

Hille, Adrian; Schupp, Jürgen

**Article — Manuscript Version (Preprint)**

## How Learning a Musical Instrument Affects the Development of Skills

Economics of Education Review

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* Hille, Adrian; Schupp, Jürgen (2015) : How Learning a Musical Instrument Affects the Development of Skills, Economics of Education Review, ISSN 0272-7757, Elsevier Science, Amsterdam, Vol. 44, pp. 56-82,  
<https://doi.org/10.1016/j.econedurev.2014.10.007>

This Version is available at:

<https://hdl.handle.net/10419/105863>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# How learning a musical instrument affects the development of skills

Adrian Hille<sup>a,b,\*</sup>, Jürgen Schupp<sup>a,c</sup>

<sup>a</sup>*German Institute for Economic Research (DIW Berlin), Socio-Economic Panel Study, Mohrenstr. 58, 10117 Berlin, Germany, Phone: +49 30 89 789 376*

<sup>b</sup>*Free University of Berlin and DIW Berlin Graduate Center, Germany*

<sup>c</sup>*Free University of Berlin and IZA, Bonn, Germany*

---

## Abstract

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.

*JEL classification:* I21, J24, Z11

*Keywords:* Music, cognitive and non-cognitive skills, educational achievement, SOEP

---

---

\*Corresponding author

Email addresses: [ahille@diw.de](mailto:ahille@diw.de) (Adrian Hille), [jschupp@diw.de](mailto:jschupp@diw.de) (Jürgen Schupp)

## 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, 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).<sup>1</sup> However, as these studies are experimental, they focus 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

---

<sup>1</sup>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).

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 German Socio-Economic Panel (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. 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 1. Any hypothesis 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.<sup>2</sup> According to Schumacher (2009), learning

---

<sup>2</sup>Some psychologists argue that personality is genetically determined and cannot be modified (Pervin et al.,

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.<sup>3</sup> 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 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.

---

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 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).

<sup>3</sup>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.

**Table 1** – *Hypotheses: Potential effects of music training*

<b>Hypothesis</b>	<b>Mechanism</b>
<i>Cognitive skills</i>	
→ Improved cognitive skills	Influence on subdomains of cognitive function, executive function or via non-cognitive skills (Schellenberg, 2004, 2011)
<i>Non-cognitive skills</i>	
→ 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)
<i>School achievement</i>	
→ Improved school grades	Positive signal to school teachers (Lareau, 2011), improved cognitive skills
<i>Cultural and social capital</i>	
→ 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)
<i>Time use</i>	
→ Changes in time use	Crowding out of positive or negative activities (Felfe et al., 2011), structure learning and time schedule (Lareau, 2011)
<i>Inequality</i>	
→ 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)

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.

### 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 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.

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.



**Table 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
<i>Number of adolescents (total: 3,941)</i>	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.

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.<sup>4</sup> 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 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.<sup>5</sup> Time-varying variables were measured as early as available for each individual, but no earlier than age five.<sup>6</sup> 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.<sup>7</sup> 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 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.

---

<sup>4</sup>We take this information from the standard SOEP household and individual questionnaires.

<sup>5</sup>However, our results are also robust to including covariates for the father, in addition to or instead of those of the mother (see Appendix C).

<sup>6</sup>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 C).

<sup>7</sup>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 C).

#### 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 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 Roman, 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.<sup>8</sup> 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 towards 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 the association of German public music schools obliges its members to propose reduced fees to individuals with low income (VDM, 2011).<sup>9</sup>

---

<sup>8</sup>These variables are retrospectively reported by the adolescent herself at age 17.

<sup>9</sup>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.

This study considers adolescents as treated, if they played music at age 8 or before and continued to do so until age 17.<sup>10</sup> 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.<sup>11</sup> 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 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 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, 2014). Second, propensity score matching allows us to use a large number of control variables even with a limited sample 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 B.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

---

<sup>10</sup>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 6.

<sup>11</sup>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.

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 B.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 B.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).<sup>12</sup> Algebraically, we estimate:

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

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.<sup>13</sup>

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 6 discusses the plausibility of this assumption.

---

<sup>12</sup>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.

<sup>13</sup>An analytical approximation would be possible, but is not recommended by Huber et al. (2014).

## 5. Results

### 5.1. Main results

Table 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 child in music classes, financial and material constraints within the household to do so, as well as characteristics of the adolescent herself 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 C.

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 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 towards 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.<sup>14</sup>

---

<sup>14</sup>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.

**Table 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 (5)	Sports (age 8 and 17, competitions) (6)	Sample size
	No controls (1)	Full controls (2)	Low (3)	High (4)			
<i>Cognitive skills (in std. dev.)</i>							
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
<i>School grades<sup>1</sup> (in std. dev.)</i>							
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
<i>Personality (in std. dev.)</i>							
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
<i>Time use (in %)</i>							
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
<i>Ambition (in %)</i>							
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

<sup>1</sup> 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. Columns (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: <sup>+</sup>  $p < 0.1$  \* \* \*  $p < 0.05$  \* \* \* \*  $p < 0.01$

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 C.11 and C.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 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 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 C.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 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.

## 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 C.

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 C.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 C.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.<sup>15</sup> 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.<sup>16</sup>

The effects of music for this small subsample are presented in column (12) of Table C.17. Due to the small sample size, none of the outcome differences are 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.

## 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

---

<sup>15</sup>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.

<sup>16</sup>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.

outcome variables of interest. As described in Section 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 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 4 shows outcome differences between musically active and inactive adolescents adding additional control variables.<sup>17</sup> Each column includes the variables mentioned in the table

---

<sup>17</sup>The propensity to give up music and other additional control variables presented in Table 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.

header as intermediate variables. As shown in column (1), our results are robust to including the probability to give up music and secondary school type as control variables.

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 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:

$$Y_i^p = \alpha + \beta \cdot \text{Music}_i + \gamma Y_i^q + \varepsilon \quad \text{for all } p, q \text{ with } p \neq q \quad (2)$$

where control observations (with  $\text{Music}_i = 0$ ) receive weights obtained by propensity score matching, as described in Section 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 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.

## 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.

**Table 4 – Outcome differences between adolescents with and without music training controlling for additional intermediate variables**

	(1)	(2)	(3)	(4)	(5)
<i>Intermediate variables additionally controlled for</i>					
Probability to give up music before 17					
Average cognitive skills	X	X	X	X	X
Average school grade		X			
Conscientiousness			X		
Perceived control				X	X
<i>Cognitive skills (in standard deviations)</i>					
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 <sup>+</sup> (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)
<i>School grades<sup>1</sup> (in standard deviations)</i>					
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 <sup>+</sup> (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 <sup>+</sup> (0.09)	-0.13** (0.06)
Mathematics grade	-0.12 <sup>+</sup> (0.07)	-0.13 (0.10)	0.02 (0.05)	-0.10 (0.10)	-0.12 <sup>+</sup> (0.07)
<i>Personality (in standard deviations)</i>					
Conscientiousness	0.24** (0.09)	0.23 <sup>+</sup> (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 <sup>+</sup> (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)	
<i>Time use (in %)</i>					
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 <sup>+</sup> (3.37)	5.80 (5.30)	4.45 (3.31)	6.39 (5.05)	5.41 (3.35)
<i>Ambition (in %)</i>					
Aim Abitur	4.23 <sup>+</sup> (2.18)	-0.64 (3.56)	3.61 <sup>+</sup> (2.10)	2.17 (3.19)	4.11 <sup>+</sup> (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 <sup>+</sup> (1.26)	2.77 (1.94)	1.94 (1.28)	1.59 (1.78)	2.19 <sup>+</sup> (1.27)

<sup>1</sup> 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 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: <sup>+</sup>  $p < 0.1$  \*  $p < 0.05$  \*\*  $p < 0.01$

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.

## Acknowledgements

This research is funded by the German Research Foundation (DFG) in the context of the Collaborative Research Center 882 “From Heterogeneities to Inequalities” (Project A1 Social Closure and Hierarchization). We wish to thank Silke Anger, Vernon Gayle, Michael Lechner, Katharina Spiess, Gert G. Wagner, an anonymous referee as well as seminar participants at DIW Berlin, the University of Bielefeld, the University of St. Gallen and the Annual Congress 2013 of the Verein für Socialpolitik for their comments. Special thanks to Deborah Bowen for her support with English language editing. We thank Annegret Arnold for valuable research assistance. All errors remain our own.

- Almlund, M., Duckworth, A. L., Heckman, J. J., Kautz, T., 2011. Personality psychology and economics. In: Hanushek, E. A., Machin, S., Woessmann, L. (Eds.), *Handbook of the Economics of Education*. Vol. 4. Elsevier Amsterdam, pp. 1–181.
- Barron, J. M., Ewing, B. T., Waddell, G. R., 2000. The effects of high school athletic participation on education and labor market outcomes. *The Review of Economics and Statistics* 82 (3), 409–421.
- Besson, M., Schon, D., Moreno, S., Santos, A., Magne, C., 2007. Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience* 25 (3-4), 399–410.
- Bialystok, E., DePape, A.-M., 2009. Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology, Human Perception and Performance* 35 (2), 565–574.
- Bilhartz, T. D., Bruhn, R. A., Olson, J. E., 1999. The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology* 20 (4), 615–636.
- BMBF, 2012. 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).
- BMBF, 2013. 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).
- Bourdieu, P., 1986. The forms of capital. In: Richardson, J. G. (Ed.), *Handbook of Theory of Research for the Sociology of Education*. Greenwood Press, pp. 241–258.
- Bourdieu, P., Passeron, J.-C., 1990. *Reproduction in education, society and culture*. Vol. 4. SAGE Publications Limited.
- Cobb-Clark, D. A., Schurer, S., 2012. The stability of big-five personality traits. *Economics Letters* 115 (1), 11–15.
- Covay, E., Carbonaro, W., 2010. After the bell: Participation in extracurricular activities, classroom behavior, and academic achievement. *Sociology of Education* 83 (1), 20–45.
- Cunha, F., Heckman, J. J., 2007. The technology of skill formation. *American Economic Review* 97 (2), 31–47.
- Cunha, F., Heckman, J. J., 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43 (4), 738–782.
- Cunha, F., Heckman, J. J., Schennach, S. M., 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78 (3), 883–931.
- DiMaggio, P., 1982. Cultural capital and school success: The impact of status culture participation on the grades of U.S. high school students. *American Sociological Review* 47 (2), 189–201.
- Donnellan, M. B., Lucas, R. E., 2008. Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging* 23 (3), 558–566.
- Eide, E. R., Ronan, N., 2001. Is participation in high school athletics an investment or a consumption good? evidence from high school and beyond. *Economics of Education Review* 20 (5), 431–442.
- Felfe, C., Lechner, M., Steinmayr, A., 2011. Sports and child development. IZA Discussion Paper 6105, Institute for the Study of Labor (IZA).
- FMSB, 2013. El Sistema. Official Site of the Simon Bolivar Music Foundation: Fundación Musical Simón Bolívar (FMSB), <http://fundamusical.org.ve/> (accessed June 16th, 2014).

- Garboui, L., Montmarquette, C., 1996. A microeconomic study of theatre demand. *Journal of Cultural Economics* 20, 25–50.
- Heckman, J., Pinto, R., 2013. 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).
- Heckman, J. J., Kautz, T., 2012. Hard evidence on soft skills. *Labour Economics* 19, 451–464.
- Heckman, J. J., Masterov, D. V., 2007. The productivity argument for investing in young children. *Applied Economic Perspectives and Policy* 29 (3), 446–493.
- Heckman, J. J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24 (3), 411–482.
- Heineck, G., Anger, S., 2010. The returns to cognitive abilities and personality traits in Germany. *Labour Economics* 17 (3), 535–46.
- Ho, Y.-C., Cheung, M.-C., Chan, A. S., 2003. Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology* 17 (3), 439–450.
- Huber, M., Lechner, M., Steinmayr, A., 2014. Radius matching on the propensity score with bias adjustment: Finite sample behaviour, tuning parameters and software implementation. *Empirical Economics*, forthcoming.
- Huber, M., Lechner, M., Wunsch, C., 2013. The performance of estimators based on the propensity score. *Journal of Econometrics* 175 (1), 1–21.
- Imai, K., Keele, L., Tingley, D., 2010. A general approach to causal mediation analysis. *Psychological Methods* 15 (4), 309–334.
- Imbens, G. W., 2014. Matching methods in practice: Three examples. IZA Discussion Paper 8049, Institute for the Study of Labor (IZA).
- Imbens, G. W., Wooldridge, J. M., 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47 (1), 5–86.
- JeKi, 2014. Jedem Kind ein Instrument: Programmbeschreibung.  
[https://www.jedemkind.de/programm/mediathek/pdf/120326\\_programminfos\\_2011\\_2012\\_mit\\_kopf.pdf](https://www.jedemkind.de/programm/mediathek/pdf/120326_programminfos_2011_2012_mit_kopf.pdf) (accessed June 16th, 2014).
- Jürges, H., Schneider, K., Büchel, F., 2005. The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association* 3 (5), 1134–1155.
- Lang, F., John, D., Lüdtke, O., Schupp, J., Wagner, G., 2011. Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods* 43 (2), 548–567.
- Lareau, A., 2011. *Unequal Childhoods: Class, Race, and Family Life*, 2nd Edition. University of California Press.
- Lareau, A., Weininger, E. B., 2003. Cultural capital in educational research: A critical assessment. *Theory and Society* 32 (5/6), 567–606.
- Lindenberg, S., 1989. Social production functions, deficits, and social revolutions: Prerevolutionary France and Russia. *Rationality and Society* 1 (1), 51–77.
- Loui, P., Kroog, K., Zuk, J., Winner, E., Schlaug, G., 2011. Relating pitch awareness to phonemic awareness in children: Implications for tone-deafness and dyslexia. *Frontiers in Psychology* 2 (111), 1–5.
- Lunn, P., Kelly, E., 2009. 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.
- McCrae, R. R., Costa, P. T., 1999. The five factor theory of personality. In: John, O. P., Robins, R. W., Pervin, L. A. (Eds.), *Handbook of personality: Theory and research* 2. Guilford, New York, pp. 139–153.
- Menninghaus, W., 2011. *Wozu Kunst? Ästhetik nach Darwin*. Suhrkamp Berlin.
- Nering, M. E., 2002. The effect of piano and music instruction on intelligence of monozygotic twins. *Dissertation Abstracts International Section A: Humanities and Social Sciences* 63 (3A), 812.
- Neville, H., 2008. Effects of music training on brain and cognitive development in under-privileged 3- to 5-year-old children: Preliminary results. In: Rich, B., Asbury, C. (Eds.), *Learning, Arts, and the Brain: The Dana Consortium Report on Arts and Cognition*. The Dana Foundation, New York/Washington, DC, pp. 105–106.
- Ormel, J., Lindenberg, S., Steverink, N., Verbrugge, L. M., 1999. Subjective well-being and social production functions. *Social Indicators Research* 46, 61–90.
- Pervin, L. A., Cervone, D., John, O. P., 2005. *Personality: Theory and Research*, 9th Edition. John Wiley & Sons, Inc.



- Pfeifer, C., Cornelissen, T., 2010. The impact of participation in sports on educational attainment: New evidence from Germany. *Economics of Education Review* 29 (1), 94–103.
- Schellenberg, E. G., 2004. Music lessons enhance IQ. *Psychological Science* 15 (8), 511–514.
- Schellenberg, E. G., 2006. Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology* 98 (2), 457–468.
- Schellenberg, E. G., 2011. Examining the association between music lessons and intelligence. *British Journal of Psychology* 102 (3), 283–302.
- Schellenberg, E. G., Moreno, S., 2010. Music lessons, pitch processing, and g. *Psychology of Music* 38 (2), 209–221.
- Schumacher, R., 2009. Pauken mit Trompeten. Lassen sich Lernstrategien, Lernmotivation und soziale Kompetenzen durch Musikunterricht fördern? *Bildungsforschung Band 32*, Bundesministerium für Bildung und Forschung (BMBF).
- Schupp, J., Herrmann, S., 2009. Kognitionspotenziale Jugendlicher. Ergänzungen zum Jugendfragebogen der Längsschnittstudie Sozio-oekonomisches Panel (SOEP). DIW-Data Documentation No. 43, DIW Berlin.
- Shahin, A. J., Roberts, L. E., Chau, W., Trainor, L. J., Miller, L. M., 2008. Music training leads to the development of timbre-specific gamma band activity. *NeuroImage* 41 (1), 113 – 122.
- Southgate, D. E., Roscigno, V. J., 2009. The impact of music on childhood and adolescent achievement. *Social Science Quarterly* 90 (1), 4–21.
- Specht, J., Egloff, B., Schmukle, S. C., 2011. 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* 101 (4), 862–882.
- Specht, J., Egloff, B., Schmukle, S. C., 2013. Everything under control? the effects of age, gender, and education on trajectories of perceived control in a nationally representative german sample. *Developmental Psychology* 49 (2), 353–364.
- Stevenson, B., 2010. Beyond the classroom: Using Title IX to measure the return to high school sports. *Review of Economics and Statistics* 92 (2), 284–301.
- Stuart, E. A., 2010. Matching methods for causal inference: A review and a look forward. *Statistical Science* 25 (1), 1–21.
- Switlick, B., Bullerjahn, C., 1999. Ursachen und Konsequenzen des Abbruchs von Instrumentalunterricht. Eine quantitative und qualitative Umfrage bei Studierenden der Universität Hildesheim. In: Knolle, N. (Ed.), *Musikpädagogik vor neuen Forschungsaufgaben*. Essen, pp. 167–195.
- Todd, P. E., Wolpin, K. I., 2003. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal* 113 (485), F3–F33.
- Vaughn, K., Winner, E., 2000. SAT scores of students who study the arts: What we can and cannot conclude about the association. *Journal of Aesthetic Education* 34 (3/4), 77–89.
- VDM, 2011. 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](http://www.musikschulen.de/medien/doks/vdm/richtlinien-des-vdm-2011_logo.pdf) (accessed June 16th, 2014).
- Wagner, G. G., Frick, J. R., Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften* 127 (1), 139–169.
- Weinhardt, M., Schupp, J., 2011. Multi-Itemskalen im SOEP Jugendfragebogen. Data Documentation 60, DIW Berlin, German Institute for Economic Research.
- Winner, E., Goldstein, T., Vincent-Lancrin, S., 2013. *Art for Art's Sake?* OECD Publishing.
- Yaish, M., Katz-Gerro, T., 2012. Disentangling “cultural capital”: The consequences of cultural and economic resources for taste and participation. *European Sociological Review* 28 (2), 169–185.

## Appendix A. 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 – Sample construction**

<b>Sample description</b>	<b>Remaining observations</b>
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
<b>Sample for estimating the propensity score</b>	<b>3,941</b>
<i>I. Outcomes available for every survey year</i>	
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
<b>Final sample I</b>	<b>3,488</b>
<i>II. Cognitive skills: measured since 2004, only for questionnaires answered in the presence of an interviewer</i>	
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
<b>Final sample II</b>	<b>1,847</b>
<i>III. Big 5 Personality traits: Only measured since 2006</i>	
Drop individuals surveyed between 2001 to 2005	1,872
Drop individuals with missing for at least one personality dimension	1,815
<b>Final sample III</b>	<b>1,815</b>

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

**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 (within control group): Mean: 0 SD: 1
Analogies	See Table A.3	2004-2012		
Maths operators	See Table A.3	2004-2012		
Average cog. skills	Mean of the above	2004-2012		
School grades				
German	Self-reported	2001-2012	1 (very good) to 6 (fail)	Normalized (within control group and school track) Mean: 0   SD: 1
Mathematics	Self-reported	2001-2012		
First foreign language	Self-reported	2001-2012		
Average grade	Mean of the above	2001-2012		
Personality				
Conscientiousness	See Table A.3	2006-2012	7-point Likert-scale	Normalized (within control group) Mean: 0 SD: 1
Openness	See Table A.3	2006-2012		
Agreeableness	See Table A.3	2006-2012		
Extraversion	See Table A.3	2006-2012		
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%

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

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

Variable	Items
<i>Cognitive skills</i>	
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
<i>Personality</i>	I see myself as someone who...
Conscientiousness	...does a thorough job ...does things effectively and efficiently ...tends to be lazy ( <i>reversed</i> )
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 ( <i>reversed</i> ) ...has a forgiving nature ...is considerate and kind to others
Extraversion	...is communicative, talkative ...is outgoing, sociable ...is reserved ( <i>reversed</i> )
Neuroticism	...worries a lot ...gets nervous easily ...is relaxed, handles stress well ( <i>reversed</i> )
Perceived control	How my life goes, depends on myself Compared to others, I have not achieved what I deserved ( <i>reversed</i> ) What one achieves is mainly a question of luck and fate ( <i>reversed</i> ) I often have the experience that others make decisions regarding my life ( <i>reversed</i> ) When I encounter difficulties I have doubts about my abilities ( <i>reversed</i> ) Opportunities in life are determined by social conditions ( <i>reversed</i> ) I have little control over the things that happen in my life ( <i>reversed</i> )

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
- *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 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.4 – List of control variables**

Variable	SOEP file	Measured at age...
<i>Motivation for music enrollment<sup>1</sup></i>		
Mother has no degree	PGEN (generated indiv. variables)	age 5 or entry <sup>2</sup>
Mother has completed only 9 years of school	PGEN	age 5 or entry <sup>2</sup>
Mother has vocational degree	PGEN	age 5 or entry <sup>2</sup>
Mother has university degree	PGEN	age 5 or entry <sup>2</sup>
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) <sup>3</sup>	P (individual questionnaire)	2005 or 2009 <sup>4</sup>
Extraversion (mother) <sup>3</sup>	P	2005 or 2009 <sup>4</sup>
Neuroticism (mother) <sup>3</sup>	P	2005 or 2009 <sup>4</sup>
Agreeableness (mother) <sup>3</sup>	P	2005 or 2009 <sup>4</sup>
Openness (mother) <sup>3</sup>	P	2005 or 2009 <sup>4</sup>
Mother attends cultural events every month	P	age 5 or entry <sup>2</sup>
Mother never attends cultural events	P	age 5 or entry <sup>2</sup>
Mother is artistically active every month	P	age 5 or entry <sup>2</sup>
Mother is never artistically active	P	age 5 or entry <sup>2</sup>
Mother's appreciation for the arts	P	2005 or 2009 <sup>4</sup>
More than 200 books at home	H (household questionnaire)	2001, 2006 or 2011 <sup>4</sup>
Fewer than 50 books at home	H	2001, 2006 or 2011 <sup>4</sup>
<i>Constraints for music enrolment</i>		
Monthly household log net income (simple, squared)	HGEN (generated household variables)	age 5 or entry <sup>2</sup>
Rooms per person at home	HGEN	age 5 or entry <sup>2</sup>
Number of siblings in the SOEP	BIOSIB	time constant
Child is the firstborn	BIOSIB	time constant
Federal state (15 dummies)	HBRUTTO	time constant
Household lives in a rural area	HBRUTTO (fieldwork information)	age 5 or entry <sup>2</sup>
<i>Adolescent's motivation to continue playing music</i>		
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

<sup>1</sup> Mother's value if available, otherwise the father's value is used.

<sup>2</sup> 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.

<sup>3</sup> For items and measurement, please refer to Table A.3.

<sup>4</sup> The earliest year with non-missing observation.

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

	<b>Mean</b>	<b>Standard deviation</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Sample size</b>
<i>Cognitive skills</i>					
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
<i>School grades<sup>1</sup></i>					
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
<i>Personality</i>					
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
<i>Time use</i>					
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
<i>Ambition</i>					
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

<sup>1</sup> 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
<i>Motivation to play music</i>					
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
<i>Constraints for playing music</i>					
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
<i>Constraints for playing music (continued)</i>					
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
<i>Adolescent's motivation to continue playing music</i>					
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
<i>Survey-specific control variables</i>					
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
<i>Missing indicators</i>					
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.

## Appendix B. 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 B.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 text of this study, financial resources of the household play an insignificant role, once the other characteristics are taken into account. In Table B.8, the marginal effect of household income is even negative, albeit insignificant.

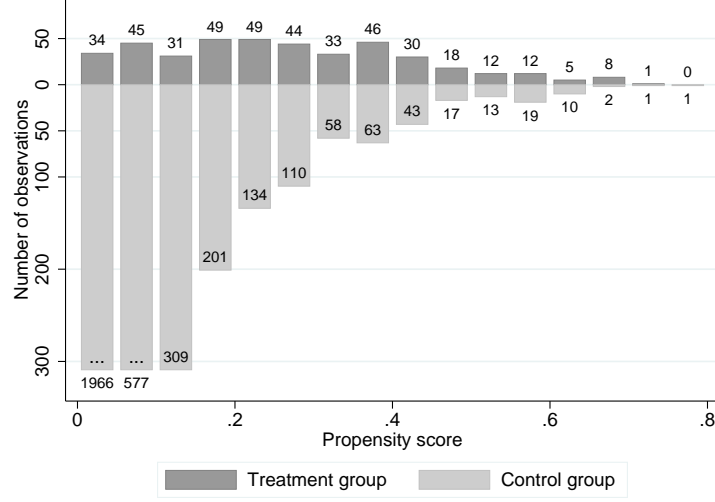
2. **Check common support.** Comparing the distribution of the propensity score between treatment and control group indicates whether we can find comparable control observations for each treated individual. Figure B.1 provides graphical evidence that this common support assumption holds. The graph shows 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.

**Table B.8** – Estimation of the propensity score (probit model, marginal effects, in percent)

<b>Play music at age 17, started at age 8 or before, take music lessons</b>		
<i>Motivation to play music</i>		
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 <sup>+</sup>	(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)
<i>Constraints for playing music</i>		
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)
<i>Adolescent's motivation to continue playing music</i>		
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: <sup>+</sup>  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

**Figure B.1 – Common support**



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 B.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:

$$A\hat{T}T = \frac{1}{N_T} \sum_{i=1}^{N_T} T_i y_i - \frac{1}{N_C} \sum_{i=1}^{N_C} (1 - T_i) \hat{w}_i y_i \quad (\text{B.1})$$

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 B.9 – Balancing of covariates after propensity score matching**

	Before matching		After matching	
	Difference	t-value	Difference	t-value
<i>Motivation to play music</i>				
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 <sup>+</sup>	-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
<i>Constraints for playing music</i>				
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
<i>Adolescent's motivation to continue playing music</i>				
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: <sup>+</sup>  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

## Appendix C. Further results

### *Robustness I: Heterogeneous effects*

Tables C.11 and C.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 C.10.

**Table C.10 – Sample sizes for heterogeneous results**

Sample definition (Column in Table C.11 / C.12)	Control group	Treatment group	% treated
Full sample from baseline specification	3,524	417	10.6
<i>Subgroups by gender:</i>			
Male (1)	1,844	151	7.6
Female (2)	1,680	266	13.7
<i>Subgroups by living area:</i>			
Urban area (3)	2,620	333	11.3
Rural area (4)	904	84	8.5
<i>Subgroups by household income:</i>			
Household income below median (5)	1,738	114	6.2
Household income above median (6)	1,786	303	14.5
<i>Subgroups by parental education:</i>			
Mother without Abitur/university degree (7)	2,068	115	5.3
Mother with Abitur/university degree (8)	1,371	298	17.9
<i>Subgroups by cultural capital:</i>			
Fewer than 200 books at home (9)	2,560	162	6.0
More than 200 books at home (10)	964	255	20.9
<i>Subgroup by type of secondary school attended:</i>			
Adolescent doesn't 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 C.11 and C.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.

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 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 C.11 – Outcome differences between adolescents with and without music training (heterogeneous effects, part 1 of 2)**

	<b>Gender</b>		<b>Living area</b>		<b>Household income</b>	
	<i>male</i> (1)	<i>female</i> (2)	<i>not rural</i> (3)	<i>rural</i> (4)	<i>&lt;median</i> (5)	<i>&gt;median</i> (6)
<i>Cognitive skills (in std. dev.)</i>						
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)
<i>School grades<sup>1</sup> (in std. dev.)</i>						
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 <sup>+</sup> (0.08)	-0.15** (0.07)	-0.11 (0.24)	-0.08 (0.17)	-0.16 <sup>+</sup> (0.08)
Foreign language grade	-0.12 (0.14)	-0.12 (0.09)	-0.14 <sup>+</sup> (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)
<i>Personality (in std. dev.)</i>						
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 <sup>+</sup> (0.11)	0.42 (0.34)	0.33 <sup>+</sup> (0.20)	0.22 <sup>+</sup> (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)
<i>Time use (in %)</i>						
Watch TV daily	-8.73 (6.39)	-9.41** (4.74)	-9.38** (3.85)	-24.71** (11.49)	-16.02 <sup>+</sup> (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)
<i>Ambition (in %)</i>						
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 <sup>+</sup> (4.19)	11.24*** (3.36)	4.72 (12.06)	14.72 <sup>+</sup> (8.27)	6.61 <sup>+</sup> (3.65)
Job success likely	1.81 (2.20)	0.69 (1.53)	-0.54 (1.29)	6.17 (3.79)	4.81 <sup>+</sup> (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)

<sup>1</sup> 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: <sup>+</sup>  $p < 0.1$  \* \*  $p < 0.05$  \* \* \*  $p < 0.01$



**Table C.12 – Outcome differences between adolescents with and without music training (heterogeneous effects, part 2 of 2)**

	Parents' education		Cultural capital		Secondary school	
	<i>low</i> (7)	<i>high</i> (8)	<i>low</i> (9)	<i>high</i> (10)	<i>lower</i> (11)	<i>upper</i> (12)
<i>Cognitive skills (in std. dev.)</i>						
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 <sup>+</sup> (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 <sup>+</sup> (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 <sup>+</sup> (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)
<i>School grades<sup>1</sup> (in std. dev.)</i>						
Average school grade	-0.21 (0.14)	-0.12 (0.08)	-0.23** (0.11)	-0.03 (0.10)	-0.23 <sup>+</sup> (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)
<i>Personality (in std. dev.)</i>						
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 <sup>+</sup> (0.18)	0.22 <sup>+</sup> (0.12)	0.43*** (0.16)	0.12 (0.15)	0.29 <sup>+</sup> (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)
<i>Time use (in %)</i>						
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 <sup>+</sup> (4.66)	14.14*** (5.44)	1.11 (5.37)	9.12 (5.65)	4.50 (4.85)
<i>Ambition (in %)</i>						
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 <sup>+</sup> (4.02)
Job success likely	1.51 (2.45)	0.64 (1.53)	3.70 <sup>+</sup> (2.06)	-0.06 (1.70)	0.74 (2.14)	1.13 (1.62)
Desired profession likely	5.47 <sup>+</sup> (2.95)	0.76 (1.73)	5.49** (2.24)	-0.79 (2.06)	1.80 (2.32)	1.51 (1.87)

<sup>1</sup> 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: <sup>+</sup>  $p < 0.1$  \*  $p < 0.05$  \*\*  $p < 0.01$

### *Robustness II: Varying treatment definitions*

Table C.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 C.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 C.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 C.13 – Sample sizes for treatment definitions**

<b>Treatment definition (Column in Table C.14)</b>	<b>Control group</b>	<b>Treatment group</b>	<b>% treated</b>
<i>Main definition:</i> Play music at age 17, started at age 8 or before, take music lessons	3,524	417	10.6
<i>Main definition without music lessons (1):</i> Play music at 17, started at age 8 or before	3,476	465	11.8
<i>Main definition without age restriction (2):</i> Play music at age 17, take music lessons	3,232	709	18.0
<i>Main definition without age restriction and lessons (3):</i> Play music at age 17	2,978	963	24.4
<i>Main definition with different age restriction (4):</i> Play music at 17, started at age 9 or later, take music lessons	3,267	257	7.3
<i>Main definition, classical music only (5):</i> Music at 17, started at 8 or before, have lessons, play classical music	3,649	154	4.0
<i>Main definition, play at least weekly (6):</i> 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 C.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 C.14 – Outcome differences between adolescents with and without music training (different treatment definitions)**

	$\leq$ age 8 & 17	age 17 lessons	age 17	$>$ age 8 & 17 lessons	$\leq$ age 8 & 17 classical lessons	$\leq$ age 8 & 17 lessons weekly
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Cognitive skills (in std. dev.)</i>						
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 <sup>+</sup> (0.06)	-0.04 (0.10)	0.12 (0.15)	-0.07 (0.09)
<i>School grades<sup>1</sup> (in std. dev.)</i>						
Average school grade	-0.13** (0.06)	-0.18*** (0.05)	-0.13*** (0.04)	-0.12 (0.08)	-0.21 <sup>+</sup> (0.12)	-0.11 <sup>+</sup> (0.07)
German grade	-0.13** (0.06)	-0.16*** (0.05)	-0.14*** (0.05)	-0.14 <sup>+</sup> (0.08)	-0.16 (0.11)	-0.12 <sup>+</sup> (0.07)
Foreign language grade	-0.12 <sup>+</sup> (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 <sup>+</sup> (0.08)	-0.14 (0.13)	-0.05 (0.07)
<i>Personality (in std. dev.)</i>						
Conscientiousness	0.20** (0.09)	0.14 <sup>+</sup> (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 <sup>+</sup> (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)
<i>Time use (in %)</i>						
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 <sup>+</sup> (3.20)	6.02** (2.78)	7.67*** (2.31)	5.73 (3.71)	9.00 (6.08)	7.72** (3.68)
<i>Ambition (in %)</i>						
Aim Abitur	3.61 <sup>+</sup> (2.14)	6.08*** (1.84)	4.87*** (1.68)	7.26** (3.02)	1.81 (3.59)	4.36 <sup>+</sup> (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 <sup>+</sup> (1.34)

<sup>1</sup> 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 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 C.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: <sup>+</sup>  $p < 0.1$  \* \*  $p < 0.05$  \* \* \*  $p < 0.01$

### *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 C.16 and C.17. Due to variations in sample composition and treatment definitions, sample sizes are not necessarily comparable between all robustness estimations. Therefore, Table C.15 presents the sample sizes of treatment and control group for each robustness specification.

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 C.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 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 C.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 C.15 – Sample sizes for robustness checks**

Estimation details (Column in Table C.16 and C.17)	Control group	Treatment group	% treated	Treatment definition
OLS estimation (1)	3,524	417	10.6	Main definition
Propensity score matching: radius matching (2), <i>caliper 0.02 instead of 0.01 in main specification</i>	3,524	417	10.6	Main definition
Propensity score matching: Kernel matching (3), <i>Epanechnikov kernel</i>	3,524	417	10.6	Main definition
Subsample: complete information for all outcomes (4), <i>one sample for all outcomes</i>	1,074	165	13.3	Main definition
Subsample: complete information for all covariates (5), <i>no missing indicators</i>	2,607	327	11.1	Main definition
Subsample: individuals in sample before treatment started (6), <i>other treated individuals dropped</i>	3,524	171	4.6	Main definition
Subsample: Individuals in sample at age five (7), <i>when covariates were measured</i>	1,345	151	10.1	Main definition
Controls: add number of siblings X household income and firstborn X household income (8), <i>keep original control variables</i>	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), <i>replace original control variables</i>	3,524	417	10.6	Play music
Treatment: play music at age 17 (10), <i>original treatment group dropped</i>	2,978	546	15.5	Play music
Treatment: play music at least sometimes (11), <i>original treatment group dropped</i>	2,067	1,405	40.5	Some music
Treatment: music age 12, but gave up before 17 (12), <i>subsample with complete music history</i>	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 C.16 and C.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.

**Table C.16 – Outcome differences between adolescents with and without music training (Robustness checks, part 1 of 2)**

	OLS (1)	Caliper matching ( $r = 0.02$ ) (2)	Kernel matching (3)	No missing outcomes (4)	No missing covariates (5)	In sample before treatment (6)
<i>Cognitive skills (in std. dev.)</i>						
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 <sup>+</sup> (0.08)	0.14 <sup>+</sup> (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)
<i>School grades<sup>1</sup> (in std. dev.)</i>						
Average school grade	-0.15*** (0.05)	-0.16*** (0.06)	-0.16*** (0.06)	-0.19 <sup>+</sup> (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 <sup>+</sup> (0.11)
Foreign language grade	-0.09 <sup>+</sup> (0.05)	-0.10 <sup>+</sup> (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 <sup>+</sup> (0.06)	-0.10 <sup>+</sup> (0.06)	-0.11 (0.12)	-0.03 (0.08)	-0.13 (0.11)
<i>Personality (in std. dev.)</i>						
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 <sup>+</sup> (0.07)	0.14 <sup>+</sup> (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)
<i>Time use (in %)</i>						
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 <sup>+</sup> (5.52)
Read books daily	7.04** (2.86)	5.23 (3.35)	6.18 <sup>+</sup> (3.21)	9.58 (6.21)	2.69 (3.82)	4.96 (5.67)
<i>Ambition (in %)</i>						
Aim Abitur	3.83 <sup>+</sup> (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 <sup>+</sup> (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)

<sup>1</sup> 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 C.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: <sup>+</sup>  $p < 0.1$  \* \* \*  $p < 0.05$  \* \* \* \*  $p < 0.01$



**Table C.17 – Outcome differences between adolescents with and without music training (Robustness checks, part 2 of 2)**

	<i>In sample at age 5</i>	<i>Interactions sibl. x HHinc</i>	<i>Father's characteristics</i>	<i>Music, but not treated</i>	<i>Some music (not never)</i>	<i>Give up before age 17</i>
	(7)	(8)	(9)	(10)	(11)	(12)
<i>Cognitive skills (in std. dev.)</i>						
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 <sup>+</sup> (0.07)	0.02 (0.06)	0.08 (0.27)
Maths operators	-0.25 <sup>+</sup> (0.15)	-0.01 (0.08)	-0.02 (0.08)	-0.09 (0.07)	-0.06 (0.06)	0.01 (0.26)
<i>School grades<sup>1</sup> (in std. dev.)</i>						
Average school grade	-0.24 <sup>+</sup> (0.13)	-0.16*** (0.06)	-0.18*** (0.06)	-0.06 (0.05)	-0.08 <sup>+</sup> (0.04)	0.20 (0.26)
German grade	-0.22 <sup>+</sup> (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 <sup>+</sup> (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 <sup>+</sup> (0.07)	-0.12 <sup>+</sup> (0.07)	-0.04 (0.06)	-0.03 (0.04)	0.20 (0.24)
<i>Personality (in std. dev.)</i>						
Conscientiousness	0.27 <sup>+</sup> (0.15)	0.21** (0.09)	0.23** (0.09)	-0.01 (0.08)	0.12 <sup>+</sup> (0.06)	0.10 (0.20)
Openness	0.25 <sup>+</sup> (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 <sup>+</sup> (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)
<i>Time use (in %)</i>						
Watch TV daily	-10.06 <sup>+</sup> (5.87)	-10.64*** (3.42)	-11.01*** (3.35)	-2.68 (2.43)	-3.08 <sup>+</sup> (1.81)	-6.31 (11.26)
Read books daily	5.66 (6.34)	6.64 <sup>+</sup> (3.42)	7.04** (3.46)	9.68*** (2.57)	0.63 (1.91)	-16.54 <sup>+</sup> (9.26)
<i>Ambition (in %)</i>						
Aim Abitur	7.19 (5.17)	3.91 <sup>+</sup> (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 <sup>+</sup> (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 <sup>+</sup> (0.77)	5.56 (4.14)
Desired profession likely	3.56 (2.36)	2.16 <sup>+</sup> (1.27)	2.01 (1.29)	0.69 (1.17)	2.28*** (0.89)	2.47 (4.44)

<sup>1</sup> 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 C.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: <sup>+</sup>  $p < 0.1$  \*  $p < 0.05$  \*\*  $p < 0.01$