sport only		Music + sport vs. music only	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.250	5	-0.018	89	-0.042	73	-0.282	5
Cognitive skills - Analogies	0.267	3	0.150	27	0.076	48	-0.184	18
Cognitive skills - Figures	0.210	14	-0.067	63	-0.021	86	-0.220	15
Cognitive skills - Math	0.139	26	-0.056	68	-0.120	39	-0.255	9
School grades - Average	0.103	28	-0.145	42	0.087	27	-0.011	91
School grades - Maths	0.003	96	-0.233	10	-0.028	76	-0.029	84
School grades - German	0.248	2	0.043	62	0.102	24	-0.139	22
School grades - 1st for. lang.	0.017	84	-0.110	59	0.108	19	0.094	41
Attends upper sec. school	0.073	12	0.014	75	0.014	76	-0.055	41
Aim to enrol at university	0.054	26	0.038	41	0.057	18	0.006	89
Attends university at age 20	0.036	50	0.027	55	0.041	41	0.007	92
Personality at age 17								
Big 5: Conscientiousness	0.001	96	-0.162	28	0.110	29	0.108	49
Big 5: Extraversion	-0.049	70	-0.053	76	0.029	79	0.081	61
Big 5: Neuroticism	-0.024	86	0.193	25	0.124	46	0.150	43
Big 5: Openness	0.468	0	0.251	8	0.105	44	-0.359	13
Big 5: Agreeableness	0.093	52	-0.014	90	0.188	13	0.092	56
Willingness to take risk	-0.004	96	-0.143	27	-0.078	47	-0.071	61
Subjective health and life style at age 18								
Satisfaction with health	-0.070	75	-0.178	31	-0.204	29	-0.130	63
Current health situation	-0.047	71	-0.073	37	-0.115	17	-0.061	69
Currently smoking	-0.066	14	0.001	96	-0.023	63	0.042	50
Other leisure activities at age 17								
Watching TV daily	-0.065	13	-0.026	52	-0.009	75	0.055	25
Playing computer games daily	-0.028	32	-0.009	74	-0.027	44	0.001	96
Reading daily	0.149	0	0.132	0	0.132	0	-0.016	78

Note: Effects presented are average treatment effects for the target population. Inference is based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Table B.12 – *Average effects of music vs. sports (male adolescents)*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	-0.246	15	-0.001	97	0.216	42	0.467	20
Cognitive skills - Analogies	-0.119	57	0.141	33	0.315	21	0.436	24
Cognitive skills - Figures	-0.204	28	-0.007	95	0.055	82	0.265	40
Cognitive skills - Math	-0.235	8	-0.125	54	0.043	87	0.282	35
School grades - Average	-0.109	41	-0.017	89	0.208	19	0.319	14
School grades - Maths	-0.234	8	-0.117	27	0.123	31	0.357	6
School grades - German	0.146	25	0.066	61	0.144	43	0.001	97
School grades - 1st for. lang.	-0.205	36	0.015	90	0.191	22	0.397	17
Attends upper sec. school	0.048	55	0.012	85	0.138	2	0.093	36
Aim to enrol at university	0.122	7	0.121	1	0.172	1	0.051	59
Attends university at age 20	0.048	38	0.081	8	0.186	5	0.138	15
Personality at age 17								
Big 5: Conscientiousness	-0.167	29	-0.233	26	0.068	67	0.235	30
Big 5: Extraversion	-0.310	7	-0.283	4	0.108	53	0.421	4
Big 5: Neuroticism	0.138	48	0.106	41	0.056	72	-0.084	72
Big 5: Openness	0.480	0	0.482	0	0.481	0	0.003	97
Big 5: Agreeableness	0.072	65	-0.076	61	0.066	56	-0.005	95
Willingness to take risk	-0.304	13	-0.171	28	-0.174	54	0.135	58
Subjective health and life style at age 18								
Satisfaction with health	0.139	46	-0.015	90	-0.291	51	-0.428	29
Current health situation	-0.073	42	-0.055	46	-0.102	45	-0.029	88
Currently smoking	-0.073	55	-0.008	93	-0.106	7	-0.033	80
Other leisure activities at age 17								
Watching TV daily	0.003	95	-0.044	31	-0.162	1	-0.166	5
Playing computer games daily	0.050	76	0.002	95	-0.083	17	-0.133	39
Reading daily	0.045	40	0.050	42	0.101	9	0.056	47

Note: Effects presented are average treatment effects for the target population. Inference is based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Table B.13 – *Average effects of music vs. sports (adolescents without recommendation for upper secondary school at the end of primary school)*

	Music only vs. sport only (1)		Music only vs. sport (competitive only) (2)		Music + sport vs. sport only (3)		Music + sport vs. music only (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.120	57	-0.032	89	-0.170	43	-0.292	31
Cognitive skills - Analogies	0.137	51	0.186	41	-0.021	89	-0.161	53
Cognitive skills - Figures	0.058	74	0.022	92	-0.258	36	-0.320	28
Cognitive skills - Math	0.100	59	-0.235	9	-0.172	46	-0.271	39
School grades - Average	-0.051	65	-0.005	95	0.035	78	0.093	53
School grades - Maths	-0.194	13	-0.235	3	-0.131	32	0.068	66
School grades - German	0.180	8	0.156	12	0.038	74	-0.138	29
School grades - 1st for. lang.	-0.076	59	0.088	44	0.139	26	0.218	22
Attends upper sec. school	0.032	49	0.017	67	0.045	35	0.014	81
Aim to enrol at university	0.012	80	0.023	54	0.081	9	0.069	27
Attends university at age 20	0.015	76	0.034	42	0.056	29	0.044	45
Personality at age 17								
Big 5: Conscientiousness	-0.222	27	-0.407	24	-0.059	65	0.141	52
Big 5: Extraversion	-0.230	24	-0.405	1	0.100	41	0.340	15
Big 5: Neuroticism	0.160	50	0.028	86	-0.187	26	-0.367	21
Big 5: Openness	0.191	22	0.156	34	0.235	4	0.019	91
Big 5: Agreeableness	0.049	82	0.033	85	-0.069	70	-0.127	57
Willingness to take risk	-0.091	67	-0.092	54	-0.080	66	0.009	96
Subjective health and life style at age 18								
Satisfaction with health	-0.288	40	-0.210	28	-0.261	44	0.019	94
Current health situation	-0.121	20	-0.183	6	-0.078	43	0.041	76
Currently smoking	-0.022	81	-0.042	49	-0.147	3	-0.126	25
Other leisure activities at age 17								
Watching TV daily	0.017	69	-0.005	89	-0.026	62	-0.044	40
Playing computer games daily	0.054	34	0.031	53	-0.007	90	-0.063	47
Reading daily	0.025	58	0.126	0	0.068	23	0.044	54

Note: Effects presented are average treatment effects for target population. Inference based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.14 – *Average effects of music vs. sports (adolescents with recommendation for upper secondary school at the end of primary school)*

	Music only vs. sport only (1)		**Music only vs. sport (competitive only)** (2)		**Music + sport vs. sport only** (3)		**Music + sport vs. music only** (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.095	48	0.002	95	0.186	4	0.085	46
Cognitive skills - Analogies	0.209	12	0.173	26	0.183	9	-0.030	85
Cognitive skills - Figures	0.171	17	0.000	97	0.190	2	0.010	94
Cognitive skills - Math	-0.094	57	-0.124	39	0.092	31	0.183	27
School grades - Average	-0.094	70	0.049	64	0.011	92	0.099	49
School grades - Maths	-0.195	35	0.111	50	0.078	52	0.270	6
School grades - German	-0.034	87	0.008	92	-0.035	67	-0.007	96
School grades - 1st for. lang.	0.030	74	-0.028	85	-0.029	72	-0.064	54
Attends upper sec. school	0.092	1	0.039	34	0.099	0	0.009	78
Aim to enrol at university	0.160	0	0.099	6	0.143	0	-0.017	71
Attends university at age 20	0.022	72	-0.006	96	0.058	30	0.037	55
Personality at age 17								
Big 5: Conscientiousness	-0.095	52	-0.181	58	0.081	39	0.181	34
Big 5: Extraversion	-0.109	56	-0.143	31	0.224	7	0.318	3
Big 5: Neuroticism	0.125	31	0.247	9	0.053	59	-0.078	53
Big 5: Openness	0.548	0	0.480	0	0.429	0	-0.119	35
Big 5: Agreeableness	-0.042	74	-0.140	33	0.215	6	0.259	7
Willingness to take risk	-0.158	24	-0.201	13	-0.020	86	0.135	33
Subjective health and life style at age 18								
Satisfaction with health	0.218	27	-0.053	77	0.071	80	-0.150	41
Current health situation	-0.143	14	-0.055	48	-0.006	95	0.135	39
Currently smoking	-0.028	71	0.055	30	-0.055	24	-0.027	81
Other leisure activities at age 17								
Watching TV daily	-0.075	15	-0.095	3	-0.069	7	0.005	92
Playing computer games daily	-0.092	4	-0.070	10	-0.058	7	0.033	47
Reading daily	0.109	4	0.102	7	0.118	0	0.006	91

Note: Effects presented are average treatment effects for the target population. Inference based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Table B.15 – *Average effects of music vs. sports (adolescents with parents who did not receive an upper secondary school degree)*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.099	43	0.050	68	-0.062	79	-0.160	47
Cognitive skills - Analogies	0.087	52	0.101	46	0.030	85	-0.054	77
Cognitive skills - Figures	0.134	26	0.021	87	-0.070	78	-0.203	33
Cognitive skills - Math	0.039	78	-0.019	89	-0.203	36	-0.242	30
School grades - Average	0.034	78	-0.035	75	-0.057	60	-0.091	52
School grades - Maths	-0.051	67	-0.103	30	0.012	91	0.070	69
School grades - German	0.075	55	0.007	94	-0.145	34	-0.221	11
School grades - 1st for. lang.	0.044	76	-0.055	60	-0.047	68	-0.099	56
Attends upper sec. school	0.067	25	0.017	71	0.138	3	0.075	53
Aim to enrol at university	0.065	23	0.041	33	0.065	25	0.006	92
Attends university at age 20	-0.017	73	0.053	13	0.051	44	0.069	27
Personality at age 17								
Big 5: Conscientiousness	-0.123	32	-0.028	86	-0.097	59	0.024	91
Big 5: Extraversion	-0.260	5	-0.111	40	0.277	4	0.535	0
Big 5: Neuroticism	0.235	12	0.228	7	-0.092	49	-0.327	11
Big 5: Openness	0.337	2	0.419	0	0.207	18	-0.130	51
Big 5: Agreeableness	0.172	27	0.115	35	0.085	39	-0.083	64
Willingness to take risk	-0.355	4	-0.103	52	0.127	22	0.481	3
Subjective health and life style at age 18								
Satisfaction with health	0.115	56	-0.176	29	-0.430	37	-0.553	30
Current health situation	-0.128	9	-0.151	9	-0.040	66	0.089	41
Currently smoking	0.033	84	0.008	90	-0.106	13	-0.139	41
Other leisure activities at age 17								
Watching TV daily	-0.060	25	-0.056	17	-0.022	68	0.036	60
Playing computer games daily	-0.051	41	-0.024	58	0.011	92	0.065	51
Reading daily	0.078	18	0.102	3	0.060	24	-0.022	77

Note: Effects presented are average treatment effects for the target population. P-values are based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.16 – *Average effects of music vs. sports (adolescents with parents who have an upper secondary school degree)*

	Music only vs. sport only (1)		**Music only vs. sport (competitive only)** (2)		**Music + sport vs. sport only** (3)		**Music + sport vs. music only** (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	-0.103	45	-0.146	34	0.136	31	0.235	14
Cognitive skills - Analogies	-0.018	86	-0.004	95	0.093	56	0.107	50
Cognitive skills - Figures	-0.156	25	-0.169	24	0.073	53	0.233	15
Cognitive skills - Math	-0.068	65	-0.148	37	0.146	39	0.207	20
School grades - Average	-0.055	53	0.029	77	0.026	89	0.079	59
School grades - Maths	-0.194	5	-0.098	33	-0.036	82	0.156	33
School grades - German	0.049	69	0.029	81	-0.013	95	-0.063	61
School grades - 1st for. lang.	0.051	57	0.149	13	0.126	26	0.072	59
Attends upper sec. school	0.074	8	-0.018	72	0.029	55	-0.045	51
Aim to enrol at university	0.113	1	0.078	13	0.078	16	-0.036	58
Attends university at age 20	0.080	21	-0.033	79	0.036	56	-0.042	63
Personality at age 17								
Big 5: Conscientiousness	-0.069	63	0.066	65	0.016	89	0.081	63
Big 5: Extraversion	-0.254	18	-0.098	55	0.044	80	0.294	16
Big 5: Neuroticism	0.035	86	-0.068	73	0.013	93	-0.014	93
Big 5: Openness	0.322	7	0.334	0	0.350	0	0.027	90
Big 5: Agreeableness	-0.073	60	-0.072	53	0.169	41	0.236	25
Willingness to take risk	-0.089	55	-0.063	84	-0.276	6	-0.185	28
Subjective health and life style at age 18								
Satisfaction with health	-0.253	27	-0.004	97	0.058	77	0.312	27
Current health situation	-0.196	5	-0.035	75	-0.129	11	0.070	52
Currently smoking	-0.086	10	-0.023	61	-0.094	10	-0.008	89
Other leisure activities at age 17								
Watching TV daily	-0.014	83	-0.020	72	-0.028	71	-0.015	78
Playing computer games daily	-0.094	3	-0.014	75	0.003	94	0.096	8
Reading daily	0.094	8	0.123	1	0.117	1	0.023	72

Note: Effects presented are average treatment effects for the target population. Inference based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Table B.17 – *Average effects of music vs. sports (adolescents with parents having an average labour market income below 1500 euros per month)*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.063	63	0.003	96	-0.243	46	-0.301	32
Cognitive skills - Analogies	0.192	23	0.065	62	-0.019	90	-0.204	37
Cognitive skills - Figures	0.049	67	0.026	87	-0.201	43	-0.245	33
Cognitive skills - Math	-0.048	71	-0.078	68	-0.364	33	-0.315	28
School grades - Average	-0.013	91	-0.014	91	0.127	35	0.145	47
School grades - Maths	-0.121	32	-0.072	45	0.035	73	0.157	30
School grades - German	0.146	15	0.093	29	0.083	52	-0.056	69
School grades - 1st for. lang.	-0.058	63	-0.054	76	0.146	44	0.208	39
Attends upper sec. school	0.105	3	0.051	23	0.098	2	-0.006	92
Aim to enrol at university	0.092	5	0.094	2	0.119	1	0.028	68
Attends university at age 20	-0.003	92	0.078	11	0.037	55	0.041	40
Personality at age 17								
Big 5: Conscientiousness	-0.186	22	-0.085	55	0.077	72	0.262	28
Big 5: Extraversion	-0.231	12	-0.166	19	0.063	61	0.292	5
Big 5: Neuroticism	0.326	2	0.314	8	-0.118	33	-0.444	1
Big 5: Openness	0.273	14	0.282	3	0.196	6	-0.080	71
Big 5: Agreeableness	0.239	5	0.082	59	0.183	23	-0.059	77
Willingness to take risk	-0.332	7	-0.275	12	0.016	93	0.349	6
Subjective health and life style at age 18								
Satisfaction with health	-0.272	22	-0.404	13	0.305	34	0.573	17
Current health situation	-0.192	2	-0.131	12	0.093	38	0.284	4
Currently smoking	-0.084	11	-0.018	74	-0.146	0	-0.062	31
Other leisure activities at age 17								
Watching TV daily	-0.100	3	-0.044	37	-0.053	30	0.047	48
Playing computer games daily	-0.003	94	0.032	52	-0.032	49	-0.030	68
Reading daily	0.130	0	0.110	2	0.181	1	0.053	49

Note: Effects presented are average treatment effects for the target population. Inference based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Table B.18 – *Average effects of music vs. sports (adolescents with parents having an average labour market income above 1500 euros per month)*

	Music only vs. sport only (1)		Music only vs. sport (competitive only) (2)		Music + sport vs. sport only (3)		Music + sport vs. music only (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.031	88	0.006	96	0.328	1	0.297	15
Cognitive skills - Analogies	0.076	76	0.079	70	0.314	2	0.237	40
Cognitive skills - Figures	0.134	67	0.050	73	0.305	1	0.171	50
Cognitive skills - Math	-0.090	64	-0.069	63	0.182	15	0.274	19
School grades - Average	-0.107	71	-0.057	84	0.117	30	0.225	26
School grades - Maths	-0.256	17	-0.259	20	0.012	93	0.269	12
School grades - German	0.147	30	0.094	48	0.149	15	0.002	98
School grades - 1st for. lang.	-0.090	64	0.052	76	0.120	35	0.211	16
Attends upper sec. school	0.013	83	-0.020	69	0.041	51	0.029	66
Aim to enrol at university	0.061	35	0.059	27	0.102	10	0.041	56
Attends university at age 20	0.072	28	0.140	5	0.047	36	-0.024	75
Personality at age 17								
Big 5: Conscientiousness	-0.293	31	0.029	83	0.183	14	0.475	7
Big 5: Extraversion	-0.063	76	-0.180	25	0.239	11	0.304	9
Big 5: Neuroticism	-0.006	96	-0.019	93	0.143	32	0.150	39
Big 5: Openness	0.233	22	0.478	1	0.454	0	0.223	37
Big 5: Agreeableness	-0.297	10	-0.130	38	0.189	18	0.487	1
Willingness to take risk	-0.122	46	-0.250	11	-0.521	2	-0.399	6
Subjective health and life style at age 18								
Satisfaction with health	0.094	69	-0.064	78	-0.207	28	-0.303	25
Current health situation	-0.094	25	-0.044	72	-0.115	15	-0.022	84
Currently smoking	-0.033	61	0.013	82	0.059	71	0.092	34
Other leisure activities at age 17								
Watching TV daily	-0.030	64	-0.042	38	-0.019	78	0.011	88
Playing computer games daily	-0.063	23	-0.067	12	-0.016	75	0.047	49
Reading daily	0.076	15	0.158	1	0.115	0	0.040	52

Note: Effects presented are average treatment effects for the target population. Inference based on 499 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.19 – *Average effects of music vs. sports (target population: all adolescents)*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.058	54	0.021	84	0.181	6	0,123	27
Cognitive skills - Analogies	0.171	11	0.109	22	0.161	16	-0.010	95
Cognitive skills - Figures	-0.001	99	0.045	64	0.172	8	0.173	18
Cognitive skills - Math	-0.008	93	-0.076	48	0.069	51	0.077	56
School grades - Average	0.070	37	0.055	45	0.172	3	0.101	30
School grades - Maths	-0.159	13	-0.078	30	0.101	16	0.260	4
School grades - German	0.236	1	0.142	4	0.112	11	-0.124	20
School grades - 1st for. lang.	0.135	19	0.088	18	0.124	18	-0.011	93
Attends upper sec. school	0.065	11	0.082	7	0.123	0	0.058	24
Aim to enrol at university	0.079	6	0.109	1	0.157	0	0.078	14
Attends university at age 20	0.029	44	0.097	3	0.100	1	0.071	12
Personality at age 17								
Big 5: Conscientiousness	-0.212	15	-0.109	39	0.054	56	0.266	10
Big 5: Extraversion	-0.099	38	-0.238	5	0.208	4	0.307	1
Big 5: Neuroticism	0.097	38	0.150	11	-0.075	59	-0.172	24
Big 5: Openness	0.370	0	0.219	20	0.333	0	-0.037	77
Big 5: Agreeableness	0.025	84	0.085	37	0.075	70	0.050	76
Willingness to take risk	-0.113	33	-0.234	4	-0.028	79	0.084	56
Subjective health and life style at age 18								
Satisfaction with health	-0.457	3	-0.240	23	-0.203	25	0.255	33
Current health situation	-0.081	42	-0.168	4	-0.061	45	0.020	87
Currently smoking	-0.018	72	-0.005	91	-0.036	41	-0.017	78
Other leisure activities at age 17								
Watching TV daily	-0.064	6	-0.066	3	-0.077	2	-0.013	78
Playing computer games daily	-0.068	5	-0.070	2	-0.059	2	0.008	85
Reading daily	0.106	0	0.174	0	0.134	0	0.027	56

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for all individuals, irrespective of whether they are active or not. Inference is based on 4999 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

Robustness analysis

Table B.20 – *Average effects of music vs. sports (additional requirement for being considered to play music: take music lessons outside of school)*

	Music only vs. sport only (1)		Music only vs. sport (competitive only) (2)		Music + sport vs. sport only (3)		Music + sport vs. music only (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.122	32	0.156	22	0.262	2	0.162	26
Cognitive skills - Analogies	0.186	15	0.226	2	0.256	2	0.089	55
Cognitive skills - Figures	0.159	18	0.189	9	0.339	0	0.198	17
Cognitive skills - Math	-0.009	94	0.001	98	0.073	52	0.100	47
School grades - Average	0.063	47	0.116	11	0.239	0	0.192	6
School grades - Maths	-0.113	23	-0.008	91	0.254	0	0.371	0
School grades - German	0.229	0	0.195	1	0.123	7	-0.086	39
School grades - 1st for. lang.	0.060	49	0.100	15	0.140	9	0.094	41
Attends upper sec. school	0.140	0	0.097	1	0.173	0	0.052	29
Aim to enrol at university	0.169	0	0.139	0	0.227	0	0.071	16
Attends university at age 20	0.073	10	0.075	8	0.122	0	0.056	29
Personality at age 17								
Big 5: Conscientiousness	-0.079	48	0.044	59	0.105	24	0.176	17
Big 5: Extraversion	-0.245	2	-0.223	4	0.057	53	0.300	2
Big 5: Neuroticism	0.081	51	0.267	2	0.081	32	0.001	98
Big 5: Openness	0.392	0	0.401	0	0.392	0	-0.001	97
Big 5: Agreeableness	0.119	28	0.125	14	0.264	0	0.142	33
Willingness to take risk	-0.181	11	-0.303	0	-0.109	25	0.063	68
Subjective health and life style at age 18								
Satisfaction with health	-0.243	14	-0.211	15	-0.091	61	0.132	59
Current health situation	-0.131	10	-0.119	8	0.032	61	0.161	8
Currently smoking	-0.071	12	-0.038	31	-0.102	0	-0.041	45
Other leisure activities at age 17								
Watching TV daily	-0.096	2	-0.111	1	-0.139	0	-0.045	37
Playing computer games daily	-0.120	0	-0.101	0	-0.074	3	0.040	59
Reading daily	0.102	2	0.126	0	0.120	0	0.024	65

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for the target population. Inference based on 999 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”)

Table B.21 – *Average effects of music vs. sports (in addition to standard definition, adolescents are considered as active if they play music or sports at least monthly)*

	Music only vs. sport only (1)		**Music only vs. sport (competitive only)** (2)		**Music + sport vs. sport only** (3)		**Music + sport vs. music only** (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.082	41	-0.050	60	0.209	2	0.128	26
Cognitive skills - Analogies	0.182	15	0.098	26	0.255	0	0.076	56
Cognitive skills - Figures	-0.003	97	0.013	90	0.326	0	0.329	1
Cognitive skills - Math	0.030	76	-0.171	6	-0.042	70	-0.073	60
School grades - Average	0.038	57	0.159	6	0.172	1	0.137	10
School grades - Maths	-0.110	14	-0.014	84	0.141	5	0.248	1
School grades - German	0.154	2	0.218	0	0.127	3	-0.021	79
School grades - 1st for. lang.	0.079	31	0.180	3	0.119	4	0.046	60
Attends upper sec. school	0.100	1	0.055	12	0.146	0	0.049	27
Aim to enrol at university	0.115	0	0.105	0	0.171	0	0.059	18
Attends university at age 20	0.056	16	0.077	2	0.073	5	0.020	71
Personality at age 17								
Big 5: Conscientiousness	-0.169	18	-0.003	98	0.091	47	0.259	4
Big 5: Extraversion	-0.223	4	-0.239	1	0.194	4	0.417	0
Big 5: Neuroticism	0.087	40	0.056	58	-0.098	31	-0.186	18
Big 5: Openness	0.495	0	0.357	0	0.443	0	-0.052	66
Big 5: Agreeableness	-0.044	69	0.096	29	0.144	15	0.188	12
Willingness to take risk	-0.100	35	-0.243	4	0.019	87	0.117	41
Subjective health and life style at age 18								
Satisfaction with health	-0.002	98	-0.238	9	-0.002	97	0.001	98
Current health situation	-0.065	58	-0.150	2	-0.015	82	0.049	65
Currently smoking	-0.115	0	-0.057	10	-0.066	16	0.047	52
Other leisure activities at age 17								
Watching TV daily	-0.081	3	-0.094	0	-0.074	5	0.006	90
Playing computer games daily	-0.057	12	-0.096	0	-0.049	13	0.007	87
Reading daily	0.105	2	0.155	0	0.131	0	0.025	58

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for the target population. Inference based on 999 bootstrap replications. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.22 – *Average effects of music vs. sports (using a different common support rule)*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.100	28	0.103	27	0.302	0	0.213	6
Cognitive skills - Analogies	0.080	58	0.198	5	0.269	0	0.201	10
Cognitive skills - Figures	0.161	10	0.167	5	0.345	0	0.191	10
Cognitive skills - Math	0.029	74	-0.065	48	0.126	15	0.107	32
School grades - Average	0.058	44	0.124	8	0.193	0	0.152	8
School grades - Maths	-0.086	26	-0.025	72	0.175	2	0.264	1
School grades - German	0.171	2	0.204	0	0.137	4	-0.012	89
School grades - 1st for. lang.	0.068	36	0.132	6	0.106	26	0.055	52
Attends upper sec. school	0.081	6	0.097	1	0.190	0	0.123	3
Aim to enrol at university	0.126	0	0.118	0	0.218	0	0.103	4
Attends university at age 20	0.056	15	0.073	3	0.128	0	0.078	9
Personality at age 17								
Big 5: Conscientiousness	-0.054	53	-0.047	63	0.106	23	0.168	9
Big 5: Extraversion	-0.121	27	-0.173	7	0.114	17	0.250	5
Big 5: Neuroticism	0.014	92	0.139	14	0.032	70	0.003	97
Big 5: Openness	0.411	0	0.377	0	0.413	0	-0.003	96
Big 5: Agreeableness	0.022	82	0.080	38	0.327	0	0.319	2
Willingness to take risk	-0.114	28	-0.236	1	-0.151	25	-0.026	84
Subjective health and life style at age 18								
Satisfaction with health	-0.134	37	-0.421	0	-0.015	89	0.128	43
Current health situation	-0.119	9	-0.175	0	-0.039	51	0.084	26
Currently smoking	-0.040	43	-0.022	53	-0.104	0	-0.075	25
Other leisure activities at age 17								
Watching TV daily	-0.097	2	-0.096	0	-0.122	0	-0.028	49
Playing computer games daily	-0.072	1	-0.088	0	-0.064	2	0.006	88
Reading daily	0.082	3	0.147	0	0.126	0	0.051	21

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for the target population. The common support quantile is now fixed to 97%. Inference based on 999 bootstrap replications (except column 2, in which p-values do not take into account that the propensity score is estimated). The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.23 – *Average effects of music vs. sports on missing outcomes*

	Music only vs. sport only		**Music only vs. sport (competitive only)**		**Music + sport vs. sport only**		**Music + sport vs. music only**	
	(1)		(2)		(3)		(4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Indicators: 1 if at least one outcome in the respective group is missing								
Missing cognitive skills	-0.006	88	0.037	61	-0.044	24	-0.039	38
Missing: One of the outcomes measured since 2006	-0.012	58	-0.001	97	0.026	28	0.037	21
Missing: Health outcomes at age 18	0.061	11	0.026	38	0.008	79	-0.051	25
Missing: Outcomes measured in every survey year	-0.023	46	0.001	96	0.015	60	0.037	26

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for the target population. Inference based on 4999 bootstrap replications.

Table B.24 – *Average effects of music vs. sports (using survey weights)*

	Music only vs. sport only (1)		**Music only vs. sport (competitive only)** (2)		**Music + sport vs. sport only** (3)		**Music + sport vs. music only** (4)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
Cognitive skills and school achievements at age 17								
Cognitive skills - Average	0.169	16	0.008	93	0.267	0	0.100	45
Cognitive skills - Analogies	0.246	4	0.093	36	0.292	0	0.047	74
Cognitive skills - Figures	0.184	16	0.051	62	0.184	4	0.002	99
Cognitive skills - Math	0.020	85	-0.114	27	0.107	26	0.088	50
School grades - Average	-0.006	95	0.070	39	0.122	10	0.135	18
School grades - Maths	-0.175	19	-0.036	60	0.090	22	0.263	1
School grades - German	0.168	3	0.115	16	0.084	20	-0.072	44
School grades - 1st for. lang.	0.039	63	0.096	19	0.095	21	0.063	52
Attends upper sec. school	0.131	0	0.065	9	0.156	0	0.032	53
Aim to enrol at university	0.097	4	0.093	0	0.182	0	0.090	7
Attends university at age 20	0.023	73	0.084	1	0.111	0	0.088	7
Personality at age 17								
Big 5: Conscientiousness	-0.161	15	-0.027	77	0.107	19	0.263	3
Big 5: Extraversion	-0.219	4	-0.181	3	0.174	1	0.390	0
Big 5: Neuroticism	0.172	9	0.173	5	-0.002	96	-0.175	12
Big 5: Openness	0.355	0	0.379	0	0.377	0	0.023	86
Big 5: Agreeableness	0.099	43	0.094	30	0.234	1	0.135	32
Willingness to take risk	-0.070	54	-0.227	3	-0.019	83	0.053	70
Subjective health and life style at age 18								
Satisfaction with health	-0.357	19	-0.140	28	0.125	48	0.481	1
Current health situation	-0.188	4	-0.093	12	0.024	76	0.214	2
Currently smoking	-0.046	32	-0.050	14	-0.128	0	-0.083	9
Other leisure activities at age 17								
Watching TV daily	-0.060	8	-0.080	1	-0.093	0	-0.034	43
Playing computer games daily	-0.074	4	-0.066	1	-0.053	7	0.020	64
Reading daily	0.091	1	0.137	0	0.144	0	0.057	22

Note: SOEP v29, 2001-2012, own calculations. Effects presented are average treatment effects for the target population. Inference based on 499 bootstrap replications (except columns 1 and 4, in which p-values do not take into account that the propensity score is estimated). The measurement of all outcomes is described in Table B.32. School grades, cognitive skills, personality, and aggregate outcomes are normalised to mean zero and variance 1 (higher value of grades is better). All other outcome variables are binary, except for the following: Satisfaction with health (0 "Worst" to 10 "Best") and current health situation (1 "Bad" to 5 "Very good").

B.3 Details on the matching estimator

Table B.25 gives the detailed matching protocol used. See Huber et al. (2013) and Huber et al. (2015) for further details on the implementation and the properties of this estimator, including trimming and common support rules.

Table B.25 – *Matching protocol for the estimation of a counterfactual outcome and the effects*

Step A-1	Choose one observation in the subsample defined by $d = 1$ and delete it from that pool.
Step B-1	Find an observation in the subsample defined by $d = 0$ that is as close as possible to the one chosen in step A-1 in terms of $(x), \tilde{x}$. 'Closeness' is based on the Mahalanobis distance.
Step C-1	Repeat A-1 and B-1 until no observation with $d = 1$ is left.
Step D-1	Compute the maximum distance (dist) obtained for any comparison between a member of the reference distribution and matched comparison observations.
Step A-2	Repeat A-1.
Step B-2	Repeat B-1. If possible, find other observations in the subsample of $d = 0$ that are at least as close as $R \cdot$ *dist* to the one chosen in step A-2. Do not remove these observations, so that they can be used again. Compute weights for all chosen comparisons observations that are proportional to their distance. Normalise the weights such that they add to one.
Step C-2	Repeat A-2 and B-2 until no participant in $d = 1$ is left.
Step D-2	For any potential comparison observation, add the weights obtained in A-2 and B-2.
Step E	Using the weights $w(x_i)$ obtained in D-2, run a weighted linear regression of the outcome variable on the variables used to define the distance (and an intercept).
Step F-1	Predict the potential outcome $y^0(x_i)$ of every observation using the coefficients of this regression: $\hat{y}^0(x_i)$
Step F-2	Estimate the bias of the matching estimator for $E(Y^0 \vert D = 1)$ as: $\sum_{i=1}^{N} \frac{(1-d_i)w_i\hat{y}^0(x_i)}{N_0} - \frac{d_i\hat{y}^0(x_i)}{N_1}$.
Step G	Using the weights obtained by weighted matching in D-2, compute a weighted mean of the outcome variables in $d = 0$. Subtract the bias from this estimate to get $E(Y^0 \vert D = 1)$.

Note: R is set to 150%. Gender is used as additional matching variable ($\tilde{x}$). The weight of the propensity score coming from the Mahalanobis metric is multiplied by 5. Control observations with weights larger than 5% are removed.

Table C.1 shows how to compute the average treatment effects on the treated (ATET). The average treatment effects on the non-treated (ATENT) are computed in the same way but the role of treated and controls are reversed in the algorithm. The average treatment effect (ATE) is computed by aggregating both effects using the treatment share ($ATE = ATET \cdot P(D = 1) + ATENT \cdot P(D = 0)$). These estimations are performed for each of the six subsamples separately. The weights used to aggregate these estimates over specific groups or the population are proportional to the number of treated (for the ATET) / controls (ATENT) / populations (ATE) in the respective subsamples.

See also Huber et al. (2015) for operational details of this estimator. The particular version of this estimator used is the RAD_MATCH Gauss package version 3.2.1. It has the feature that sampling weights may be accounted for in general. Furthermore, bootstrap inference as described in Huber et al. (2015) is based on weights that are combination of sampling weights, matching weights as well as regression weights. The improved bootstrap smoother as proposed by Racine and MacKinnon (2007) is used to economize on the required bootstrap replications. In addition, the variable *gender* is used as additional variables in the Mahalanobis step, in which the propensity score is overweighed by a factor of 5. The distance measure is set to 150%.

B.4 Robustness with respect to the validity of the selection on observables assumption

Sensitivity test – Ichino et al. (2008)

The test relies on the simulation of a confounder, U, that is correlated to the outcomes and to the treatment. The sensitivity of the specification to a violation of the CIA is assessed by comparing results with the simulated confounder U and results without. The confounder U is binary and independent of included covariates. The confounder is simulated such that its probability (p^{ij}) for taking a value of 1 varies for the four strata defined by the outcome and the treatment i.e. $p^{ij} = P(U = 1|D = i, Y = j), i, j \in \{0, 1\}$ (Y denotes the outcome variable and D denotes the treatment). Setting p^{ij} to 0.5 for all cases would therefore be the same as simulating a confounder U that does not bias the results. This probability scenario is similar to the "without the confounder U" scenario (scenario 1). It is used as the baseline when comparing results. To simulate a confounder U which p^{ij} varies with treatment and outcome, we use the correlation pattern of an important observed confounder: gender (scenario 2). In other words, the probability (p^{ij}) is similar to the probability for gender, which varies among the strata (and among the different treatments that we use). By doing so we test the sensitivity of our specification to a missing unobserved confounder that is correlated to treatment and potential outcome in the same way as gender is. We choose gender because it is the most important confounder in terms of size and significance in the p-score. Furthermore, we believe that the unobserved confounders are closely related to gender (behaviours, parents' education and preferences, for example). A third scenario is simulated in which we use the correlation pattern of another important covariate: the recommendation for upper secondary school (scenario 3). This covariate is also an important confounder which is used as a proxy for cognitive skills. We believe that the unobserved confounders might be correlated with this proxy.

Following Ichino et al. (2008) suggestion, we repeat the simulations several times and take the mean of the effects over the simulations. We draw 19 times U (i.e. we have 19 random realisations of U for each observation) within each bootstrap replication and use 99 bootstrap replications for each scenario. Then, we compare the results between the baseline ($p^{ij} = 0.5, \forall i, j$) and the two biased scenario (scenario 2 and 3). The deviations from the baseline results and their significance are reported in Table B.26 for scenario 2 and in Table B.27 for scenario 3. The deviation is computed for the ATE for the target population with linear bias correction for the 4 treatments. The deviation is never significant at the 5% threshold. Our specification is not sensitive to a violation of the CIA in the tested scenario.

Table B.26 – *Average effects of music vs. sports (scenario 2: a confounder similar to gender)*

	Music only vs. sport only	**Music only vs. sport (competitive) only**	**Music + sport vs. sport only**	**Music + sport vs. music only**
	(1)	(2)	(3)	(4)
	Difference to baseline	*Difference to baseline*	*Difference to baseline*	*Difference to baseline*
Cognitive skills and school achievement at age 17				
Cognitive skills - Average	0.007	-0.020	0.001	-0.015
Cognitive skills - Analogies	-0.028	-0.025	0.031	-0.003
Cognitive skills - Figures	0.005	0.011	-0.008	-0.017
Cognitive skills - Math	0.033	-0.027	-0.012	-0.006
School grades - Average	-0.004	0.024	0.008	-0.016
School grades - Maths	-0.012	0.002	0.014	-0.016
School grades - German	-0.001	0.038	0.005	-0.019
School grades - 1st for. lang.	0.008	0.009	0.001	-0.005
Attends upper sec. school	-0.008	-0.011	0.003	0.014
Attends university at age 20	0.003	-0.007	-0.006	0.028
Aim to enrol at university	0.001	-0.002	0.005	0.012
Personality at age 17				
Big 5: Conscientiousness	0.013	0.004	-0.034	0.012
Big 5: Extraversion	0.040	0.018	-0.041	0.085
Big 5: Neuroticism	0.007	-0.007	0.031	0.007
Big 5: Openness	-0.024	0.008	-0.012	0.081
Big 5: Agreeableness	0.015	-0.013	-0.006	-0.008
Willingness to take risk	0.035	0.002	-0.006	0.028
Subjective health and life style at age 18				
Satisfaction with health	-0.008	-0.050	0.026	0.052
Current health situation	0.017	0.015	-0.003	-0.032
Currently smoking	-0.010	0.001	0.003	0.007
Other leisure activities at age 17				
Watching TV daily	0.001	-0.005	0.001	0.010
Playing computer games daily	0.004	-0.010	-0.004	0.001
Reading daily	-0.002	-0.009	-0.008	0.015

Note: SOEP v29, 2001-2012, own calculations. Difference of effects in confounding scenario to baseline scenario. For the confounding scenario, a confounder U is simulated with probabilities: $P(U=1|Y=1,D=1)$; $P(U=1|Y=0,D=1)$; $P(U=1|Y=1,D=0)$; $P(U=1|Y=0,D=0)$ according to those of the variable gender. Inference based on 99 bootstrap replications and 19 draws of simulated binary confounder; quantile method, smoothed version, linear bias adjustment, symmetric p-values used. None of the differences are statistically significant at the 10% level. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills and personality are normalised. All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Table B.27 – *Average effects of music vs. sports (scenario 3: a confounder similar to the rec-ommendation for upper secondary school)*

	Music only vs. sport only	**Music only vs. sport (com-petitive) only**	**Music + sport vs. sport only**	**Music + sport vs. music only**
	(1)	(2)	(3)	(4)
	Difference to baseline	*Difference to baseline*	*Difference to baseline*	*Difference to baseline*
Cognitive skills and school achievement at age 17				
Cognitive skills - Average	-0.027	0.004	-0.011	0.010
Cognitive skills - Analogies	-0.033	0.002	0.001	0.005
Cognitive skills - Figures	-0.009	0.015	-0.017	-0.033
Cognitive skills - Math	-0.020	-0.007	-0.014	0.026
School grades - Average	0.002	-0.014	0.008	-0.002
School grades - Maths	0.003	-0.012	0.007	-0.006
School grades - German	0.004	-0.008	0.001	0.001
School grades - 1st for. lang.	-0.005	-0.014	0.012	0.007
Attends upper sec. school	-0.011	-0.007	-0.003	0.007
Attends university at age 20	-0.018	-0.001	-0.015	0.016
Aim to enrol at university	-0.008	-0.001	0.000	0.001
Personality at age 17				
Big 5: Conscientiousness	-0.026	0.027	-0.014	0.000
Big 5: Extraversion	0.009	-0.026	-0.020	0.092
Big 5: Neuroticism	0.032	-0.008	0.007	0.015
Big 5: Openness	0.007	-0.003	-0.030	0.028
Big 5: Agreeableness	0.034	-0.011	0.000	-0.081
Willingness to take risk	0.025	-0.048	-0.009	0.064
Subjective health and life style at age 18				
Satisfaction with health	0.012	-0.024	0.001	0.064
Current health situation	0.015	0.002	-0.012	-0.019
Currently smoking	-0.007	0.000	0.004	0.007
Other leisure activities at age 17				
Watching TV daily	-0.013	0.002	0.000	0.033
Playing computer games daily	0.006	-0.006	0.004	0.000
Reading daily	0.001	0.003	-0.011	0.003

Note: SOEP v29, 2001-2012, own calculations. Difference of effects in confounding scenario to baseline scenario. For the confounding scenario, a confounder U is simulated with the following probabilities: $P(U=1|Y=1,D=1)$; $P(U=1|Y=0,D=1)$; $P(U=1|Y=1,D=0)$; $P(U=1|Y=0,D=0)$ according to those of the variable recommendation for upper secondary school. Inference based on 99 bootstrap replications and 19 draws of simulated binary confounder; quantile method, smoothed version, linear bias adjustment, symmetric p-values used. None of the differences are statistically significant at the 10% level. The measurement of all outcomes is described in Table B.32. School grades, cognitive skills and personality are normalised. All other outcome variables are binary, except for the following: Satisfaction with health (0 “Worst” to 10 “Best”) and current health situation (1 “Bad” to 5 “Very good”).

Instrumental variable estimation

Concepts and instruments used

We need an instrumental variable (IV) as an exogenous source of variation in the choice of an extracurricular activity. This IV should induce the individual to choose music rather than sports: we use an indicator of whether either parent played a musical instrument in their youth. It is a binary indicator coded to "1" if either parent declares having played a musical instrument, "0" otherwise. In order to identify a causal effect, this IV must be strong and valid.

We argue that parental engagement in musical activities during their own adolescence is a strong IV for the child's decision to play music rather than or in addition to sports. The relevance of that variable can be justified both theoretically and empirically. Parents play an important role in choosing their children's extracurricular activities, given that they provide financial, logistic and psychological support (Gustafson and Rhodes, 2006). Following a classical Becker-type reasoning (Becker, 1965), the decision to enrol in music classes or a sports club is likely to be motivated by maximizing utility, both with respect to current pleasure and future investments in skills. Assuming that individuals have imperfect information, personal experience is likely to influence their choice with respect to their child's leisure activity. In addition, the choice of activity is likely to depend on taste. This reinforces the relevance of our IV. Taste is shaped either by education or repeated consumption (Garboua and Montmarquette, 1996; Lunn and Kelly, 2009; Seaman, 2006). Hence, the child is likely to share the same taste as her parents. Hille and Schupp (2015) argue that the parents' taste for the arts is a very strong predictor of their child's enrolment into music practice.

The crucial assumption for an IV to be valid is that it has no direct effect on the outcome. This assumption is known as exogeneity or mean exclusion restriction (Frölich, 2007). The mean exclusion restriction may be invalid due to reverse causality, observable and unobservable characteristics determining the IV as well as a direct effect of the IV through a channel other than the treatment. We discuss each of these potential violations in the following.

Reverse causality can be ruled out since the IV is measured prior to the child's birth. Still, the IV – childhood musical activity of the parents – may affect the outcome due to the influence of its determinants. If education or other parental characteristics – besides providing an incentive to engage with music – influence the child's skill development, the exogeneity assumption is violated. Parental participation in artistic activities is linked to their level of education and their socio-economic environment. These characteristics are correlated with their labour-market success and thus with their child's development. We take this endogeneity into account by adding covariates (the same we use for in the selection-on-observables framework in the main part of this paper).

A final source of violation of the exogeneity assumption exists, if the IV has a direct effect on the outcomes due to the effects of parental activities on parental skills. If being active in music has affected parental skill development differently than being active in sports and if this has consequences on the child's skill development, our IV is no longer exogenous. Music activities might have benefited the parents' skills during childhood and youth. While personality has been found malleable during childhood (e.g. Cobb-Clark and Schurer, 2012; Donnellan and Lucas, 2008; Heckman and Kautz, 2012; Specht et al., 2011), it is considered stable among adults (Pervin et al., 2005). We control for parental personality, which may have been affected by childhood leisure activities. Some of the advantages resulting from parental differences in past leisure activity are taken into account by controlling for education, income and labour market situation as described above.

Finally, the identification of the local average treatment effect using IV requires the assumption of monotonicity (Frölich, 2007). We need to assume that no adolescent chooses to become active in sports rather than music, even though their parents were engaged in artistic activities several years before. While we cannot rule out the existence of rebellious children, we argue that, given the age at which children decide between music and sports, they are unlikely to drive our results and that the monotonicity assumption is satisfied due to our sample choice.

Sample selection and descriptive statistics

The subsample for which the IV is available is roughly half the size of the original sample and somewhat different. The high-income group is overrepresented and immigrants as well as East Germans are underrepresented. Being part of the SOEP design, these differences in sample composition are exogenous to our questions of interest. We expect mean outcomes to differ from the original sample.

Results

As suggested by Frölich (2007), we estimate the LATE non-parametrically by using a propensity score to reduce dimensionality. This is necessary, given that the sample is relatively small and that we need to control for a large number of covariates. The estimation of the LATE with covariates (propensity score in our case) is equivalent to the computation of two matching estimators: the effect of the instrument on the outcome divided by the effect of the instrument on the treatment (conditional on a set of covariates). This effect concerns solely the subpopulation of compliers: individuals who react to the instrument (i.e. adolescents who play music because at least one of their parents was artistically active at the age of 16).

The results of the estimations are reported in Table B.28. The first stage represents the effect of the instrument on the probability of being treated. The fact that at least one of the parents played a musical instrument during their adolescence is a good predictor of the probability that the adolescent herself is engaged in music. As stated above, the LATE is different from the ATE; therefore, the results cannot be directly compared to the main results of the paper. We find that practicing music instead of sports has a detrimental effect on school grades in mathematics for the complier population (columns 1 and 2 of Table B.28). Last, practicing music instead of doing sport competitively increases the openness of the adolescent.

Table B.28 – *Average effects of music vs. sports (IV estimations)*

	Music only vs. sport only (1)		**Music only vs. sport (competitive) only** (2)		**Music + sport vs. sport only** (3)	
	Effect	*p-value in %*	*Effect*	*p-value in %*	*Effect*	*p-value in %*
First stage	0.071	1	0.144	0	0.099	0
Cognitive skills and school achievement at age 17						
Cognitive skills - Average	0.593	89	0.293	88	0.424	94
Cognitive skills - Analogies	-2.460	65	-1.214	60	-1.760	66
Cognitive skills - Figures	-0.962	84	-0.475	81	-0.688	85
Cognitive skills - Math	4.456	54	2.199	48	3.188	81
School grades - Average	-0.907	17	-0.448	13	-0.649	23
School grades - Maths	-2.614	6	-1.290	2	-1.870	18
School grades - German	-0.326	68	-0.161	63	-0.234	67
School grades - 1st for. lang.	-0.226	79	-0.112	76	-0.162	88
Attends upper sec. school	0.126	75	0.062	69	0.090	81
Attends university at age 20	-0.561	18	-0.277	11	-0.402	22
Aim to enrol at university	-0.057	86	-0.028	85	-0.041	88
Personality at age 17						
Big 5: Conscientiousness	-0.130	56	-0.064	50	-0.093	75
Big 5: Extraversion	-0.014	94	-0.007	94	-0.010	96
Big 5: Neuroticism	0.150	48	0.074	40	0.107	66
Big 5: Openness	0.297	26	0.147	9	0.213	27
Big 5: Agreeableness	0.213	27	0.105	18	0.152	67
Willingness to take risk	0.944	69	0.466	66	0.676	71
Subjective health and life style at age 18						
Satisfaction with health	-0.429	86	-0.212	83	-0.307	87
Current health situation	-0.313	75	-0.155	72	-0.224	83
Currently smoking	-0.334	55	-0.165	49	-0.239	61
Other leisure activities at age 17						
Watching TV daily	-0.092	82	-0.045	79	-0.066	84
Playing computer games daily	0.412	43	0.203	30	0.294	75
Reading daily	0.371	48	0.183	37	0.266	42

Note: SOEP v29, 2001-2012, own calculations. Inference based on 399 bootstrap replications. Contrary to the estimations in other tables of this paper, school grades, cognitive skills and personality are not normalised, but measured from 0 to 20 for cognitive skill, from 6 to 1 for school grades and from 0 to 1 for personality. The measurement of all other outcomes is described in Table B.32. All other outcome variables are binary, except for some health measures as indicated.

Characteristics of dropouts

Potential threat to identification:

Differences in reasons to give up music or sports

Our treatment definition considers adolescents as playing music or sports if they do so at age 17 and have started at age 14 or earlier. With other words, those who played sports or a musical instrument as a child and gave up before age 17 (before answering the SOEP youth questionnaire), are not considered treated. Moreover, we have no information on sports and music participation for these adolescents, given that questions on starting age and other details of involvement with music or sports are only asked if the adolescent states to be still engaged with the activity at age 17. When describing our identification strategy, we argue that this is not a problem, because the reasons for giving up music and sports are similar, and we therefore consider similar adolescents as not being active, even though they gave up their activity before age 17 and might differ from those who never played music or sports.

Possible test:

Subsample with information on history of music/sports involvement

Since 2006, parents are asked about extracurricular activities for each of their children under the age of 16. These questions are included biannually in the SOEP household questionnaire. Using these answers, we can compare the characteristics of adolescents who gave up music or sports before answering the SOEP youth questionnaire for a subsample of 752 individuals. For them, we have biannual information on music and sports participation at least from age 13/14.

For both music and sports, we constructed a variable taking the values 0 “Never played”, 1 “Played at least at the time of one SOEP interview between age 11 and 14, but does not play at age 17 (is not treated)” and 2 “Treated (plays at age 17 and started at age 14 or before)”. Table B.29 illustrates the coding of this variable. According to this definition, only 65 of the 752 adolescents (8.6%) gave up music before age 17 and 133 adolescents (17.7%) gave up sports.

Table B.30 shows the history of sports and music participation between age 14 and 17 by treatment status within the reduced sample described above. The table shows that among those who only play sports, 15% also played music at age 14, but gave up before answering the SOEP youth questionnaire at age 17. Among music participants, even 45% gave up sports between age 14 and 17. With other words, only 85% of the sports and 55% of the music participants were not involved with the other activity between age 14 and 17. These figures are comparable given the prevalence of both activities among adolescents: for every four adolescents who are active in music or sports according to our definition (playing at age 17 and having started at age 14 or before), approximately another three were active at age 14, but gave up before age 17. Among those who play neither of the activities according to our definition, 43% actually played sports and 8% played music a couple of years earlier, but gave up before answering the youth questionnaire.

Whether the existence of adolescents who gave up one activity within the treatment group of those active in the other threatens our identification strategy depends on whom we actually compare. If the reasons to give up music and sports are similar, these 'partly treated' individuals should not be of concern. Although we have no information on the reasons for giving up music or sports activities, we can compare the characteristics adolescents who gave up one of these activities between age 14 and 17 to those who did not play at age 14, as well as those who we consider as treated (who started at age 14 or before and still play at 17).

Table B.31 shows covariates and outcomes for different histories of sports and music participation. In general, adolescents who gave up music or sports are similar. In terms of socio-economic status, they are situated between those who still play at age 17 and those who never played between age 14 (or 12) and 17. The only exception is migration background: Parents of adolescents who gave up sports are more likely to have a migration background than parents of other adolescents. With respect to the parents' ISEI status and education, drop-outs are more similar to treated than to never-playing individuals.

With respect to individual characteristics, drop-outs do have some notable characteristics. In contrast to the finding for SES stated above, school grades of drop-outs are more similar to those who never played than to treated adolescents. This might reflect the fact that difficulties at school are an important reason to stop playing music or sports. Moreover, music drop-outs are slightly less conscientious and more emotionally stable. Drop-outs of both music and sports also have a higher willingness to take risks and are slightly more often smokers. Table B.31 is restricted to variables that help characterize

Table B.29 – *Possible patterns of music/sports participation and how we summarize it*

Possible patterns of music/sports participation and treatment status			**Coding of new variable**
Playing at age 11 or 12	*Playing at age 13 or 14*	*Playing at age 16/17 and having started at age 14 or before (Treated)*	
No	No	No	Never played
(information missing)	No	No	
No	(information missing)	No	
Yes	No	No	Gave up before age 17
No	Yes	No	
Yes	Yes	No	
(information missing)	Yes	No	
Yes	(information missing)	No	
		Yes	Treated

Note: Each line represents one possible pattern of for the three variables "Playing at age 11 or 12", "Playing at age 13 or 14" and "Treated" and how they were interpreted when coding the variable "Giving up".

Table B.30 – *History of music and sports participation by treatment status*

	Comparison groups used in this study			
	Neither sports nor music	*Sports only*	*Music only*	*Music and sports*
Total sample (N)	1374	1640	333	501
Reduced sample (N) (information on sports/music available since age 12 or 14)	219	311	84	138
Actual history of sports participation (in % of reduced sample)				
Never played	57	0	55	0
Gave up between age 14 and 17	43	0	45	0
Started at age 14 or before, played at age 17	0	100	0	100
Actual history of music participation (in % of reduced sample)				
Never played	92	85	0	0
Gave up between age 14 and 17	8	15	0	0
Started at age 14 or before, played at age 17	0	0	100	100

Note: SOEP v29, 2001-2012, own calculations. History of sports and music participation between age 12 (14 for some adolescents) and 17 by treatment group status. For each activity, the values show the percent of adolescents among the reduced sample, who never played the activity between age 14 and 17, who played at age 14, but no longer did so at age 17, and those who played at age 14 and 17.

adolescents, who give up music or sports during adolescence. Many further characteristics are very similar across these groups of extracurricular participation. Despite the small differences stated above, music and sports drop-outs do not differ much and should therefore not threaten the main results of this paper.

Table B.31 – *Characteristics of never-players, drop-outs and active adolescents by activity*

	Sports			**Music**		
	Never played	*Gave up*	*Treated Treated*	*Never played*	*Gave up*	*Treated Treated*
Covariates						
Female	0.59	0.47	0.45	0.47	0.49	0.53
Recommendation for upper sec. school	0.23	0.40	0.51	0.34	0.49	0.59
Household net wealth (in 100,000 Euros)	2.19	1.49	2.36	1.57	2.20	3.42
ISEI socio-economic status (0 to 90)	42	48	50	44	54	54
Parent with upper sec. school degree	0.26	0.34	0.43	0.28	0.55	0.53
Parent with migration background	0.21	0.29	0.19	0.23	0.18	0.18
Willingness to take risks (mother, 0 to 10)	3.84	4.09	4.03	3.99	4.36	3.92
Outcome variables						
Average cognitive skills	-0.18	-0.11	0.13	-0.11	0.04	0.27
Average mark	0.05	-0.05	0.14	-0.04	0.07	0.34
Conscientiousness	-0.07	-0.03	0.08	0.02	-0.11	0.09
Neuroticism	0.11	-0.01	-0.03	0.02	-0.17	0.03
Willingness to take risks	-0.20	0.14	0.02	-0.03	0.12	0.01
Currently smoking (age 18)	0.18	0.23	0.13	0.17	0.19	0.12

Note: SOEP v29, 2001-2012, own calculations. Covariates and outcomes for different histories of sports and music participation. Most covariates are binary, except if otherwise indicated (household wealth and ISEI status). Most outcome variables are normalized to have mean zero and variance 1, except ‘currently smoking’ (binary).

B.5 Measurement of outcome variables

Table B.32 – *Description of outcome variables*

Outcome variables	**Description**
Cognitive skills and school achievement at age 17	
Cognitive skills - Average	Mean value of analogies, figures and math (see below)
Cognitive skills - Analogies	Identify word pairs (e.g. forest – tree is equivalent to meadow – ? [grass])
Cognitive skills - Figures	Identify the following symbol in a row of symbols
Cognitive skills - Math	Insert maths operators in small calculus problems
School grades - Average	Mean value of maths, German and foreign language grade (see below)
School grades - Maths	Mathematics grade in latest school report (self-reported)
School grades - German	German grade in latest school report (self-reported)
School grades - 1st for. lang.	1st foreign language grade in latest school report (self-reported)
Attends upper sec. school	The adolescent attends an upper secondary school at age 17
Aim to enroll at university	Self-reported aim (at age 17) to attend higher education
Attends university at age 20	Available for adolescents who are still part of the SOEP at age 20 (about 2/3 of full sample)
Personality at age 17	
	Big 5: Mean value of 3 or 4 items. For each, the adolescent states the degree of approval on a 7-point Likert scale. The items are: “I see myself as someone who...”
Big 5: Conscientiousness	...does a thorough job; ...does things effectively; ...tends to be lazy (reversed)
Big 5: Extraversion	...is communicative, talkative; ...is outgoing, sociable; ...is reserved (reversed)
Big 5: Neuroticism	...worries a lot; ...gets nervous easily; ...is relaxed, handles stress well (reversed)
Big 5: Openness	...is original, comes up with new ideas; ...values artistic experiences; ...has an active imagination; ...is eager for knowledge
Big 5: Agreeableness	...is sometimes somewhat rude to others (reversed); ...has a forgiving nature; ...is considerate and kind to others
Willingness to take risk	0 (not at all willing to take risks) ... 10 (very willing to take risks)
Subjective health and life style at age 20	
Satisfaction with health	0 (not at all satisfied) ... 10 (very satisfied)
Current health situation	1 “bad” 2 “less good” 3 “satisfactory” 4 “good” 5 “very good”
Currently smoking	Yes/no
Other leisure activities at age 17	
Watching TV daily	Yes/no
Playing computer games daily	Yes/no
Reading daily	Yes/no

Note: SOEP v29, 2001-2012, own calculations. The exact definitions of each activity are given in Table B.1.

Appendix to “Improving access to non-formal education – evidence from the German ‘Educational package’”

C.1 Estimation sample

Treatment of missing covariates

1. *Mother’s work hours, labor market status and age at birth (originally 105 missings):* These variables are replaced with the father’s values if the child lived only with the father in 2010 or if it lived with both parents, but only the father answered the FiD individual questionnaire. Given that the mother was either not present in the household, or did not take the time to respond to the individual questionnaire, I assume that the father takes the most important household decisions (also with respect to the child). Thus, the father’s information on these variables might be relevant. In cases, where the mother answered the individual questionnaire, but did not provide information on these variables, I leave them missing. This step reduces the number of missings for mother’s work hours to 37 (see step 5 below), mother’s labor market status to 1 and mother’s age to 0.

2. *Migration background of the child (originally 11 missings):* This variable is replaced by that of the parents in the following way: a) if the child lived with only one parent in 2010, it is replaced with that parent’s value (5 cases); b) if the child lived with both parents in 2010 and both have valid information on migration background, it is set to 1 if one of the parents has a migration background (4 cases); c) if the child lived with both parents in 2010 and parental migration background is missing for at least one parent, the variable is left missing (2 cases).

3. *One parent with Abitur and one parent with university degree (92 missings):* I recode missing values to 0 and additionally include a missing indicator, which is 1 if one of the two variables is missing.

4. *Household expenditure on education, culture and leisure (116 missings):* I recode missing values to the variable’s mean and include a missing indicator, which is 1 if

one of the three variables is missing. In only 6 cases, both parental education and household expenditures are missing.

5. *Mother's work hours in 2010 (37 missings):* I recode missing values to the variable's mean and include a missing indicator, which is 1 if mother's work hours are missing. In only 3 cases, both mother's work hours and parental education are missing, in 4 cases, both mother's work hours and household expenditures are missing.

6. *Child's migration background, mother's labor force status, school grade:* I drop the 9 individuals for whom this information is missing even after the corrections desribed above (see Table C.1).

Table C.1 – *Sample construction*

Sample...	**Remaining observations**			
Basic sample				
All children (below 17) in FiD	11252			
...drop if born after 2007 (too young)[1]	8256			
	Sample 2010-2012		*Sample 2010-2013*	
...in sample as a child in 2010 and 2012/2013[2]	4243		3483	
Treatment				
...without missing treatment in 2010 and 2012/2013	4243		3482	
Covariates				
...drop children living neither with father nor mother (2010)	4239		3479	
...drop if the child's migration background or school grade, or the mother's work hours are missing	4232		3474	
	All children	Students	All children	Students
Outcomes				
...part of outcome sample subgroup	4232	2440	3474	1913
...without missings on any outcome in any year	4221	2432	3462	1902
Final sample	**4221**	**2432**	**3462**	**1902**
Sample for propensity score matching estimation[3]	**4216**	**2386**	**3457**	**1871**

[1] Due to the focus of FiD on very young children, the sample is considerably reduced in this step.
[2] The strong reduction in sample size is due to the fact that 1959 children enter the sample only after 2010 (exogenous). 1910 children are no longer in the sample in 2012 (361 because they are no longer children, 832 because they dropped out of the study and 717 because parents did not provide information for the child). Among the 760 children, who are in the sample in 2012, but no longer in 2013, 223 are no longer children, 21 left the data, and for 516, parents did not provide any information (which could also mean that they have left the data, but are still considered as survey households). 144 children are left out of Sample 2010-2013, because parents answered the relevant questions of the household questionnaire in 2013, but not in 2012.
[3] Excludes treatment observations with a propensity score higher than the highest propensity score in the control group. Figure C.1 illustrates, how many treated observations are "off support".

Table C.2 – *Detailed description of all variables*

Variable	**Description**	**Type**	**Available**
TREATMENT VARIABLE			
Eligibility Educational package	Household receives basic income for jobseekers, social assistance, housing assistance, or complentary child allowance	Binary	2010, 2012, 2013
OUTCOME VARIABLES			
Sports outside school	Schoolchildren, directly asked in household questionaire	Binary	2010, 2012, 2013
Music/sports outside school	Schoolchildren, constructed from 2 questions (sports, music) in household questionnaire	Binary	2010, 2012, 2013
Music/sports at/outside	Schoolchildren, constructed from 4 questions (sports/music; at school/outside school)	Binary	2010, 2012, 2013
Sports outside	Outside school or kindergarden, all children, constructed from 2 questions (for schoolchildren and pre-school children)	Binary	2010, 2012, 2013
Music/sports at/outside	Outside school or kindergarden, all children, constructed from 8 questions (schoolchildren/pre-school children; music/sports; at school or kindergarden/outside school or kindergarden)	Binary	2010, 2012, 2013
GROUP VARIABLES FOR HETEROGENEITY ANALYSIS			
Household income >/< median	Median calculated among eligible children	Binary	2012, 2013
Single parent home	Child lives with only one parent: constructed from comparison of child's household ID with that of each parent. "Yes" if household ID matches 1 parent and the other parent is unknown or has a different household ID. Children living with neither parent are dropped (see Table C.1)	Binary	2012, 2013
Monthly household expenditure for culture above median	Median calculated among eligible children	Binary	2012, 2013
Female	0 "male", 1 "female", fieldwork information	Binary	–
VARIABLES USED IN SECTION 4.5 (SUBSECTION ON TAKE-UP)			
Take-up of Educational package	Question in household questionnaire for each child: did child receive subsidy from Educational package? One item per subsidy: cultural participation, school trips (one day), class trips (several days), transportation to school, school lunch. Information for school and class trips were summarized in one variable, given that both subsidies are very similar	Binary	2012, 2013
CONTROL VARIABLES			
Individual characteristics			

Table C.2 to be continued.

Table C.2 continued.

Variable	Description	Type	Available
Female	0 “male”, 1 “female”, fieldwork information	Binary	–
Migration background (child)	Direct or indirect migration background, from recruitment interview, replaced with that of the parents if missing	Binary	–
Attending school	Information from parents asked in household questionnaire	Binary	2010
School type	Information from parents asked in household questionnaire, includes answer options primary school, lower secondary school, medium secondary school, upper secondary school, comprehensive school, other school type, coded 0 if not in school	Dummy	2010
School grade	Information from parents asked in household questionnaire, coded 0 if not in school	Binary	2010
Full-day school	Information from parents asked in household questionnaire, 0 “No” 1 “Yes”, coded 0 if not in school	Binary	2010
Birth year	Fieldwork information	Dummy	–
Household characteristics			
Nb of children under 16 in household	Fieldwork information	Continuous	2010
Single parent home	Child lives with only one parent: constructed from comparison of child’s household ID with those of each parent. “Yes” if household ID matches 1 parent and the other parent is unknown or has a different household ID. Children living with neither parent are dropped (see Table C.1)	Binary	2010
Parent with Abitur	At least 1 parent has upper sec. school degree (university entrance certificate), from generated individual information of parents, takes value from 2012 if available, otherwise replaced in the following order: 2011, 2010, 2013. For each year constructed in 4 steps: 1) 1 if either parent has Abitur if information non-missing for both parents 2) 1 if either parent has Abitur and child living with both parents 3) 1 if mother has Abitur and child lives with mother only 4) 1 if father has Abitur and child lives with father only	Binary	–
Parent with university degree	At least 1 parent has university degree, otherwise coding identical to previous variable	Binary	–
Log monthly net household income	From household questionnaire, if missing: 1st imputed value (provided by FiD)	Continuous	2010
Monthly household expenditure for education	Euros, asked in household questionnaire	Continuous	2010
Monthly household expenditure for culture	Euros, asked in household questionnaire	Continuous	2010

Table C.2 to be continued.

Table C.2 continued.

Variable	**Description**	**Type**	**Available**
Monthly household expenditure for leisure	Euros, asked in household questionnaire	Continuous	2010
Household lives in rural area	Population $<$ 20,000, from household questionnaire	Binary	2010
Household lives in large city	Population $>$ 100,000, from household questionnaire	Binary	2010
Federal state	Fieldwork information	Dummy	2010
FiD subsample	Fieldwork information, categories: single-parent families, multi-child families, low-income families, cohort samples of 2007, 2008, 2009, 2010	Dummy	–
Missing: Parental education	"Parent with Abitur" or "Parent with university degree" missing	Binary	–
Missing: Household expenditure	Monthly household expenditure for education, culture or leisure missing	Binary	2010
Mother-related covariates			
Age of mother at birth	Birth year child – birth year mother, fieldwork information, replaced with father's if missing	Continuous	–
Mother's work hours	Actual work hours if available, otherwise agreed upon work hours, 0 if not working, from mother's individual questionnaire, replaced with father's if missing	Continuous	2010
Mother not working	From mother's individual questionnaire. "Yes" if labour force status is "not working", replaced with father's if missing	Binary	2010
Missing: Mother's work hours		Binary	2010

C.2 Summary statistics

Table C.3 – *Summary statistics by eligibility (2012)*

	All children		Schoolchildren	
	Non-eligible	*Eligbile*	*Non-eligible*	*Eligbile*
Individual-level covariates				
Female	0.49	0.51	0.50	0.53
Migration background	0.32	0.40	0.29	0.36
Attends school	0.57	0.61	1.00	1.00
Primary school	0.31	0.36	0.55	0.60
Lower secondary school track	0.04	0.06	0.07	0.09
Middle secondary school track	0.08	0.07	0.14	0.12
Upper secondary school track	0.09	0.04	0.16	0.06
Comprehensive secondary school	0.02	0.04	0.04	0.06
Other school	0.46	0.44	0.04	0.08
School grade	2.36	2.34	4.17	3.82
Full-day school	0.12	0.18	0.21	0.29
Birth year	2002.10	2001.87	1999.47	1999.63
Household-level covariates				
Nb of children in household	2.48	2.74	2.52	2.76
Single-parent household	0.21	0.44	0.30	0.50
One parent with Abitur	0.37	0.08	0.33	0.07
Parent with university degree	0.33	0.06	0.29	0.06
Log monthly net household income	7.90	7.45	7.87	7.47
Monthly HH exp: education	28.34	11.54	30.50	11.71
Monthly HH exp: culture	18.01	8.90	21.43	9.97
Monthly HH exp: leisure	66.90	30.01	72.13	32.61
Rural area (population $<$ 20k)	0.51	0.39	0.52	0.38
Large city (population $>$ 100k)	0.24	0.28	0.24	0.29
Federal state	7.38	7.60	7.30	7.62
FiD subsample	63.96	64.14	64.50	64.44
Missing: parental education	0.04	0.03	0.04	0.02
Missing: household expenditure	0.04	0.04	0.05	0.04
Mother-related covariates				
Work hours	14.81	7.00	16.45	8.48
Mother not working	0.40	0.68	0.33	0.62
Age at birth	30.24	27.75	29.77	27.61
Missing: mother's work hours	0.02	0.00	0.02	0.01
Number of observations	3094	1127	1748	684

Source: FiD v4, own calculations. Summary statistics by eligibility in 2013 are presented in main text (Table 4.3).

Table C.4 – *Patterns of eligibility (sample restricted to schoolchildren)*

Sample 2010-12				Sample 2010-13			
Pattern of eligibility				*Pattern of eligibility*			
2010	*2012*	*N*	*%*	*2010*	*2013*	*N*	*%*
No	No	1454	59.8	No	No	1109	58.3
Yes	No	294	12.1	Yes	No	268	14.1
No	Yes	132	5.4	No	Yes	108	5.7
Yes	Yes	552	22.7	Yes	Yes	417	21.9

Source: FiD v4, own calculations. Patterns of eligibility for all children are found in main text (Table 4.1).

Table C.5 – *Summary statistics of all variables used*

	Mean	**Standard deviation**	**Min**	**Max**	**Number missings**
Treatment variables					
Eligible (2012)	0.27	0.44	0.00	1.00	0
Eligible (2013)	0.26	0.44	0.00	1.00	497
Outcomes					
Sports outside school (2010)	0.56	0.50	0.00	1.00	1784
Music or sports outside school (2010)	0.64	0.48	0.00	1.00	1784
Music or sports, at or outside school (2010)	0.74	0.44	0.00	1.00	1784
Sports outside school or kindergarten (2010)	0.49	0.50	0.00	1.00	0
Music or sports, at or outside school or kindergarten (2010)	0.62	0.49	0.00	1.00	0
Sports outside school (2012)	0.56	0.50	0.00	1.00	1060
Music or sports outside school (2012)	0.63	0.48	0.00	1.00	1060
Music or sports, at or outside school (2012)	0.73	0.45	0.00	1.00	1060
Sports outside school or kindergarten (2012)	0.54	0.50	0.00	1.00	0
Music or sports, at or outside school or kindergarten (2012)	0.72	0.45	0.00	1.00	0
Sports outside school (2013)	0.57	0.49	0.00	1.00	1397
Music or sports outside school (2013)	0.64	0.48	0.00	1.00	1397
Music or sports, at or outside school (2013)	0.73	0.44	0.00	1.00	1397
Sports outside school or kindergarten (2013)	0.56	0.50	0.00	1.00	759
Music or sports, at or outside school or kindergarten (2013)	0.73	0.44	0.00	1.00	759
Variables for heterogeneity analysis					
HH income > median (among eligible 2012)	0.78	0.41	0.00	1.00	0
HH income > median (among eligible 2013)	0.74	0.44	0.00	1.00	759
Single parent household (2012)	0.27	0.44	0.00	1.00	2
Single parent household (2013)	0.27	0.44	0.00	1.00	757
Exp. f. culture > median (among eligible 2012)	0.70	0.46	0.00	1.00	0
Exp. f. culture > median (among eligible 2013)	0.71	0.45	0.00	1.00	759
Female	0.49	0.50	0.00	1.00	0
Variables on take-up of Educational package					
Membership fees (2012)	0.04	0.19	0.00	1.00	0
Membership fees (2013)	0.04	0.20	0.00	1.00	497
School or class trips (2012)	0.07	0.26	0.00	1.00	0
School or class trips (2013)	0.09	0.29	0.00	1.00	497

Table C.5 to be continued.

Table C.5 continued.

	Mean	Standard deviation	Min	Max	Number missings
Transport to school (2012)	0.03	0.16	0.00	1.00	0
Transport to school (2013)	0.03	0.16	0.00	1.00	497
School lunch (2012)	0.07	0.26	0.00	1.00	0
School lunch (2013)	0.07	0.26	0.00	1.00	497
Tutoring (2012)	0.01	0.12	0.00	1.00	0
Tutoring (2013)	0.01	0.11	0.00	1.00	497
Covariates (measured in 2010 or time-constant)					
Female	0.49	0.50	0.00	1.00	0
Migration background	0.34	0.47	0.00	1.00	0
Attends school	0.58	0.49	0.00	1.00	0
Primary school	0.32	0.47	0.00	1.00	0
Lower secondary school track	0.05	0.21	0.00	1.00	0
Middle secondary school track	0.08	0.27	0.00	1.00	0
Upper secondary school track	0.08	0.27	0.00	1.00	0
Comprehensive secondary school	0.03	0.16	0.00	1.00	0
Other school	0.45	0.50	0.00	1.00	0
School grade	2.35	2.59	0.00	13.00	0
Full-day school	0.13	0.34	0.00	1.00	0
Birth year	2002.04	3.55	1996.00	2007.00	0
Nb of children in household	2.55	1.09	1.00	8.00	0
Single-parent household	0.27	0.44	0.00	1.00	0
One parent with Abitur	0.30	0.46	0.00	1.00	0
Parent with university degree	0.26	0.44	0.00	1.00	0
Log monthly net household income	7.78	0.49	5.91	9.77	0
Monthly HH exp: education	23.86	97.37	0.00	3000.00	0
Monthly HH exp: culture	15.57	50.88	0.00	2000.00	0
Monthly HH exp: leisure	57.05	107.31	0.00	2000.00	0
Rural area (population $<$ 20k)	0.47	0.50	0.00	1.00	0
Large city (population $>$ 100k)	0.25	0.43	0.00	1.00	0
Federal state	7.44	3.73	1.00	16.00	0
FiD subsample	64.00	1.55	61.00	65.00	0
Work hours	12.72	15.34	0.00	75.00	0
Mother not working	0.47	0.50	0.00	1.00	0
Age at birth	29.57	5.64	15.00	56.00	0
Missing: parental education	0.04	0.19	0.00	1.00	0
Missing: household expenditure	0.04	0.20	0.00	1.00	0
Missing: mother's work hours	0.01	0.11	0.00	1.00	0

Source: FiD v4, own calculations.

C.3 Propensity score matching

The predicted probability to be eligible is called the propensity score. For each observation in the treatment group (being eligible), I perform radius matching with a caliper of 2%, which means that I match all individuals from the control group, who have the same predicted probability to be treated (plus/minus 2%). In the matching step I impose common support, which means that treatment observations are dropped, if they have a propensity score higher than the highest of the control group. For each treated, all control observations receive identical weights that sum to 1. Next, I estimate the first-differences specification described in equation (4.3), in which control observations are weighted by the weights resulting from propensity score matching. Standard errors are obtained by bootstrapping with 999 replications and are clustered at the household level.

Table C.6 – *Selection into eligibility for educational package: Probit model (2012, full model)*

	All children		**Schoolchildren**	
	Marginal effect	*P-value in %*	*Marginal effect*	*P-value in %*
Individual-level covariates				
Female	0.02	19	0.02	17
Migration background	0.05	1	0.06	3
Attends school	0.08	15		
School type (Reference: primary school)				
Lower secondary school	-0.00	98	-0.01	88
Medium secondary school	-0.02	63	-0.03	53
Upper secondary school	-0.05	16	-0.06	13
Comprehensive school	0.00	97	0.01	90
Other	0.06	9	0.06	12
School grade	-0.01	34	-0.01	28
Full-day school	0.04	6	0.04	9
Birth year (Reference: 1996)				
1997	0.01	73	0.01	84
1998	0.00	96	0.01	86
1999	-0.05	32	-0.06	28
2000	-0.02	75	-0.03	64
2001	-0.06	38	-0.07	34
2002	-0.05	53	-0.07	43
2003	-0.03	70	-0.05	57
2004	-0.02	84		
2005	-0.09	24		
2006	-0.05	59		
2007	-0.04	61		
Household-level covariates				
Nb of children in household	0.12	0	0.12	0
Single-parent household	0.12	0	0.13	0
One parent with Abitur	-0.11	0	-0.14	0

Table C.6 to be continued.

Table C.6 continued.

	Marginal effect	*P-value in %*	*Marginal effect*	*P-value in %*
Parent with university degree	-0.07	2	-0.09	4
Log monthly net household income	-0.36	0	-0.35	0
Monthly HH exp: education	-0.00	42	-0.00	50
Monthly HH exp: culture	-0.00	55	-0.00	75
Monthly HH exp: leisure	-0.00	6	-0.00	11
Rural area (population $<$ 20k)	-0.06	1	-0.08	1
Large city (population $>$ 100k)	0.01	59	0.01	74
Federal state (Reference: Schleswig-Holstein)				
Hamburg	-0.23	0	-0.22	1
Lower Saxony	-0.14	4	-0.09	27
Breme	-0.04	77	0.15	37
North Rhine-Westphalia	-0.15	2	-0.09	24
Hesse	-0.20	0	-0.13	12
Rhineland-Palatina	-0.22	0	-0.22	1
Baden Wuerttemberg	-0.19	1	-0.14	7
Bavaria	-0.24	0	-0.19	1
Saarland	-0.25	1	-0.15	25
Berlin	-0.06	52	0.04	72
Brandenburg	-0.10	22	-0.05	56
Mecklenburg-Vorpommern	0.07	57	0.26	7
Saxony	-0.09	22	-0.02	83
Saxony-Anhalt	-0.03	75	0.08	40
Thuringia	-0.01	91	0.15	20
FiD sample (Reference: Birth cohort 2007)				
Birth cohort 2008	-0.04	30	-0.09	11
Birth cohort 2009	-0.03	41	-0.05	37
Birth cohort 2010	0.03	50	0.03	62
Screening (sampled 2010)	0.01	65	0.01	89
Missing: parental education	-0.08	10	-0.11	11
Missing: household expenditure	-0.02	67	-0.09	11
Mother-related covariates				
Mother's work hours (in 10h/week)	-0.02	5	-0.00	3
Mother not working	0.04	17	0.05	23
Age at birth	-0.00	3	-0.00	31
Missing: mother's work hours	-0.26	0	-0.24	1
Number of observations	4221		2432	
Efron's R-square (in %)	40.7		39.7	

Source: FiD v4, own calculations. Marginal effects are calculated at the mean for each covariate.

Table C.7 – *Selection into eligibility for educational package: Probit model (2013, full model)*

	All children		**Schoolchildren**	
	Marginal effect	*P-value in %*	*Marginal effect*	*P-value in %*
Individual-level covariates				
Female	-0.02	22	-0.00	96
Migration background	0.07	0	0.09	0
Attends school	0.08	27		
School type (Reference: primary school)				
Lower secondary school	-0.02	63	-0.03	53
Medium secondary school	-0.06	17	-0.08	12
Upper secondary school	-0.11	2	-0.12	2
Comprehensive school	-0.06	29	-0.05	38
Other	0.04	43	0.05	38
School grade	-0.02	9	-0.04	1
Full-day school	0.00	89	0.01	79
Birth year (Reference: 1997)				
1998	0.02	74	0.01	86
1999	-0.07	21	-0.11	10
2000	-0.09	20	-0.16	5
2001	-0.10	21	-0.18	5
2002	-0.16	7	-0.25	1
2003	-0.17	6	-0.27	1
2004	-0.15	10		
2005	-0.19	4		
2006	-0.16	9		
2007	-0.17	8		
Household-level covariates				
Nb of children in household	0.10	0	0.10	0
Single-parent household	0.09	0	0.11	0
One parent with Abitur	-0.13	0	-0.17	0
Parent with university degree	-0.02	61	0.02	68
Log monthly net household income	-0.35	0	-0.31	0
Monthly HH exp: education	-0.00	34	-0.00	34
Monthly HH exp: culture	0.00	53	0.00	94
Monthly HH exp: leisure	-0.00	23	-0.00	34
Rural area (population < 20k)	-0.06	2	-0.08	3
Large city (population > 100k)	0.04	21	0.01	71
Federal state (Reference: Schleswig-Holstein)				
Hamburg	-0.19	1	-0.21	0
Lower Saxony	-0.08	18	-0.03	73
Breme	-0.02	84	0.18	33
North Rhine-Westphalia	-0.09	12	-0.03	71
Hesse	-0.13	4	-0.09	23
Rhineland-Palatina	-0.14	3	-0.14	9
Baden Wuerttemberg	-0.11	8	-0.07	38
Bavaria	-0.18	0	-0.15	3
Saarland	-0.07	63	0.12	56
Berlin	-0.02	80	0.03	76
Brandenburg	0.03	66	0.10	29
Mecklenburg-Vorpommern	0.10	37	0.14	26
Saxony	-0.02	77	0.04	70
Saxony-Anhalt	-0.03	74	0.05	61
Thuringia	-0.01	91	0.14	20

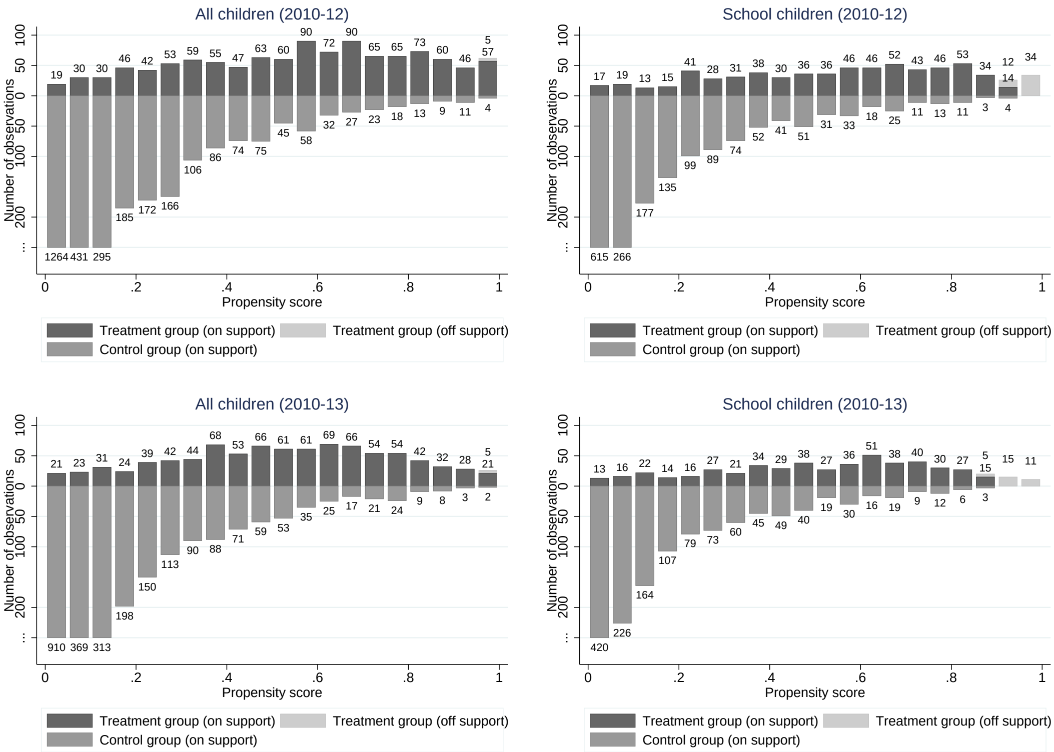
Table C.7 to be continued.

Table C.7 continued.

	Marginal effect	*P-value in %*	*Marginal effect*	*P-value in %*
FiD sample (Reference: Birth cohort 2007)				
Birth cohort 2008	-0.00	96	-0.05	47
Birth cohort 2009	0.02	65	0.02	80
Birth cohort 2010	0.07	14	0.13	13
Screening (sampled 2010)	0.00	93	-0.02	72
Missing: parental education	-0.17	2	-0.21	3
Missing: household expenditure	-0.03	60	-0.07	29
Mother-related covariates				
Mother's work hours (in 10h/week)	-0.04	0	-0.01	0
Mother not working	-0.01	79	-0.02	63
Age at birth	-0.00	2	-0.00	9
Missing: mother's work hours	-0.25	1	-0.29	1
Number of observations	3462		1902	
Efron's R-square (in %)	37.0		35.7	

Source: FiD v4, own calculations. Marginal effects are calculated at the mean for each covariate.

Figure C.1 – *Common support graphs*



Source: FiD v4, own calculations.

Table C.8 – *Balancing of covariates after propensity score matching (Sample 2010-12)*

	All children		**Schoolchildren**	
	Difference	*P-value in %*	*Difference*	*P-value in %*
Individual-level covariates				
Female	-0.02	42	-0.02	56
Migration background	0.01	50	0.02	55
Attends school	-0.01	63	0.00	.
Primary school	-0.01	49	0.01	84
Lower secondary school track	-0.00	76	-0.00	84
Middle secondary school track	0.00	96	-0.01	54
Upper secondary school track	-0.00	94	0.01	71
Comprehensive secondary school	-0.00	85	-0.00	94
Other school	0.02	37	0.01	71
School grade	-0.03	80	-0.03	78
Full-day school	-0.01	66	-0.02	35
Birth year	0.06	69	0.07	57
Household-level covariates				
Nb of children in household	0.08	18	-0.06	44
Single-parent household	0.01	49	0.03	32
One parent with Abitur	0.01	63	-0.00	77
Parent with university degree	-0.01	37	-0.02	17
Log monthly net household income	0.05	1	0.02	36
Monthly HH exp: education	-2.94	16	-4.38	12
Monthly HH exp: culture	0.96	17	0.18	87
Monthly HH exp: leisure	-0.30	86	-2.63	29
Rural area (population $<$ 20k)	-0.01	66	-0.03	29
Large city (population $>$ 100k)	0.01	74	0.03	24
Federal state	0.20	25	0.05	85
FiD subsample	-0.07	25	0.05	48
Missing: parental education	0.00	66	0.00	86
Missing: household expenditure	-0.02	4	-0.03	4
Mother-related covariates				
Work hours	-1.08	5	-1.34	9
Mother not working	0.02	22	0.02	42
Age at birth	0.15	56	0.02	96
Missing: mother's work hours	0.00	84	0.00	66
Number of observations	4216		2386	

Source: FiD v4, own calculations.

Table C.9 – *Balancing of covariates after propensity score matching (Sample 2010-13)*

	All children		Schoolchildren	
	Difference	*P-value in %*	*Difference*	*P-value in %*
Individual-level covariates				
Female	-0.01	60	-0.03	42
Migration background	-0.03	27	-0.00	93
Attends school	-0.01	73	0.00	.
Primary school	0.00	86	-0.01	70
Lower secondary school track	-0.00	72	0.00	97
Middle secondary school track	-0.00	86	0.01	69
Upper secondary school track	-0.01	41	-0.01	71
Comprehensive secondary school	0.01	44	0.01	45
Other school	0.00	93	-0.00	94
School grade	-0.05	66	0.02	84
Full-day school	-0.01	58	0.01	71
Birth year	0.04	81	-0.03	83
Household-level covariates				
Nb of children in household	0.06	35	-0.14	10
Single-parent household	0.01	80	0.03	38
One parent with Abitur	-0.02	19	-0.02	36
Parent with university degree	-0.02	20	-0.02	28
Log monthly net household income	0.04	8	-0.01	80
Monthly HH exp: education	-2.66	18	-2.55	35
Monthly HH exp: culture	0.30	71	-0.78	44
Monthly HH exp: leisure	-1.94	35	-4.07	20
Rural area (population $<$ 20k)	0.03	20	-0.01	80
Large city (population $>$ 100k)	-0.02	32	-0.00	96
Federal state	-0.04	84	0.13	61
FiD subsample	-0.09	18	0.07	35
Missing: parental education	-0.01	14	-0.01	28
Missing: household expenditure	-0.02	2	-0.02	11
Mother-related covariates				
Work hours	-0.39	49	0.01	99
Mother not working	0.00	82	-0.01	77
Age at birth	-0.17	55	0.17	67
Missing: mother's work hours	-0.00	24	-0.00	76
Number of observations	3457		1871	

Source: FiD v4, own calculations.

C.4 Further outcomes

C.1 Results for all years combined

Table C.10 – *Regression results (all years)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Regression 1: Fixed effects model					
Treatment effect 2012	-0.01	-0.03	-0.02	-0.03	-0.02
	(66)	(29)	(43)	(19)	(33)
Treatment effect 2013	-0.00	0.01	-0.00	-0.02	-0.01
	(95)	(78)	(89)	(37)	(83)
N	5613	5613	5613	10371	10371
Regression 2: Fixed effects model, subgroup eligible 2010					
Treatment effect 2012	0.00	-0.01	-0.01	0.01	0.03
	(90)	(79)	(75)	(69)	(37)
Treatment effect 2013	0.01	0.02	-0.00	0.01	0.03
	(86)	(67)	(100)	(62)	(38)
N	1968	1968	1968	3540	3540

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses.

C.2 Heterogeneous results 2013

Table C.11 – *Heterogeneous results (treatment: eligibility 2013)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
HH income < median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	-0.05	-0.04	-0.00	-0.02	0.06
	(31)	(41)	(93)	(57)	(13)
N	516	516	516	879	879
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.10	-0.09	0.04	-0.02	0.08
	(26)	(28)	(64)	(79)	(19)
N	506	506	506	845	845
HH income > median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	0.07	0.10**	0.07	-0.00	0.02
	(14)	(4)	(14)	(97)	(67)
N	1355	1355	1355	2578	2578
Regression 2: First differences model & propensity score matching					
Treatment effect	0.10	0.09	-0.04	0.04	-0.04
	(11)	(13)	(57)	(40)	(48)
N	1325	1325	1325	2510	2510

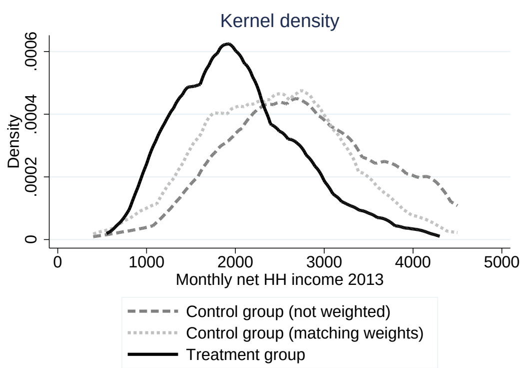
Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Table C.12 – *Heterogeneous results (treatment: eligibility 2013)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Both parents present (2013)					
Regression 1: First differences model					
Treatment effect	0.10**	0.11**	0.09*	0.02	0.04
	(3)	(1)	(5)	(59)	(22)
N	1239	1239	1239	2534	2534
Regression 2: First differences model & propensity score matching					
Treatment effect	0.07	0.06	0.00	0.02	-0.01
	(22)	(28)	(97)	(56)	(78)
N	1218	1218	1218	2486	2486
Single parent household (2013)					
Regression 1: First differences model					
Treatment effect	-0.05	-0.03	-0.04	-0.03	0.01
	(29)	(55)	(38)	(50)	(73)
N	631	631	631	922	922
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.06	-0.05	0.01	-0.05	-0.00
	(58)	(66)	(90)	(59)	(99)
N	604	604	604	887	887

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Figure C.2 – *Density in treatment and control group for local polynomial graph (Figure 4.5)*



Source: FiD v4, own calculations. Epanechnikov kernel with the optimal bandwidth.

Table C.13 – *Heterogeneous results (treatment: eligibility 2013)*

	Schoolchildren			**All children**	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Exp. f. culture < median (among eligible 2013)					
REGRESSION 1: FIRST DIFFERENCES MODEL					
Treatment effect	-0.02	-0.02	-0.04	-0.04	-0.04
	(77)	(74)	(46)	(31)	(30)
N	509	509	509	1000	1000
REGRESSION 2: FIRST DIFFERENCES MODEL & PROPENSITY SCORE MATCHING					
Treatment effect	0.02	0.02	-0.06	-0.01	-0.02
	(87)	(85)	(55)	(79)	(76)
N	498	498	498	967	967
Exp. f. culture > median (among eligible 2013)					
REGRESSION 1: FIRST DIFFERENCES MODEL					
Treatment effect	0.03	0.05	0.03	-0.02	0.01
	(49)	(20)	(40)	(57)	(87)
N	1362	1362	1362	2457	2457
REGRESSION 2: FIRST DIFFERENCES MODEL & PROPENSITY SCORE MATCHING					
Treatment effect	-0.05	-0.02	-0.05	-0.01	0.04
	(48)	(71)	(38)	(84)	(38)
N	1363	1363	1363	2458	2458

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Table C.14 – *Heterogeneous results (treatment: eligibility 2013)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Male					
Regression 1: First differences model					
Treatment effect	0.01	0.03	0.03	-0.04	0.00
	(81)	(43)	(51)	(21)	(91)
N	936	936	936	1776	1776
Regression 2: First differences model & propensity score matching					
Treatment effect	0.01	-0.01	0.03	0.01	0.09
	(92)	(90)	(66)	(82)	(11)
N	924	924	924	1764	1764
Female					
Regression 1: First differences model					
Treatment effect	0.02	0.03	-0.01	-0.01	-0.01
	(59)	(55)	(89)	(78)	(79)
N	935	935	935	1681	1681
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.05	-0.04	-0.10	-0.04	-0.08
	(55)	(59)	(24)	(51)	(20)
N	936	936	936	1669	1669

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

C.3 Main results for 2012

Table C.15 – *Regression results (treatment: eligibility 2012)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Regression 1: First differences model					
Treatment effect	0.00	-0.01	-0.00	-0.03	-0.03
	(90)	(76)	(87)	(20)	(22)
N	2386	2386	2386	4216	4216
Regression 2: First differences model, subgroup eligible 2010					
Treatment effect	0.01	0.02	0.01	-0.00	0.01
	(82)	(72)	(77)	(93)	(71)
N	804	804	804	1399	1399
Regression 3: First differences model & propensity score matching					
Treatment effect	-0.04	-0.04	-0.00	-0.05	-0.02
	(32)	(31)	(98)	(16)	(48)
N	2386	2386	2386	4216	4216
Regression 4: First differences model & propensity score matching, subgroup eligible 2010					
Treatment effect	0.00	0.00	0.14*	-0.03	0.01
	(95)	(100)	(7)	(49)	(84)
N	785	785	785	1340	1340

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

C.4 Heterogeneous results for 2012

Table C.16 – *Heterogeneous results (treatment: eligibility 2012)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
HH income < median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	0.02	0.03	0.05	0.01	0.03
	(73)	(61)	(31)	(88)	(43)
N	565	565	565	906	906
Regression 2: First differences model & propensity score matching					
Treatment effect	0.09	0.07	0.17*	0.07	0.12*
	(30)	(46)	(10)	(27)	(9)
N	496	496	496	839	839
HH income > median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	0.01	-0.01	0.00	-0.03	-0.01
	(73)	(75)	(91)	(37)	(72)
N	1821	1821	1821	3310	3310
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.04	-0.06	-0.08	-0.06	-0.08*
	(46)	(29)	(21)	(12)	(7)
N	1771	1771	1771	3203	3203

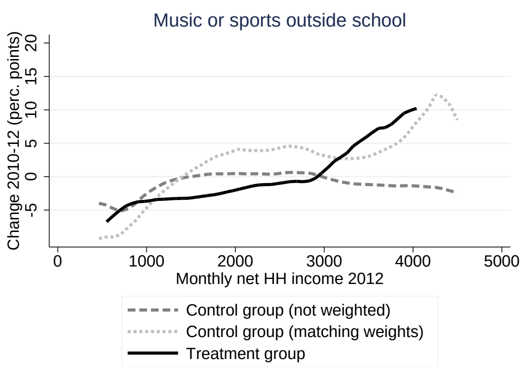
Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Table C.17 – *Heterogeneous results (treatment: eligibility 2012)*

	Schoolchildren			All children	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Both parents present (2013)					
Regression 1: First differences model					
Treatment effect	0.05	0.03	0.03	-0.00	-0.01
	(17)	(48)	(52)	(92)	(81)
N	1572	1572	1572	3079	3079
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.01	-0.04	-0.07	-0.05	-0.05
	(84)	(46)	(27)	(22)	(21)
N	1519	1519	1519	2984	2984
Single parent household (2013)					
Regression 1: First differences model					
Treatment effect	-0.06	-0.05	-0.03	-0.04	0.02
	(15)	(23)	(54)	(29)	(65)
N	812	812	812	1135	1135
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.02	-0.01	0.08	0.06	0.08
	(79)	(92)	(38)	(31)	(20)
N	780	780	780	1051	1051

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Figure C.3 – *Changes in music/sports participation by household income (local polynomial estimations)*



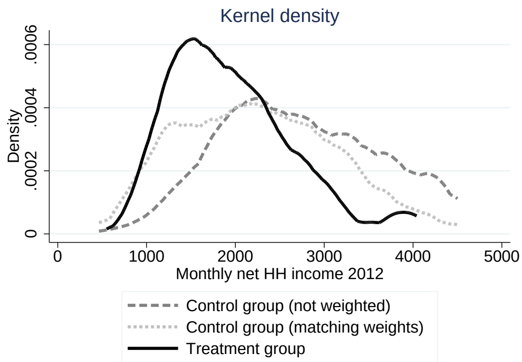
Source: FiD v4, own calculations. Figure C.4 shows the kernel density for the entire distribution of household income. Epanechnikov kernel with a bandwidth of 500 (household income). Children with a household income larger than 4500 Euros were deleted from the sample, because for them, the graph would become very noisy due to the small density of very rich families.

Table C.18 – *Heterogeneous results (treatment: eligibility 2012)*

	Schoolchildren			**All children**	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Exp. f. culture < median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	-0.04	-0.05	-0.01	-0.03	-0.04
	(47)	(32)	(77)	(33)	(31)
N	635	635	635	1255	1255
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.05	-0.12	-0.03	-0.03	-0.03
	(53)	(20)	(75)	(60)	(58)
N	622	622	622	1195	1195
Exp. f. culture > median (among eligible 2013)					
Regression 1: First differences model					
Treatment effect	0.01	0.00	-0.01	-0.03	-0.04
	(71)	(91)	(84)	(31)	(14)
N	1751	1751	1751	2961	2961
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.07	-0.07	-0.02	-0.06	-0.04
	(18)	(20)	(68)	(17)	(36)
N	1750	1750	1750	2947	2947

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Figure C.4 – *Density in treatment and control group for local polynomial graph (Figure C.3)*



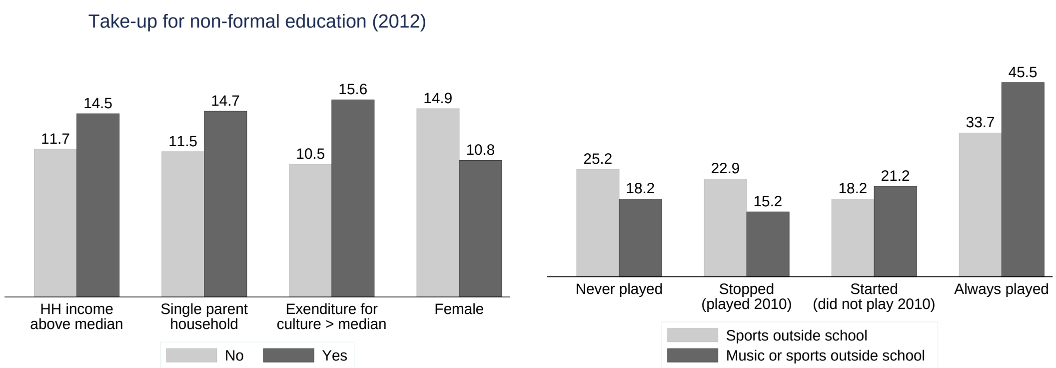
Source: FiD v4, own calculations. Epanechnikov kernel with the optimal bandwidth.

Table C.19 – *Heterogeneous results (treatment: eligibility 2012)*

	Schoolchildren			**All children**	
	Sports outside (1)	*Music/sports outside* (2)	*Music/Sports at/outside* (3)	*Sports outside* (4)	*Music/sports at/outside* (5)
Male					
Regression 1: First differences model					
Treatment effect	-0.03	-0.03	-0.03	-0.04	-0.03
	(47)	(43)	(48)	(15)	(30)
N	1190	1190	1190	2142	2142
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.13*	-0.14**	-0.10	-0.07	-0.03
	(5)	(4)	(21)	(15)	(60)
N	1173	1173	1173	2131	2131
Female					
Regression 1: First differences model					
Treatment effect	0.03	0.01	0.02	-0.01	-0.02
	(37)	(72)	(59)	(68)	(45)
N	1196	1196	1196	2074	2074
Regression 2: First differences model & propensity score matching					
Treatment effect	-0.01	-0.00	-0.01	-0.01	-0.06
	(88)	(99)	(86)	(80)	(26)
N	1198	1198	1198	2046	2046

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Propensity score matching results were obtained with radius matching (caliper 0.02). Standard errors were computed with 999 bootstrap replications.

Figure C.5 – *Heterogeneity in take-up of the subsidy for non-formal education and pattern of activities among beneficiaries of the subsidy (2012)*



Source: FiD v4, own calculations.

C.5 Robustness tests (2013 results)

Table C.20 – *Robustness with respect to estimation method (2013)*

	Music or sports outside of school				
		by household income		*by household composition*	
	All children (1)	< median (2)	> median (3)	both parents (4)	single parent (5)
Caliper changed to 1% (rather than 2%)					
Treatment effect	-0.01	-0.09	0.07	0.06	-0.00
	(79)	(35)	(34)	(40)	(97)
Kernel matching (instead of radius matching)					
Treatment effect	-0.01	-0.11	0.08	0.07	-0.08
	(82)	(23)	(25)	(32)	(52)
Entropy balancing (instead of propensity score matching)					
Treatment effect	-0.00	0.02	0.07	0.07	-0.03
	(94)	(81)	(24)	(19)	(55)
Simple first differences model including survey weights					
Treatment effect	0.04	-0.05	0.11*	0.09	-0.01
	(30)	(46)	(8)	(11)	(80)
First differences model with propensity score matching, including survery weights					
Treatment effect	0.01	-0.07	0.09	0.05	-0.01
	(88)	(41)	(19)	(51)	(91)

Source: FiD v4, own calculations. Significance: * 10% ** 5% *** 1%. P-values (in %) in parentheses. Standard errors of propensity score matching results were computed with 999 bootstrap replications. Results for 2012 can be provided by the author on request.

Appendix to "How a universal music education program affects time use, behavior, and school attitude"

D.1 Data

Construction of data set

Table D.1 – *Detailed description of all variables*

Variable	**Description**	**Type**	**Data source**
Treatment variables			
Treatments for main results			
T1: Jeki share > 50%	Main treatment, see Table 5.2	Binary	Primary schools
T2: Music sc. fees in lowest third	Main treatment, see Table 5.2	Binary	Music schools
Combined treatment	1: T1 = 1 (JeKi share >50%) 2: T1 = 0 and T2 = 1 (No JeKi, music school fees in lowest third) 3: T1 = 0 and T2 = 0 (No JeKi, music school fees above lowest third)	Categories	Music schools Primary schools
Treatments for robustness checks			
JeKi share > 5%	T1, but cutoff at 5%	Binary	Primary schools
JeKi share > 50% (no radius limit)	T1, but includes schools farther than 5km	Binary	Primary schools
JeKi share (continuous)	T1, but not transformed into binary	Continuous	Primary schools

Table D.1 to be continued.

Table D.1 continued.

Variable	Description	Type	Data source
JeKi share (weighted by distance)	T1, but not transformed into binary, each school receives weight, which linearly decreases with distance to child's residence	Continuous	Primary schools
JeKi at closest school	Yes/no in year of school enrollment	Binary	Primary schools
JeKi at 1 of 5 closest schools	Yes/no in year of school enrollment	Binary	Primary schools
Fees below median	T2, average replaced by median	Binary	Music schools
Music school fees < average	T2, 33rd percentile replaced by average	Binary	Music schools
Fees (in Euros)	T2, without calculating 33rd percentile and transforming into binary	Continuous	Music schools
Fees (in Euros, alt. def.)	Like previous, but alternative calculation of fees: Fee revenues divided by number of classes instead of number of students	Continuous	Music schools
Fees in lowest third (alt. def.)	T2, with alternative calculation of fees	Binary	Music schools
Placebo treatments			
Jeki share > 50% (placebo)	T1, but as if JeKi started 1 year earlier	Binary	Primary schools
Fees in lowest third (placebo)	Random draw of same number of music schools as in T2	Binary	Music schools
OUTCOME VARIABLES			
Music activities			
Plays music	Yes, if "Attends music club at school" or "Music or singing lessons outside school"	Binary	SOEP/FiD[1]
Plays music at school	Yes, if "Attends music club at school"	Binary	SOEP/FiD[1]
Plays music outside school	Yes, if "Music or singing lessons outside school"	Binary	SOEP/FiD[1]
Strengths and Difficulties Questionnaire (SDQ)			
	Average for following items: Child...		
No emotional symptoms	...is often unhappy or dejected *(Reversed)* ...is nervous or clingy in new situations, loses self-confidence easily *(R)* ...has many fears, becomes frightened easily *(R)*	3-point Likert scale (normalized)	SOEP/FiD[2]

Table D.1 to be continued.

Table D.1 continued.

Variable	**Description**	**Type**	**Data source**
No conduct problems	...often has tantrums, has a temper *(R)* ...quarrels a lot with other children, picks on them *(R)*	3-point Likert scale (normalized)	SOEP/FiD[2]
Not hyperactive	...is agitated, hyperactive, cannot sit still *(R)* ...is fidgety *(R)* ...is easily distracted and lacks concentration *(R)* ...finishes tasks, is able to concentrate ...thinks before acting	3-point Likert scale (normalized)	SOEP/FiD[2]
No peer problems	...is a loner, usually plays by herself *(R)* ...is popular with other children ...is often made fun of or picked on by other children *(R)* ...gets along better with adults than with other children *(R)*	3-point Likert scale (normalized)	SOEP/FiD[2]
Pro-social behavior	...is considerate ...likes to share with others (sweets, toys, crayons) ...is helpful if others are hurt, sick or sad ...helps others of his/her own accord (parents, teachers, other children)	3-point Likert scale (normalized)	SOEP/FiD[2]
Attitude toward school			
Likes to go to school	“fully applies” to “does not apply”	4-point Likert scale (normalized)	SOEP/FiD[2]
Likes to study	“fully applies” to “does not apply”	4-point Likert scale (normalized)	SOEP/FiD[2]
Follows lessons well	“fully applies” to “does not apply”	4-point Likert scale (normalized)	SOEP/FiD[2]
Gets along well with classmates	“fully applies” to “does not apply”	4-point Likert scale (normalized)	SOEP/FiD[2]
Gets along well with teacher	“fully applies” to “does not apply”	4-point Likert scale (normalized)	SOEP/FiD[2]

Table D.1 to be continued.

Table D.1 continued.

Variable	Description	Type	Data source
Combined non-cognitive skills scores			
Non-cognitive skills (weighted)	Summary index of all outcomes from SDQ and attitude toward school (see above). Based on Anderson (2008)	Continuous (normalized)	SOEP/FiD[2]
Difficulties (from SDQ)	Average of the 4 "difficulties" variables from the SDQ (see below): No emotional symptoms, no conduct problems, not hyperactive, no peer problems (higher number = fewer problems)	Continuous (normalized)	SOEP/FiD[2]
GROUP VARIABLES FOR HETEROGENEITY ANALYSIS			
Household income >/< median[3]	Median calculated within sample	Binary	SOEP/FiD[4]
University/No university	At least 1 parent has university degree	Binary	SOEP/FiD[5]
High/low education (parents)	At least 1 parent has upper sec. school degree (university entrance certificate)	Binary	SOEP/FiD[5]
Migration background (child)	Yes if direct or indirect migration background	Binary	SOEP/FiD[6]
Gender	Male/female	Binary	SOEP/FiD[6]
CONTROL VARIABLES			
Individual characteristics			
Female	0 "male", 1 "female"	Binary	SOEP/FiD[6]
Migration background (child)	Direct or indirect migration background	Binary	SOEP/FiD[6]
Late school enrollment	Year of school enrolment > Birth year + 6	Binary	SOEP/FiD[6]
Year of school enrollment	Source 1: Directly asked at age 8 Source 2: Directly asked at age 10 Source 3: Legal school entry year calculated from birth month & cutoff date	Categorical (fixed effects)	SOEP/FiD[2,6]
Parental, family and household characteristics			
Age of mother at birth	Birth year child – birth year mother	Continuous	SOEP/FiD[6]
Parent with Abitur	At least 1 parent has upper sec. school degree (university entrance certificate)	Binary	SOEP/FiD[5]
Parent with university degree	At least 1 parent has university degree	Binary	SOEP/FiD[5]
Actual work hours of parents[3]	Average of both parents, rescaled such that 40h = 1	Continuous	SOEP/FiD[5]

Table D.1 to be continued.

Table D.1 continued.

Variable	**Description**	**Type**	**Data source**
Only one parent at home[3]	Parent in household if household ID (parent)[7] = household ID (child)	Binary	SOEP/FiD[6]
Nb of children under 16 in HH[3]	Calculated from fieldwork information	Continuous	SOEP/FiD[1]
Rural area[3]	Population in town of residence < 20,000 (fieldwork information)	Binary	SOEP/FiD[9]
Large city[3]	Population in town of residence > 100,000 (fieldwork information)	Binary	SOEP/FiD[9]
Log monthly net HH income[3]	If missing: 1st imputed value	Continuous	SOEP/FiD[4]
Missing: Work hours (parents)	"Actual work hours of parents" missing[11]	Binary	–
Missing: Mother's age	"Age of mother at birth" missing[11]	Binary	–
Missing: Migration background	"Migr. background (child)" missing[11]	Binary	–
County-level characteristics (merged for the county of residence[9])			
Share of employed (%)	2012	Continuous	INKAR[8]
Yearly GDP/person (1000 euros)	2012	Continuous	INKAR[8]
Av. monthly HH income (euros)	2012, per adult member of the household	Continuous	INKAR[8]
Share highly qualified workers (%)	2012	Continuous	INKAR[8]
School age population share (%)	2012	Continuous	INKAR[8]
Pupils share (%)	2012	Continuous	INKAR[8]
Share foreign pupils (%)	2012	Continuous	INKAR[8]
Share students with Abitur (%)	2012	Continuous	INKAR[8]
Share students without degree (%)	2012	Continuous	INKAR[8]
Share children in daycare (%)	2012	Continuous	INKAR[8]
Share village population (%)	2012	Continuous	INKAR[8]
Population within 100km radius	In 1000, 2012	Continuous	INKAR[8]
Youth unemployment rate (%)	2012	Continuous	INKAR[8]
Child poverty rate (%)	2012	Continuous	INKAR[8]
Net incoming commuter share (%)	Workers, 2012	Continuous	INKAR[8]
County identifier[3]	Used to include fixed effects	Categorical	SOEP/FiD[9]
Music school characteristics[10]			
Participation at JeKi	Acts as JeKi cooperating partner	Binary	Music schools
Number of students	Taking classes at music school	Continuous	Music schools
Total revenues (1000 Euros)	Per year	Continuous	Music schools
Cooperations with schools	Number per year	Continuous	Music schools

Table D.1 to be continued.

Table D.1 continued.

Variable	**Description**	**Type**	**Data source**
Share federal state subsidies	Share of total revenues	Continuous	Music schools
Share basic classes	Among total classes given	Continuous	Music schools
Share instrument lessons	Among total classes given	Continuous	Music schools
Share ensemble lessons	Among total classes given	Continuous	Music schools
Missing: Music school information	Any music school covariate missing[11]	Binary	–

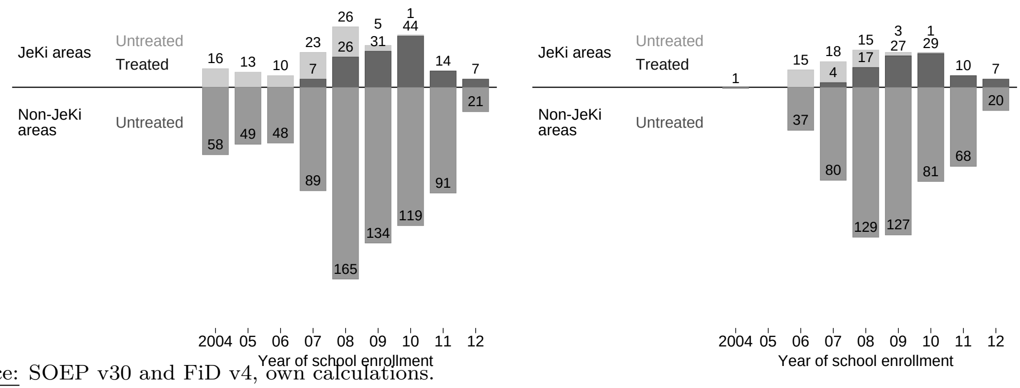
[1] The KIDLONG file contains all child-related information, which was survey in the SOEP or FiD household questionnaire.
[2] Mother-child questionnaires answered by parent in the year the child turns 6, 8 and 10.
[3] Time-varying covariates were measured in the first year they were available between age 7 and 11. They were measured at age 7 (63% of the sample), 8 (13%), 9 (12%) or 10 (12%).
[4] The HGEN file contains household-related information generated from answers to the household questionnaire.
[5] The PGEN file contains person-related information generated from answers to current and previous individual questionnaires.
[6] The PPFAD file contains person-related meta data available for SOEP and FiD.
[7] The parent identifiers were taken from the KIDLONG file.
[8] The INKAR file contains detailed statistics on various regional levels and is provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR, 2015).
[9] Regional information on the place of residence of SOEP and FiD households is stored on a secure server on site at DIW Berlin.
[10] These numbers are either for year 2008 (for children whose covariates were measured in 2010 or before) or 2012 (for children whose covariates were measured in 2011 or later).
[11] Missing covariates were replaced by 0 (binary variables) or their mean (continuous variables). Missing indicators were included as additional control variables. The results are robust to only examining the subsample without missing covariates.

Table D.2 – *Sample construction*

Sample...	**Remaining observations**	
...including all children observed at least once between age 8 and 10 (in SOEP/FiD, includes the entire Federal Republic of Germany)	11302	
...with complete information on federal state	11086	
...restricted to North Rhine-Westphalia	2445	
	Music activities	*Non-cognitive skills*
Outcomes		
...which received questions on the outcomes[1]	1018	736
...with complete information on all outcomes (no missing answers)	1018	697
Treatments		
...with complete information on both treatments[2]	997	689
Covariates		
...with complete information on individual covariates[3]	997	689
...with complete information on regional covariates	997	689
...with complete information on music school covariates[4]	997	689
Final sample	**997**	**689**
Valid observations in both outcome groups	548	
Valid observations in at least one of the two outcome groups	1138	

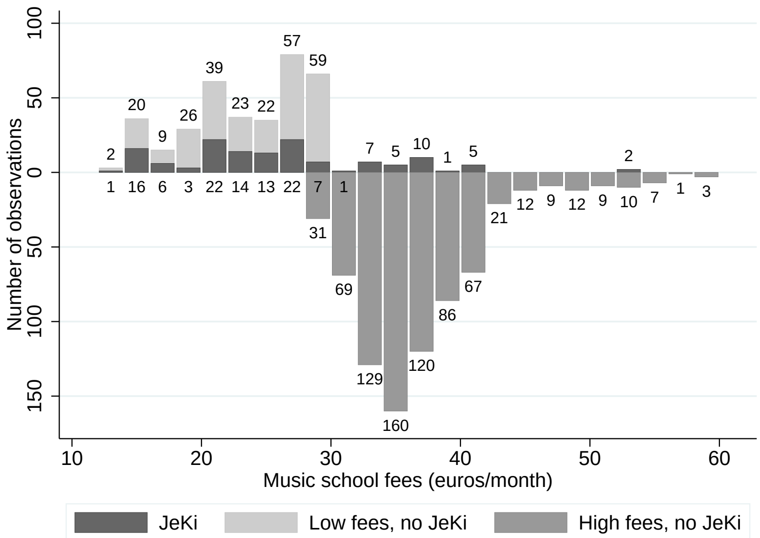
[1]To receive the questions on music activities, parents had to answer the SOEP or FiD household questionnaire in 2006, 2008, 2010, 2011, 2012 or 2013. To receive questions on the non-cognitive skills studied in this paper (SDQ and school attitude), parents had to answer the FiD mother-child questionnaire at age 8 (administered every year between 2010 and 2013) or the SOEP or FiD mother-child questionnaire at age 10 (administered every year between 2010 and 2013 for the FiD sample, and in 2012 and 2013 for the SOEP sample). In other years, the relevant questions were not part of the survey, the drop in observations compared to the previous line is therefore exogenous.
[2]The treatment status is not available for some individuals due to missing information on the number of students at some primary schools (needed to calculate the JeKi share), and the fee revenues at some music schools (needed to calculate the average fees).
[3]The only covariates with missing values were the parents' work hours (7% missing), the mother's age (2% missing) and the child's migration background (1% missing). To prevent a further reduction of the sample size, I replaced missing values by 0 for binary and by the sample mean for continuous variables and added missing indicators in all regressions for these variables.
[4]These covariates are only included in the estimations on below-average music school fees. 3% of the observations with missing values were replaced by the sample mean. A missing indicator was included in all regressions.

Figure D.1 – *Number of observations in treatment and control group for the two outcome group subsamples*



Source: SOEP v30 and FiD v4, own calculations.

Figure D.2 – *Distribution of children by treatment status and music school fee level*



Source: SOEP v30 and FiD v4, own calculations. The figure shows the distribution of observations within 2 Euro-windows of monthly music school fees. The bars with three different shadings represent the three groups of the combined treatment. The two main treatments (JeKi share above 50% and music school fees below 33rd percentile) are visible as well. In the first, JeKi students (darkest gray) are simply compared to everyone else. In the second, I contrast areas with low and high music school fees (left and right part of the graph), irrespective of whether JeKi exists. Average music school fees in North Rhine-Westphalia vary between 12 and 60 Euros per month. The 33rd percentile is at 29.4 Euros, which explains why both bars, representing low and high fee levels, are visible in the interval ranging from 28 Euros to 30 Euros.

Summary statistics

Table D.3 – *Summary statistics of all treatment variables*

	Mean	Standard deviation	Minimum	Maximum	Sample size
Main treatment definitions					
Jeki share > 50%	0.12	0.32	0.00	1.00	1138
Music sc. fees in lowest third	0.32	0.46	0.00	1.00	1138
Alternative definitions of JeKi					
Jeki share > 5%	0.19	0.39	0.00	1.00	1138
Jeki share > 50% (no radius limit)	0.12	0.32	0.00	1.00	1138
JeKi share (continuous)	0.12	0.28	0.00	1.00	1138
JeKi share (weighted by distance)	0.12	0.28	0.00	1.00	1138
JeKi at closest school	0.11	0.32	0.00	1.00	1138
JeKi at 1 of 5 closest schools	0.19	0.39	0.00	1.00	1138
Alternative definitions for music school fees					
Music school fees < median	0.61	0.49	0.00	1.00	1138
Music school fees < average	0.43	0.49	0.00	1.00	1138
Monthly fees (in euros)	32.56	8.15	13.07	58.35	1138
Monthly fees (in euros, alt. def.)	27.19	12.66	9.84	284.31	1138
Fees in lowest third (alt. def.)	0.32	0.47	0.00	1.00	1138
Placebo treatments					
Jeki share > 50% (placebo)	0.16	0.36	0.00	1.00	1138
Fees in lowest third (placebo)	0.44	0.50	0.00	1.00	1138

Source: SOEP v30 and FiD v4, own calculations.

Table D.4 – *Summary statistics of all outcome variables*

	Mean	**Standard deviation**	**Minimum**	**Maximum**	**Sample size**
Music activities					
Plays music	0.37	0.48	0.00	1.00	997
Non-cognitive skills					
Strengths and Difficulties Questionnaire					
No emotional symptoms	-0.04	1.01	-3.51	0.90	689
No conduct problems	-0.02	0.99	-3.29	0.88	689
Not hyperactive	-0.04	1.00	-2.78	1.29	689
No peer problems	-0.01	1.01	-3.85	0.90	689
Pro-social behavior	0.01	1.00	-3.03	1.08	689
Attitude towards school					
Likes to go to school	-0.04	1.01	-3.72	0.82	689
Likes to study	-0.03	1.00	-2.83	1.05	689
Follows lessons well	-0.04	1.02	-3.75	0.78	689
Gets along well...					
...with classmates	0.05	0.97	-1.39	1.05	689
...with teacher	-0.00	0.99	-2.52	0.60	689
Aggregated non-cognitive skills scores					
Non-cog. skills (aggregate)	-0.02	0.58	-2.15	0.94	689
Difficulties (from SDQ)	-0.04	1.00	-3.90	1.40	689

Source: SOEP v30 and FiD v4, own calculations. Outcomes in the groups “Strengths and Difficulties Questionnaire” and “Attitude toward school” were normalized with mean 0 and a standard deviation of 1 within the control group of the JeKi program treatment (i.e. children living in areas, which never had the JeKi program).

Table D.5 – *Summary statistics of all control variables*

	Mean	Standard deviation	Minimum	Maximum	Sample size
Individual and family characteristics					
Female	0.50	0.50	0.00	1.00	1138
Migration background (child)	0.35	0.48	0.00	1.00	1138
Late school enrollment	0.53	0.50	0.00	1.00	1138
Age of mother at birth	29.8	5.3	16.0	47.0	1138
Parent with Abitur	0.34	0.47	0.00	1.00	1138
Parent with university degree	0.28	0.45	0.00	1.00	1138
Actual work hours of parents	0.57	0.32	0.00	1.75	1138
Only one parent at home	0.15	0.36	0.00	1.00	1138
Nb of children under 16 in HH	2.6	1.1	1.0	8.0	1138
Rural area	0.14	0.35	0.00	1.00	1138
Large city	0.40	0.49	0.00	1.00	1138
Log monthly net HH income	7.9	0.5	6.0	10.5	1138
Missing: Work hours (parents)	0.07	0.26	0.00	1.00	1138
Missing: Mother's age	0.02	0.13	0.00	1.00	1138
Missing: Migration background	0.01	0.10	0.00	1.00	1138
Year of school enrollment	2008	2	2004	2012	1138
County-level characteristics					
Share of employed (%)	52	3	44	57	1138
Yearly GDP/person (1000 Euros)	33	9	21	70	1138
Av. monthly HH income/person (Euros)	1701	158	1359	2177	1138
Share highly qualified workers (%)	6	3	3	18	1138
School age population share (%)	14	1	11	17	1138
Pupils share (%)	12	1	10	14	1138
Share foreign pupils (%)	9	5	2	21	1138
Share students with Abitur (%)	36	5	25	52	1138
Share students without degree (%)	5	1	3	9	1138
Share children in daycare (%)	18	3	12	26	1138
Share village population (%)	6	12	0	68	1138
Population within 100km radius	960	523	217	1916	1138
Youth unemployment rate (%)	19	2	15	23	1138
Child poverty rate (%)	18	7	7	35	1138
Net incoming commuter share (%)	-3	20	-46	43	1138
County identifier	5509	299	5111	5978	1138
Music school characteristics					
Participation at JeKi	0.29	0.45	0.00	1.00	1138
Number of students	2169	2094	112	11347	1138
Total revenues (1000 Euros)	1891	1798	41	7294	1138
Cooperations with schools	16.8	19.0	0.0	91.0	1138
Share federal state subsidies	0.01	0.01	0.00	0.10	1138
Share basic classes	0.19	0.09	0.00	0.53	1138
Share instrument lessons	0.62	0.13	0.00	0.94	1138
Share ensemble lessons	0.14	0.08	0.00	0.94	1138
Missing: Music school information	0.03	0.16	0.00	1.00	1138

Source: SOEP v30 and FiD v4, own calculations.

Table D.6 – *Outcome variables by treatment status (in %)*

	Control group	Treatment group	Difference	p-value
Music activities				
Plays music	0.36	0.47	0.11	1
Alternative treatment: Music school fees in lowest third				
Plays music	0.36	0.41	0.05	24
Non-cognitive skills				
Strengths and Difficulties Questionnaire				
No emotional symptoms	-0.02	-0.18	-0.16	13
No conduct problems	-0.04	0.01	0.05	56
Not hyperactive	-0.02	-0.23	-0.21	1
No peer problems	-0.02	-0.05	-0.03	82
Pro-social behavior	-0.01	0.08	0.09	42
Attitude towards school				
Likes to go to school	-0.05	-0.03	0.02	79
Likes to study	-0.05	0.00	0.05	62
Follows lessons well	-0.02	-0.14	-0.11	37
Gets along well...				
...with classmates	0.02	0.15	0.13	18
...with teacher	-0.01	0.11	0.12	16
Aggregated non-cognitive skills scores				
Non-cog. skills (aggregate)	-0.02	-0.04	-0.02	62
Difficulties (from SDQ)	-0.04	-0.17	-0.13	23

Source: SOEP v30 and FiD v4, own calculations. Treatment: Share of JeKi students among all students in the closest five schools is larger than 50% (except in the second row, where the treatment is: Average fees at the closest public music school are in the lowest third among all music schools within the federal state).

Table D.7 – *Individual and county-level background characteristics by treatment group status (in %): Combined treatment*

	Mean for each treatment status			Differences (t-tests)					
	T1	**T2**	**T3**	**T2 - T1**		**T3 - T1**		**T3 - T2**	
	JeKi	*Low fees, no JeKi*	*High fees no JeKi*	*Diff.*	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*
County-level characteristics									
Av. monthly HH income/person (Euros)	1584	1680	1729	95	1	145	0	50	13
Yearly GDP/person (1000 Euros)	31.7	33.1	33.6	1.4	61	1.9	45	0.5	83
Share students with Abitur (%)	37.0	35.1	36.7	-1.8	5	-0.3	81	1.6	7
Share students without degree (%)	5.9	5.2	4.9	-0.7	2	-1.1	0	-0.4	16
Youth unemployment rate (%)	18.5	19.1	19.6	0.6	26	1.1	4	0.5	13
Share foreign pupils (%)	11.6	9.1	8.1	-2.5	2	-3.4	0	-1.0	24
Child poverty rate (%)	25.4	18.7	16.5	-6.7	0	-8.9	0	-2.2	15
Individual-level characteristics									
Monthly net HH income (Euros)	3082	2944	3201	-137	60	120	63	257	12
Parent with Abitur	35.6	27.6	35.3	-7.9	19	-0.3	96	7.6	6
Parent with university degree	24.4	22.2	30.7	-2.3	69	6.3	20	8.5	5
Migration background (child)	31.9	37.0	35.0	5.1	38	3.1	59	-2.0	65
Only one parent at home	17.0	17.5	14.2	0.5	92	-2.8	40	-3.3	26
Nb of children under 16 in HH	2.7	2.6	2.5	-0.1	66	-0.2	8	-0.1	35

Source: SOEP v30 and FiD v4, own calculations.

Table D.8 – *Outcomes by treatment group status: Combined treatment*

	Mean by treatment status			Differences (t-tests)					
	T1	**T2**	**T3**	**T2 - T1**		**T3 - T1**		**T3 - T2**	
	JeKi	*Low fees, no JeKi*	*High fees no JeKi*	*Diff.*	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*
MUSIC ACTIVITIES									
Plays music	0.47	0.37	0.36	-0	9.34	-0.11	1.15	-0	84.28
NON-COGNITIVE SKILLS									
Strengths and Difficulties Questionnaire									
No emotional symptoms	-0.18	-0.19	0.03	-0	94.67	0.22	5.31	0	3.22
No conduct problems	0.01	-0.17	0.01	-0	5.85	-0.00	97.47	0	3.19
Not hyperactive	-0.23	0.05	-0.05	0	0.91	0.18	3.54	-0	33.80
No peer problems	-0.05	-0.08	0.00	-0	85.12	0.05	69.38	0	42.69
Pro-social behavior	0.08	-0.01	-0.00	-0	48.71	-0.08	47.28	0	94.06
Attitude towards school									
Likes to go to school	-0.03	-0.06	-0.05	-0	79.86	-0.02	84.20	0	91.20
Likes to study	0.00	-0.10	-0.03	-0	52.54	-0.03	72.70	0	59.26
Follows lessons well	-0.14	0.03	-0.04	0	27.19	0.10	46.15	-0	51.13
Gets along well...									
...with classmates	0.15	0.07	0.00	-0	58.56	-0.15	12.23	-0	55.60
...with teacher	0.11	-0.07	0.01	-0	9.27	-0.10	28.13	0	48.73
Aggregated non-cognitive skills scores									
Non-cog. skills (aggregate)	-0.04	-0.04	-0.01	0	97.38	0.03	54.19	0	50.85
Difficulties (from SDQ)	-0.17	-0.14	-0.00	0	84.29	0.17	14.89	0	21.82

Source: SOEP v30 and FiD v4, own calculations.

Table D.9 – *Covariates for each estimation sample*

	Music activities		Non-cognitive skills		Combined sample	
	Mean	*Std. dev.*	*Mean*	*Std. dev.*	*Mean*	*Std. dev.*
Individual and family characteristics						
Female	0.50	0.50	0.51	0.50	0.49	0.50
Migration background (child)	0.35	0.48	0.40	0.49	0.41	0.49
Late school enrollment	0.55	0.50	0.34	0.47	0.32	0.47
Age of mother at birth	29.8	5.3	29.7	5.4	29.8	5.4
Parent with Abitur	0.33	0.47	0.34	0.47	0.33	0.47
Parent with university degree	0.29	0.45	0.27	0.45	0.28	0.45
Actual work hours of parents	0.57	0.32	0.58	0.32	0.58	0.31
Only one parent at home	0.13	0.34	0.23	0.42	0.21	0.41
Nb of children under 16 in HH	2.6	1.1	2.8	1.1	2.8	1.1
Rural area	0.14	0.35	0.16	0.37	0.16	0.37
Large city	0.39	0.49	0.40	0.49	0.39	0.49
Log monthly net HH income	7.9	0.5	7.9	0.5	7.9	0.5
Missing: Work hours (parents)	0.08	0.26	0.08	0.28	0.09	0.29
Missing: Mother's age	0.01	0.11	0.02	0.13	0.01	0.10
Missing: Migration background	0.01	0.10	0.01	0.10	0.01	0.10
Year of school enrollment	2008	2	2009	2	2009	1
County-level characteristics						
Share of employed (%)	52	3	52	3	52	3
Yearly GDP/person (1000 Euros)	33	9	34	9	34	9
Av. monthly HH income/person (Euros)	1702	159	1704	155	1708	155
Share highly qualified workers (%)	6	3	6	3	6	3
School age population share (%)	14	1	14	1	14	1
Pupils share (%)	12	1	12	1	12	1
Share foreign pupils (%)	9	4	9	5	9	4
Share students with Abitur (%)	36	5	36	5	36	5
Share students without degree (%)	5	1	5	1	5	1
Share children in daycare (%)	18	3	18	3	18	3
Share village population (%)	6	12	6	12	6	12
Population within 100km radius	962	521	951	535	952	535
Youth unemployment rate (%)	19	2	19	2	19	2
Child poverty rate (%)	18	7	18	7	18	7
Net incoming commuter share (%)	-3	20	-2	20	-2	19
County identifier	5510	301	5506	297	5508	300
Music school characteristics						
Participation at JeKi	0.29	0.45	0.28	0.45	0.28	0.45
Number of students	2163	2111	2352	2287	2389	2360
Total revenues (1000 Euros)	1886	1810	2027	1893	2053	1937
Cooperations with schools	16.8	19.2	18.4	20.4	18.7	21.1
Share federal state subsidies	0.01	0.01	0.01	0.01	0.01	0.01
Share basic classes	0.19	0.09	0.19	0.09	0.18	0.09
Share instrument lessons	0.62	0.13	0.63	0.14	0.63	0.14
Share ensemble lessons	0.14	0.08	0.13	0.08	0.13	0.08
Missing: Music school information	0.02	0.15	0.02	0.15	0.02	0.13
Sample size	997		689		548	

Source: SOEP v30 and FiD v4, own calculations.

D.2 Further results

Further information for main results

Table D.10 – *Number of observations (main results)*

	Full sample		by household income				by gender			
			< median		> median		Male		Female	
	(1)		(2)		(3)		(4)		(5)	
	T	*C*	*T*	*C*	*T*	*C*	*T*	*C*	*T*	*C*
MUSIC ACTIVITIES										
Plays music	129	868	56	446	73	422	61	382	68	486
Alternative treatment: Music school fees in lowest third										
Plays music	314	683	161	341	153	342	149	294	165	389
NON-COGNITIVE SKILLS										
Strengths and Difficulties Questionnaire										
No emotional symptoms	94	595	40	301	54	294	39	288	55	307
No conduct problems	94	595	40	301	54	294	39	288	55	307
Not hyperactive	94	595	40	301	54	294	39	288	55	307
No peer problems	94	595	40	301	54	294	39	288	55	307
Pro-social behavior	94	595	40	301	54	294	39	288	55	307
Attitude towards school										
Likes to go to school	94	595	40	301	54	294	39	288	55	307
Likes to study	94	595	40	301	54	294	39	288	55	307
Follows lessons well	94	595	40	301	54	294	39	288	55	307
Gets along well...										
...with classmates	94	595	40	301	54	294	39	288	55	307
...with teacher	94	595	40	301	54	294	39	288	55	307
Combined non-cognitive skills scores										
Non-cog. skills (aggregate)	94	595	40	301	54	294	39	288	55	307
Difficulties (from SDQ)	94	595	40	301	54	294	39	288	55	307

Source: SOEP v30 and FiD v4, own calculations. Sample sizes for main estimations presented in Table 5.4.

Table D.11 – *R squares (main results), in %*

	Full sample	by household income		by gender	
		< median	**> median**	**Male**	**Female**
	(1)	(2)	(3)	(4)	(5)
MUSIC ACTIVITIES					
Plays music	14.4	13.0	13.0	9.7	14.5
Alternative treatment: Music school fees in lowest third					
Plays music	14.4	14.3	11.4	9.2	17.2
NON-COGNITIVE SKILLS					
Strengths and Difficulties Questionnaire					
No emotional symptoms	2.8	8.9	-0.4	3.0	0.4
No conduct problems	6.6	9.1	2.2	17.2	9.3
Not hyperactive	16.6	15.7	8.6	20.1	15.8
No peer problems	7.0	2.7	9.7	5.6	4.7
Pro-social behavior	5.8	8.2	-2.2	15.5	-1.1
Attitude towards school					
Likes to go to school	8.6	6.0	4.1	10.8	2.8
Likes to study	6.7	6.0	-2.7	4.8	7.0
Follows lessons well	7.3	0.7	8.2	3.8	8.9
Gets along well...					
...with classmates	6.9	3.2	8.9	7.1	4.4
...with teacher	7.3	2.1	10.2	11.2	1.6
Combined non-cognitive skills scores					
Non-cog. skills (aggregate)	15.2	16.1	2.1	19.4	13.6
Difficulties (from SDQ)	12.9	16.5	4.5	21.3	12.9

Source: SOEP v30 and FiD v4, own calculations. R-squares for main estimations presented in Table 5.4.

Table D.12 – *Effects of treatments (main results from Table 5.4, including full set of outcomes for low music school-treatment)*

	Full sample				HH income < median				HH income > median			
	JeKi		**Music sc. fees**		**JeKi**		**Music sc. fees**		**JeKi**		**Music sc. fees**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Diff.	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*	*Diff.*	*p-value*
Music and sports activities												
Plays music	0.14	2	0.10	5	0.06	52	0.17	1	0.20	3	0.03	68
Non-cognitive skills												
Strengths and Difficulties Questionnaire												
No emot. symptoms	-0.04	84	-0.12	33	0.02	91	-0.27	17	-0.21	49	0.09	62
No conduct problems	0.40	3	-0.19	16	0.38	14	-0.21	34	0.59	9	-0.11	61
Not hyperactive	0.15	37	0.09	45	0.28	33	0.22	20	0.32	29	0.09	68
No peer problems	-0.03	86	-0.07	57	-0.23	49	-0.22	30	-0.03	88	0.01	93
Pro-social behavior	0.19	33	0.12	39	0.03	92	0.25	30	0.18	62	0.06	73
Attitude towards school												
Likes to go to school	0.25	14	0.13	41	0.53	5	0.06	77	-0.06	84	0.25	15
Likes to study	0.34	12	0.05	71	0.94	0	-0.09	61	0.02	95	0.19	38
Follows lessons well	0.04	86	0.20	8	0.23	63	0.35	4	0.09	75	0.18	31
Gets along well...												
...with classmates	0.14	46	-0.03	83	0.00	99	-0.22	19	0.09	76	0.18	45
...with teacher	0.27	2	-0.08	62	0.52	7	-0.28	28	-0.02	95	0.10	62
Combined non-cognitive skills scores												
Non-cog. skills (aggregate)	0.17	4	0.02	79	0.29	8	-0.02	82	0.10	44	0.10	39
Difficulties (from SDQ)	0.18	41	-0.10	42	0.19	51	-0.15	42	0.26	42	0.03	89
Covariates included:												
Individual char.	✓		✓		✓		✓		✓		✓	
Regional char.	✓		✓		✓		✓		✓		✓	
Music school char.			✓				✓				✓	
Time fixed effects	✓		✓		✓		✓		✓		✓	
County fixed effects	✓		✓		✓		✓		✓		✓	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how the treatment in the column header affects the outcome denoted in the first column. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" and "Play sports" are binary variables indicating whether the child played a musical instrument/sports at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1. The exact definitions of all outcome variables are given in Table D.1.

Heterogeneity

Table D.13 – *Effects of treatments (OLS estimates): Heterogeneity – JeKi share above 50%*

	University (parents)				**Education (parents)**				**Migration backgr.**			
	No		**Yes**		**Low**		**high**		**Yes**		**No**	
	(1)		(2)		(3)		(4)		(5)		(6)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*
MUSIC AND SPORTS ACTIVITIES												
Plays music	0.18	3	0.10	51	0.20	0	0.03	69	0.23	1	0.14	11
Alternative treatment: Music school fees in lowest third												
Plays music	0.08	11	0.22	8	0.07	28	0.18	9	0.09	40	0.09	14
Sample size	711		286		664		333		350		647	
NON-COGNITIVE SKILLS												
Strengths and Difficulties Questionnaire												
No emot. symptoms	0.14	51	-0.52	24	0.11	56	-0.85	2	-0.36	43	-0.14	50
No conduct problems	0.54	0	-0.15	58	0.44	2	0.20	39	0.38	35	0.25	21
Not hyperactive	0.21	26	0.05	92	-0.02	88	0.22	55	0.22	26	-0.01	98
No peer problems	-0.02	91	-0.16	73	0.17	38	-0.71	1	0.01	98	-0.15	53
Pro-social behavior	0.28	26	-0.46	3	0.28	37	-0.16	62	0.57	10	-0.02	97
Attitude towards school												
Likes to go to school	0.37	14	0.07	76	0.33	25	-0.04	91	-0.00	99	0.58	4
Likes to study	0.40	17	0.21	65	0.25	39	0.14	79	-0.00	99	0.32	35
Follows lessons well	-0.00	99	0.37	20	-0.09	78	0.21	50	-0.32	18	0.21	53
Gets along well...												
...with classmates	0.06	76	0.05	92	0.24	22	0.15	72	-0.09	79	0.27	26
...with teacher	0.30	6	0.10	80	0.24	13	0.23	38	0.07	71	0.47	2
Aggregated non-cognitive skills scores												
Non-cog. skills (aggregate)	0.24	2	-0.04	78	0.19	8	-0.08	53	0.05	77	0.17	26
Difficulties (from SDQ)	0.33	12	-0.27	45	0.24	15	-0.34	20	0.10	76	-0.00	100
Sample size	502		187		457		232		275		414	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. All regressions are estimated using difference-in-differences, and additionally control for individual and regional characteristics, as well as county fixed effects. Effects of the alternative treatment (music school fees in lowest third) are estimated with standard OLS (no difference-in-differences), additionally including music school characteristics as control variables. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Robustness checks

Table D.14 – *Effects of treatments (OLS estimates): Variations in treatment definition – JeKi share above 50%*

	JeKi share >5% (1) *Diff.*	*p-value*	**JeKi sh. >50% (no radius limit)** (2) *Diff.*	*p-value*	**JeKi share (continuous)** (3) *Diff.*	*p-value*	**JeKi sh. (weighted by distance)** (4) *Diff.*	*p-value*	**JeKi in closest school** (5) *Diff.*	*p-value*	**JeKi in 1 of 5 closest schools** (6) *Diff.*	*p-value*
Music and sports activities												
Plays music	0.04	50	0.16	1	0.15	6	0.12	15	0.05	42	0.04	50
Sample size	997		997		997		997		997		997	
Non-cognitive skills												
Strengths and Difficulties Questionnaire												
No emot. symptoms	0.34	35	0.02	93	0.07	83	0.05	88	-0.15	42	0.34	35
No conduct problems	0.05	85	0.37	2	0.38	8	0.44	3	0.38	9	0.05	85
Not hyperactive	-0.12	52	0.09	57	-0.15	51	-0.17	48	-0.08	66	-0.12	52
No peer problems	-0.01	96	-0.05	75	-0.25	15	-0.31	5	-0.08	70	-0.01	96
Pro-social behavior	0.27	32	0.16	41	0.07	73	0.10	60	0.07	60	0.27	32
Attitude towards school												
Likes to go to school	0.16	38	0.24	15	0.08	69	0.07	69	0.27	31	0.16	38
Likes to study	-0.06	85	0.33	13	0.07	80	0.09	76	0.15	62	-0.06	85
Follows lessons well	-0.11	65	0.01	95	-0.14	54	-0.06	78	0.01	98	-0.11	65
Gets along well...												
...with classmates	0.08	72	0.09	65	0.03	91	0.05	84	0.06	69	0.08	72
...with teacher	0.26	9	0.26	3	0.18	24	0.18	26	0.50	0	0.26	9
Aggregated non-cognitive skills scores												
Non-cog. skills (aggregate)	0.07	58	0.15	5	0.02	83	0.03	76	0.10	37	0.07	58
Difficulties (from SDQ)	0.09	71	0.16	44	0.04	88	0.03	90	0.03	88	0.09	71
Sample size	689		689		689		689		689		689	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program (different definitions) affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. All regressions are estimated using difference-in-differences, and additionally control for individual and regional characteristics, as well as county fixed effects. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.15 – *Treatment effects: Variations in treatment definition – Music school fees*

	Fees < median (1)		**Fees < average** (2)		**Fees (in euros)** (3)		**Fees (in euros, alt. def.)** (4)		**Fees < average (alt. def.)** (5)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*	*Diff.*	*p*
Plays music	-0.05	27	0.03	63	-0.00	46	-0.00	38	-0.01	86
Sample size	997		997		997		997		997	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how different definitions of low music school fees affect music activities. All regressions are estimated using OLS, controlling for individual, regional and music school characteristics, as well as time and county fixed effects. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.16 – *Treatment effects (OLS estimates) – Outcomes measured at age 8 or 10 only*

	Outcomes measured at age 8 only (1)		**Outcomes measured at age 10 only** (2)	
	Difference	*p-value*	*Difference*	*p-value*
MUSIC ACTIVITIES				
Plays music	0.07	45	0.08	59
Alternative treatment: Music school fees in lowest third				
Plays music	0.05	32	0.02	72
Sample size	769		554	
NON-COGNITIVE SKILLS				
Strengths and Difficulties Questionnaire				
No emotional symptoms	0.46	55	-0.24	30
No conduct problems	0.70	24	0.65	0
Not hyperactive	0.46	19	-0.02	95
No peer problems	-0.15	73	-0.05	82
Pro-social behavior	-0.10	61	0.05	81
Attitude towards school				
Likes to go to school	-0.22	52	0.28	15
Likes to study	0.52	5	0.17	49
Follows lessons well	0.30	39	-0.16	53
Gets along well...				
...with classmates	0.23	39	0.20	35
...with teacher	0.07	70	0.14	48
Aggregated non-cognitive skills scores				
Non-cog. skills (aggregate)	0.06	42	0.13	15
Difficulties (from SDQ)	0.59	42	0.15	58
Sample size	383		467	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. All regressions are estimated using difference-in-differences, and additionally control for individual and regional characteristics, as well as county fixed effects. Effects of the alternative treatment (music school fees in lowest third) are estimated with standard OLS (no difference-in-differences), additionally including music school characteristics as control variables. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.17 – *Effects of treatments (OLS estimates) – Variations of "Playing music" variable and effects on missing indicators*

	JeKi share > 50% (1)		**Music school fees in lowest third** (2)	
	Difference	*p-value*	*Difference*	*p-value*
VARIATIONS IN THE DEFINITION OF MUSIC ACTIVITIES				
Music at school	0.20	0	0.04	39
Music outside school	-0.04	38	0.04	30
Sample size	997		997	
Music at school (age 8)	0.20	5	0.01	90
Music outside school (age 8)	-0.11	2	0.03	47
Sample size	769		769	
Music at school (age 10)	0.02	83	-0.06	40
Music outside school (age 10)	0.07	48	0.00	95
Sample size	554		554	
EFFECTS ON MISSING COVARIATES				
Missing: Work hours (parents)	0.03	22	0.05	10
Missing: Mother's age	-0.02	9	0.01	31
Missing: Migration background	0.02	20	-0.01	30
Missing: Music school information	-0.05	7	-0.00	88
Sample size	1138		1138	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program and low music school fees affect the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definitions. All regressions are estimated using difference-in-differences (for JeKi), and control for individual and regional characteristics, as well as county fixed effects. Effects of the alternative treatment (music school fees in lowest third) are estimated with standard OLS (no difference-in-differences), additionally including music school characteristics as control variables. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. All outcome variables are binary. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.18 – *Effects of treatments – Different estimation method*

	Propensity score matching (1)		**Entropy balancing** (2)		**Exclude missing covariates** (3)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*
MUSIC AND SPORTS ACTIVITIES						
Plays music	0.15	6	0.11	16	0.13	5
Alternative treatment: Music school fees in lowest third						
Plays music	0.07	9	0.08	5	0.07	16
Sample size	984		997		881	
NON-COGNITIVE SKILLS						
Strengths and Difficulties Questionnaire						
No emotional symptoms	-0.25	26	-0.15	49	-0.28	24
No conduct problems	0.31	16	0.41	5	0.45	2
Not hyperactive	0.06	78	0.15	47	0.21	30
No peer problems	-0.20	38	-0.04	85	-0.16	35
Pro-social behavior	0.04	84	0.12	58	0.24	26
Attitude towards school						
Likes to go to school	0.20	36	0.18	39	0.26	17
Likes to study	0.34	12	0.39	6	0.32	21
Follows lessons well	0.22	34	0.14	49	0.06	79
Gets along well...						
...with classmates	-0.00	98	0.04	84	0.21	29
...with teacher	0.31	12	0.26	19	0.20	10
Aggregated non-cognitive skills scores						
Non-cog. skills (aggregate)	0.11	40	0.16	20	0.15	12
Difficulties (from SDQ)	-0.01	98	0.15	50	0.10	66
Sample size	675		689		603	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. Regressions are estimated with different estimation methods as specified in the column headers. The first uses propensity score matching (radius matching with a caliper of 2%), the second applies entropy balancing (Hainmueller, 2012; Hainmueller and Xu, 2013), and the third excludes all observations with at least one missing covariate. In order to ensure common support, estimations with propensity score matching only use a subset of the full list of covariates from Table D.1: gender, migration background (child), late school enrollment, parent with Abitur, parent with university degree, actual work hours of parents, only one parent at home, missing: Work hours (parents), missing: Migration background, school age population share (%), pupils share (%), share foreign pupils (%), share children in daycare (%). Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. “Plays music” is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.19 – *Effects of treatments (OLS estimates) – Different specifications*

	Including survey weights (1)		**Excluding covariates** (2)		**OLS without DiD** (3)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*
MUSIC AND SPORTS ACTIVITIES						
Plays music	0.03	81	0.16	2	0.15	3
Alternative treatment: Music school fees in lowest third						
Plays music	-0.01	88	0.05	24		
Sample size	906		997		997	
NON-COGNITIVE SKILLS						
Strengths and Difficulties Questionnaire						
No emotional symptoms	0.10	71	-0.03	87	-0.09	68
No conduct problems	0.42	3	0.43	3	0.33	1
Not hyperactive	0.07	70	0.16	36	-0.06	72
No peer problems	0.03	91	-0.10	64	-0.12	29
Pro-social behavior	0.13	46	0.20	33	0.09	52
Attitude towards school						
Likes to go to school	0.34	5	0.30	6	0.11	39
Likes to study	0.27	36	0.36	5	0.16	43
Follows lessons well	-0.13	63	0.09	69	0.03	87
Gets along well...						
...with classmates	0.03	90	0.09	53	-0.02	90
...with teacher	0.18	51	0.37	1	0.17	13
Aggregated non-cognitive skills scores						
Non-cog. skills (aggregate)	0.15	12	0.19	4	0.06	45
Difficulties (from SDQ)	0.23	37	0.18	43	0.03	84
Sample size	682		689		689	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. The first regression weights observations using survey weights, which are provided in the SOEP and FiD data to make them representative for the German population. The second regression excludes all covariates. The third regression includes all covariates, but estimates a simple OLS without the JeKi group variable that is necessary to obtain a difference-in-differences estimator. For all regressions, control variables include individual and regional characteristics, as well as time and county fixed effects. Effects of the alternative treatment (music school fees in lowest third) additionally control for music school characteristics. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.20 – *Effects of treatments (OLS estimates) – Placebo treatment and placebo outcomes*

	Placebo treatment (1)		**Outcomes measured pre-treatment (age 6)** (2)	
	Difference	*p-value*	*Difference*	*p-value*
MUSIC ACTIVITIES				
Plays music	-0.02	81	0.01	94
Alternative treatment: Music school fees in lowest third				
Plays music	0.03	42	0.10	5
Sample size	997		343	
NON-COGNITIVE SKILLS				
Strengths and Difficulties Questionnaire				
No emotional symptoms	0.20	56	-0.24	52
No conduct problems	-0.02	94	-0.60	15
Not hyperactive	0.12	51	-0.03	94
No peer problems	0.21	49	-0.14	75
Pro-social behavior	-0.01	99	-0.40	8
Attitude towards school				
Likes to go to school	0.32	41		
Likes to study	-0.19	55		
Follows lessons well	0.06	80		
Gets along well...				
...with classmates	0.29	13		
...with teacher	0.35	16		
Aggregated non-cognitive skills scores				
Non-cog. skills (aggregate)	0.11	33	-0.32	25
Difficulties (from SDQ)	0.17	40	-0.61	28
Sample size	689		164	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Table D.1 for more details on the placebo treatment variables were defined. All regressions are estimated using difference-in-differences, and additionally control for individual and regional characteristics, as well as county fixed effects. Effects of the alternative treatment (music school fees in lowest third) are estimated with standard OLS (no difference-in-differences), additionally including music school characteristics as control variables. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. "Plays music" is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.

Table D.21 – *Effects of treatments (OLS estimates) – Sensitivity to individual school choice*

	Primary schools outside municipality excluded (1)		**Excluding movers (no moving between first observation and last outcome)** (2)		**Excluding movers (no moving between age 5 and last outcome)** (3)	
	Diff.	*p*	*Diff.*	*p*	*Diff.*	*p*
Music and sports activities						
Plays music	0.17	1	0.15	1	0.11	44
Alternative treatment: Music school fees in lowest third						
Plays music			0.10	8	0.12	20
Sample size	997		854		366	
Non-cognitive skills						
Strengths and Difficulties Questionnaire						
No emotional symptoms	-0.00	98	0.15	59	-0.56	61
No conduct problems	0.38	2	0.35	26	0.19	85
Not hyperactive	0.11	49	0.06	69	-0.21	85
No peer problems	-0.07	69	0.15	60	1.66	9
Pro-social behavior	0.16	40	0.25	43	0.28	82
Attitude towards school						
Likes to go to school	0.21	24	0.27	21	-0.12	91
Likes to study	0.30	16	0.07	76	-0.37	66
Follows lessons well	-0.01	98	-0.18	40	-0.53	62
Gets along well...						
...with classmates	0.13	54	0.04	85	0.96	41
...with teacher	0.26	3	0.41	1	1.43	29
Aggregated non-cognitive skills scores						
Non-cog. skills (aggregate)	0.14	6	0.15	14	0.17	74
Difficulties (from SDQ)	0.16	44	0.25	39	0.27	77
Sample size	689		512		91	

Source: SOEP v30 and FiD v4, own calculations. Each cell is a separate OLS regression estimating how access to the JeKi program affects the outcome denoted in the first column. Please refer to Section 5.4 for more details on the treatment definition. The first regression restricts the treatment definition to schools within the municipality (new treatment definition: JeKi share among students attending the 5 closest primary schools within municipality is larger than 50%). The second regression excludes all individuals who moved between the first time they enter the data and the year the last outcome is observed (age 8, 9 or 10). The third regression excludes all individuals who moved at least once between age 5 and 10. Here, the sample size is very small, given that only few children are actually observed in the data for such a long period. For most individuals, this is due to the fact that, so far, only 4 years of FiD data are available, by construction, children can thus not have been observed from age 5 to 10. For all regressions, control variables include individual and regional characteristics, as well as time and county fixed effects. Effects of the alternative treatment (music school fees in lowest third) additionally control for music school characteristics. Standard errors are clustered at the county level. The number of observations, as well as the R-squares for each regression can be provided on request. “Plays music” is a binary variable indicating whether the child played a musical instrument at or outside school at age 9. All other outcome variables are normalized with mean 0 and a standard deviation of 1 within the control group. The exact definitions of all outcome and control variables are given in Table D.1.