Freie Universität Berlin

The long-term effect of age at school entry on competencies in adulthood

Katja Görlitz
Merlin Penny
Marcus Tamm

School of Business & Economics
Discussion Paper
Economics
2019/2

The long-term effect of age at school entry on competencies in adulthood

Katja Görlitz[a], Merlin Penny[b], and Marcus Tamm[c,*]

[a] FU Berlin, DIW, RWI and IZA
FU Berlin, Boltzmannstr. 20, 14195 Berlin, Germany, katja.goerlitz@fu-berlin.de

[b] FU Berlin
FU Berlin, Boltzmannstr. 20, 14195 Berlin, Germany, m.penny@fu-berlin.de

[c] RWI and IZA
RWI, Invalidenstr. 112, 10115 Berlin, Germany, tamm@rwi-essen.de

Abstract. The previous literature has shown that children who enter school at a more advanced age outperform their younger classmates on competency tests taken between kindergarten and Grade 10. This study analyzes whether these effects of school starting age continue into adulthood. Based on data on math and language test scores for adults in Germany, the identification of the long-term causal effects exploits state and year variation in school entry regulations. The results show that there are no effects of school starting age (SSA) on competencies in math and text comprehension. However, the long-term SSA effect is sizable on receptive vocabulary.

JEL classification: I21, J21, J31

Keywords: school starting age, education, cognitive competencies, instrumental variable estimates

[*] This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Adults, doi:10.5157/NEPS:SC6:8.0.0. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. Katja Görlitz and Marcus Tamm acknowledge financial support from the Deutsche Forschungsgemeinschaft (GO 2168/1-2 and TA 829/2-2).

1. Introduction

Many countries determine the age at which children may legally start school by defining cut-off dates for enrollment. All children born before the cut-off are supposed to enter school in a given year, while those born after the cut-off are expected to wait until the start of the next school year. This leads to considerable between-child variation in the school starting age (SSA) within a class. A vast empirical literature shows that SSA has important effects on children's school performances. Children that are enrolled at a higher SSA outperform their younger classmates in mathematics, reading, and writing (see, e.g., Bedard and Dhuey, 2006; Elder and Lubotsky, 2009; Smith, 2009). Most countries also give parents legal options to enroll their children with a one-year delay. These options usually allow parents of children, who would have had a low SSA if they were enrolled regularly, to delay school entry by one year, making the child a high SSA student. This practice, which is often called "red-shirting" in the academic literature, has become increasingly common in recent years in the US and Germany (Deming and Dynarski, 2008; Statistisches Bundesamt, 2017, 2018). Parents might consider red-shirting to be an optimal decision if enrolling their children with a lower SSA is accompanied by lifetime disadvantages for their children. However, previous literature, which documents the negative SSA effects, focuses mostly on the test score differentials of schoolchildren. Only a handful of papers go beyond Grade 10 and show that SSA effects become much smaller when looking at IQ and SAT scores at around the age of 18 (see, e.g., Black et al., 2011; Hurwitz et al., 2015).

This paper is the first to investigate whether SSA test score differentials continue into adulthood or fade away after leaving school. It evaluates SSA effects on adult competencies, as measured in comprehensive tests administered as part of a representative survey of individuals between 23 and 71 years old. Analyzing the persistence of SSA effects provides important information for parents. It also contributes to the literature investigating the long-run effects of SSA on individuals' wages and employment (e.g., Black et al., 2011; Fredriksson and Öckert, 2014; Larsen and Solli, 2017). Individuals with a lower SSA have higher wages and better employment perspectives shortly after entering the labor market. This result is due to a lower SSA being accompanied by longer tenure and actual experience when holding the age constant. However, Fredriksson and Öckert (2014) and Larsen and Solli (2017) conclude that SSA has no effects on cumulative earnings over an individual's life using data for Sweden and Norway, respectively. Since one of the potential channels for long-run effects on wages and employment is SSA-induced differences in adult competencies, our study should be viewed as complementary to this stream of literature.[1]

Our study also contributes to the literature showing that SSA test score differentials decrease as children progress through school (see, e.g., Bedard and Dhuey, 2006; Elder and Lubotsky, 2009). Even though many studies support this view, there is still "some disagreement about whether the effects are attenuated by middle school" (Cook and Kang, 2018, p. 2). Apart from our main analysis, which sheds further light on the potential attenuation by presenting the long-term effects of SSA, our literature section makes an additional contribution. It graphically

[1] Other papers on the long-run effects of SSA are only slightly connected to our research question. These papers involve studies investigating the effects of SSA on crime (e.g., Cook and Kang, 2016; Landersø et al., 2017), fertility (e.g., McCrary and Royer, 2011; Skirbekk et al., 2004), and marriage outcomes (e.g., Lefgren and McIntyre, 2006).

analyzes how the estimates of SSA on students' test scores presented in the previous literature differ by grade level.

The identification of the effect of SSA on competencies relies on an instrumental variable strategy that exploits the state- and year-specific rules given by the cut-off dates. The empirical model controls for a full set of month-of-birth dummies. This is important because Buckles and Hungerman (2013) point out that the distribution of family background differs by children's months of birth and differences in SSA might capture some of these effects. This capture is possible in our study because the school-entry regulations were subject to several changes at the state and year levels, allowing separate identification of the month-of-birth effects. In addition, our model can account for "age-at-test effects". Such an accounting is not possible in most of the previous studies that analyze SSA effects on test scores for school-aged children. If all children take the test at the same point in time, being older at school entry means automatically taking each test at an older age. Black et al. (2011) show that SSA effects become much smaller after controlling for the age at the time of testing. We are able to account for the age at the time of testing because the competency tests were not taken at a particular date for all individuals, but the time interval for the interviews stretched out over several months and the test-taking date was unrelated to the date of birth and the state where individuals went to school.

This paper contributes to the literature by additionally disentangling relative from absolute age effects, which is important from a policy perspective. Relative age measures the age difference compared to the ages of the other students within the cohort. Absolute age refers to the age (and, thus, maturity) when starting school. The previous literature is generally not able to separate these two potential channels when analyzing SSA differences. This inability to separate comes about due to the fact that the relative age at school entry is linearly related to the absolute age. One exception is the work of Cascio and Schanzenbacher (2016), which separates absolute from relative SSA effects. They show that absolute age significantly increases a combined math and reading test score in Grade 8, while relative age has a statistically negative impact on the test score. Our analysis provides estimates for relative and absolute SSA effects by exploiting the fact that several states experienced changes in cut-off dates over time. However, to compare our results to the previous literature, our baseline results provide SSA effects without separating the two effects.

Our analysis shows that the impact of SSA on math and text comprehension measured in adulthood are considerably smaller than what the literature has shown for children in school. Further, both estimates are statistically insignificant. In contrast, the effect of SSA on receptive vocabulary is sizable in adulthood, with a one-year-higher SSA increasing competency by around a third of a standard deviation. These findings survive several tests of robustness. When disentangling the effect of SSA into an absolute and a relative age effect, we find that receptive vocabulary is affected solely by absolute age.

The remainder of the paper is organized as follows. Section 2 gives an overview of the previous literature on SSA effects on test scores and illustrates how the estimated effects differ by grade level. Section 3 describes the data and the school entry regulations, and Section 4 presents the estimation strategy. The fifth section presents and discusses the results, and the final section offers a conclusion.

2. Previous literature

There exists a vast empirical literature that estimates the causal effects of SSA on test scores. The majority of papers rely on instrumental variable strategies for identification. The instruments exploit the variation in individuals' dates of birth and cut-off rules. Both determine at which age a student should legally start school. Alternatively, some studies use the information on date of birth and the cut-off rules for implementing a regression discontinuity design. The previous literature uses test scores measured at different grade levels as outcomes. The earliest test score differences are measured during kindergarten, where the variation derives from kindergarten entry cut-off dates. Several studies show that the oldest kindergarten children score significantly better in reading and math tests compared to the youngest children within the class (Datar, 2006; Elder and Lubotsky, 2009; Lubotsky and Kaestner, 2016; Cascio and Schanzenbach, 2016). Being one year older within a kindergarten class increases math and reading/writing test scores by 0.43 to 0.87 and 0.42 to 0.58 of a standard deviation, respectively.

These early advantages of being the oldest when entering school continue into higher classes (Bedard and Dhuey, 2006; Fredriksson and Öckert, 2006; Puhani and Weber, 2007; McEwan and Shapiro, 2008; Ponzo and Scoppa, 2014; Cook and Kang, 2016; Dhuey et al., 2017; Attar and Cohen-Zada, 2017; Koppensteiner, 2018). However, when comparing the magnitude of the relationship, this advantage seems to decrease when children progress through school. In Grades 9 and 10, the test score differential from being one year older at school entry is 0.10 to 0.20 of a standard deviation in math and 0.15 to 0.24 in reading/writing (Smith, 2009; Black et al., 2011; Peña, 2017).

We provide further evidence on how SSA test score differential evolves by graphically illustrating how SSA estimates presented in the literature differ by grade level. The studies considered for the graphs come from an extensive literature search performed using EconLit.[2] Figure 1 shows the relationship for math and reading/writing scores in Panels A and B, respectively. It illustrates that the test score advantage conferred on the oldest students decreases with grade level. It is largest in kindergarten – even more so for math compared to reading/writing. The dashed lines indicate the best fit for the functional relationship between

[2] The search on EconLit was done on November 20th, 2018. It included the following keywords: "school entry age", "kindergarten entry age", "school entrance age", "kindergarten entrance age", "school starting age", "kindergarten starting age", "age at school entry", "age at kindergarten entry", "age at school start", "age at kindergarten start", "enrollment cutoff", "age effect school", "relative age school", and "relative age performance". From the studies found in this manner, we kept only those in which the identification strategy exploits cut-off rules and the dependent variable is a test score (rather than grades given by the teacher). To be able to consider all estimates from these studies in one figure for math and in one figure for reading/writing, the estimates additionally had to fulfil the following criteria: i) they are provided separately for math and reading/writing tests; ii) they are provided separately by grade level; iii) they can be interpreted in terms of standard deviations of the test score's distribution (i.e., summary statistics that at least allowed a corresponding interpretation had to be provided); and iv) the estimates are provided for the entire population of students rather than just separately for subgroups, such as boys and girls. Table A1 in the appendix contains the full list of studies that are used for the figures.

SSA test score differentials and grade levels.[3] The lines suggest a declining but non-linear relationship. Based on the graphs, extrapolation beyond Grade 10 suggests that the math test score differentials will continue to decline, while there is no clear conclusion for reading/writing.

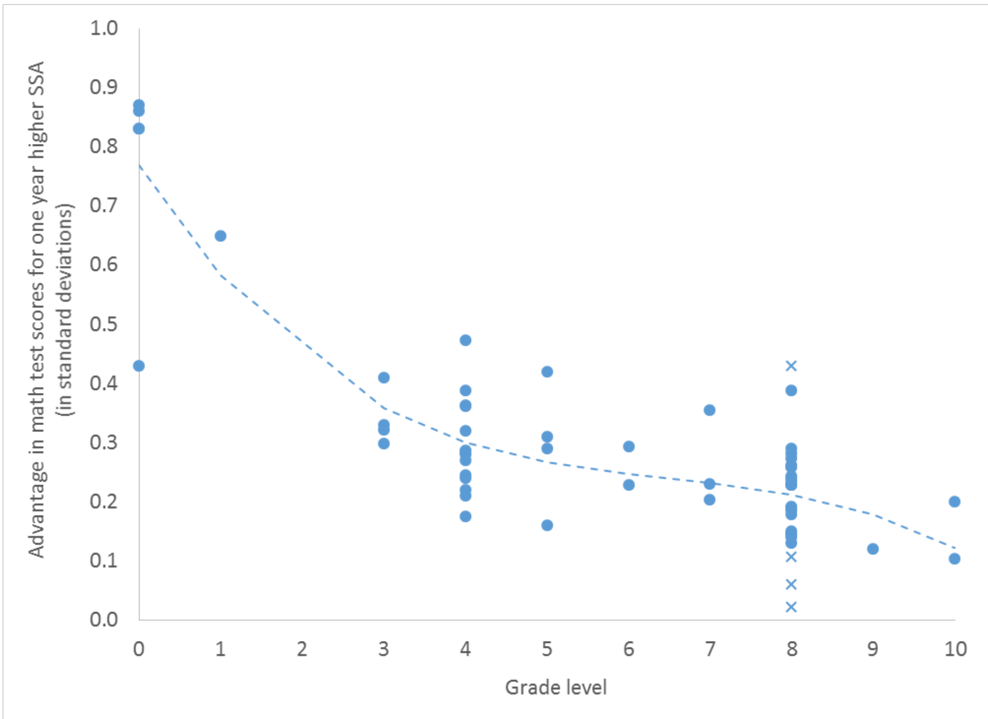
There are only a handful of papers that go beyond Grade 10. However, most of these studies analyze competencies at around the age of 18 and do not extend to higher ages. Further, they sometimes use tests that do not represent competencies but, rather, abilities, or they present their results for a selective sample of students. Therefore, we could not include these results in Figure 1. The following paragraph will summarize these findings.

Using cut-off dates for identification, Black et al. (2011) provide evidence of the impact of SSA on men’s IQ scores measured at 18 years of age. They find relatively small negative effects, i.e., those who are older within class have slightly lower IQ scores. When using IQ scores as an outcome, one caveat is that it is unclear to what extent IQ represents competencies acquired at school and to what extent it represents innate abilities. Cascio and Lewis (2006) provide reduced form estimates of the effect of SSA on AFQT scores of individuals aged 15 to 19. While the effects are not significant for whites, they are slightly negative for blacks. Fletcher and Kim (2016) estimate the reduced form effects of the kindergarten entry cut-offs that differ by state on state-specific averages of the test scores in math and reading. They find no effect of kindergarten entry age on test scores in Grade 12. Nam (2014) exploits cut-off rules using Korean data. The results indicate that SSA has positive effects on math and reading/writing test scores in Grades 6 to 8 but that these differences do not persist when students graduate from high school. Implementing a regression discontinuity design based on cut-off dates, Hurwitz et al. (2015) show that SSA has no significant impact on the SAT scores of college-bound students. However, both Nam (2014) and Hurwitz et al. (2015) analyze test scores of students who voluntarily participated in the test, thus representing a selective sample of students who intend to enroll in college after high school graduation. The same is true for Pellizzari and Billari (2012), who look at the performance of university students. They find that younger students outperform their older fellow students, which they explain via fewer social activities and, thus, more learning time for the younger students. In conclusion, it is an open question whether and how the effects of SSA on competencies linger into adulthood.

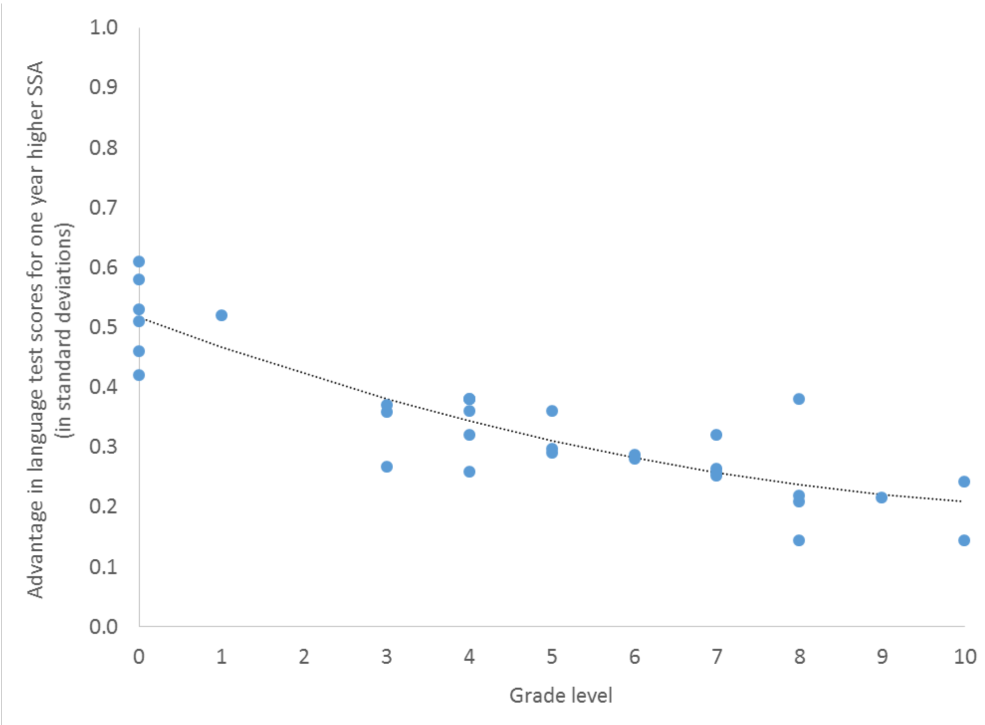
[3] To decide on the functional form of the relationship, we have run linear, quadratic, cubic, and exponential regressions. According to the AIC criteria, a cubic specification best fits the relationship for math scores, and a quadratic model works best for reading/writing.

Figure 1. Effects of school starting age on test scores by grade level

Panel A: Math test scores



Panel B: Reading/writing test scores



Notes: Grade 0 refers to kindergarten. The dots represent statistically significant estimates and the crosses statistically insignificant estimates. The dashed blue line presents the best fit of a model, where grade level enters as a third polynomial. The dotted black line presents the best fit of a quadratic relationship between reading test scores and grade level. The AIC criteria was used to decide on the functional form of the relationships. Table A1 in the appendix shows the full list of studies considered in the figures.

3. Data and school enrollment regulations

3.1 Data

The analysis is based on the adult cohort of the National Educational Panel Study (NEPS-SC6). The NEPS-SC6 includes information on the educational, occupational, and family formation processes for individuals born between 1944 and 1986. It covers detailed information from birth through adult life (Blossfeld et al., 2011), including information on state and date of school entry. This last information allows us to determine accurately the date when children should have entered school according to official regulations and when they have actually done so.

The data also contain information on competencies measured in adulthood. For our analysis, we use one test for mathematical competencies and two tests for language competencies, i.e., for text comprehension and receptive vocabulary. The different competency tests were collected in the NEPS-SC6 data in different waves. The tests were part of the interviews in wave 2010/2011, in which individuals were randomly selected to take the math or the text comprehension test or both, and 2012/2013, in which text comprehension tests were administered to individuals who did not take the test in 2010/2011. In 2014/2015, tests for receptive vocabulary were part of the interview for all respondents. These competency tests capture basic competencies in everyday life situations (Weinert et al., 2011). The math competency tests were designed to describe respondents' abilities to use and apply mathematics flexibly in realistic situations. They cover four content areas: *data and probability*, *quantities*, *shape and space*, and *change and relationship* (Schnittjer and Durchhardt, 2015). To provide one example from the area *data and probability*, respondents are asked whether they understand the statistics on side effects from the package inserts of a pharmaceutical product. The test for text comprehension uses different types of text from which respondents must find information in the text, draw text-related conclusions, and reflect and assess (Gehrer et al., 2012). The receptive vocabulary test is similar to the Peabody Picture Vocabulary Test. Respondents have to assign pictures to a single word given by the interviewer by choosing from four possibilities. The correlation between the three test scores is high, with correlation coefficients varying between 0.47 and 0.55, but there is substantial independent variation in each of these outcomes, i.e., we are guaranteed that the three tests scores measure different dimensions of competencies.[4] In order to ease interpretation of the results, the test scores are normalized to having a mean of zero and a variance of one (z-scores).

Figure A1 in the appendix illustrates the distribution of the three test scores. Even though the tests are designed to capture basic competencies, there is ample variation at the upper and lower end of the competency distributions. The histogram of the receptive vocabulary test score shows that its distribution is highly left-skewed. Therefore, the results section also includes findings for receptive vocabulary that analyze whether this skewness presents a problem. Specifically, we transform the receptive vocabulary test scores using a Box-Cox-transformation. The transformed and standardized test scores have a distribution that is much more similar to a normal distribution (see Figure A2 in the appendix). Because the interpretation

[4] The correlation coefficients are 0.55 for math and text comprehension, 0.47 for math and receptive vocabulary and 0.51 for text comprehension and receptive vocabulary.

of the transformed test scores is less straightforward than that for the untransformed test scores, the main specification for receptive vocabulary is based on the untransformed test scores. However, the robustness section documents that the main results remain unchanged when using the transformed test scores.

The analysis focuses on individuals who entered primary school in West Germany. East Germany (including Berlin) is dropped from the analysis because the East German schooling system differed considerably from the schooling system in the West for this time period. Also, the East German cut-off dates for school entry did not differ between regions and over time, inhibiting separate identification of the effects of age at school entry from month-of-birth effects. The analysis considers information on around 3,700 individuals for mathematical literacy, around 5,900 individuals for text comprehension, and around 6,000 individuals for receptive vocabulary.[5] Descriptive statistics of the three samples are provided in Table 1. Differences between the samples are small, except for age at test, which is due to the fact that the different competencies were assessed in different waves. The average age at test is between 46 and 50 years. Half of the respondents are women. On average, they entered school at 6½ years old and received an average of 13½ years of education.

Table 1. Descriptive statistics

	Sample: Mathematical literacy		Sample: Text comprehension		Sample: Receptive vocabulary	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Female	0.499	0.500	0.498	0.500	0.503	0.500
Age at school entry	6.48	0.49	6.48	0.49	6.48	0.49
Age at test	45.80	11.19	46.68	11.38	49.61	11.09
Years of education	13.58	2.70	13.39	2.79	13.54	2.76
Observations	3,678		5,855		6,053	

Notes: The reason for the differences in the number of observations by competency test is that the tests were taken in different waves.

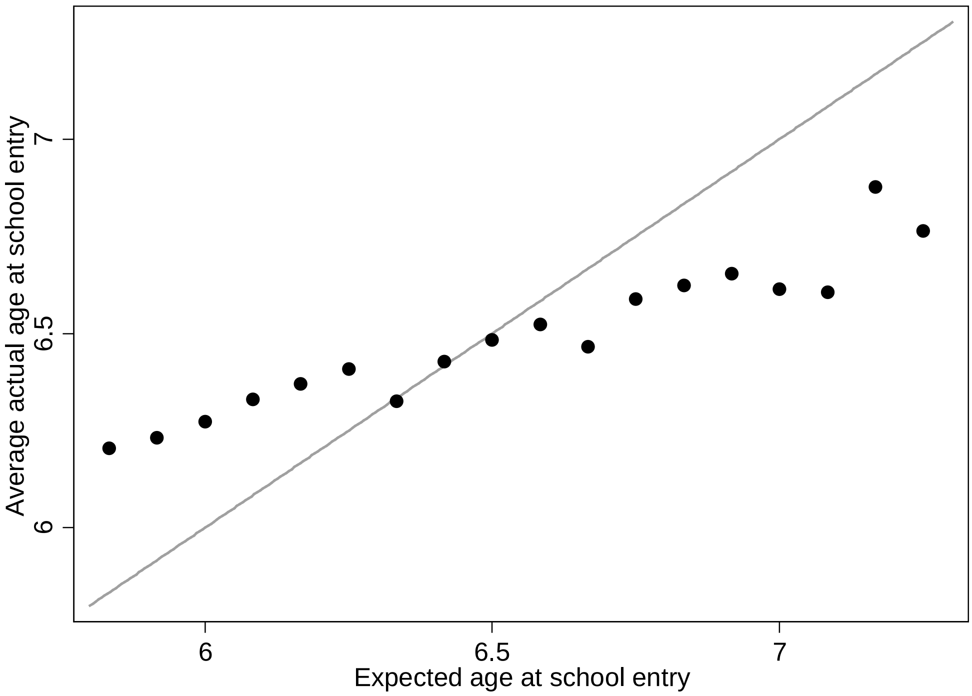
3.2 School enrollment regulations

The analysis exploits variation in school entry cut-off dates that legally define at which age children are supposed to enter school. In Germany, children turning age 6 before the cut-off date enter school at the beginning of the school year, while those turning 6 after the cut-off enter school one year later. The cut-off date is determined by the states. The age at school entry does not depend only on the cut-off date and one's birthday; it also depends on the date when the school year starts, which is also legally determined at the state level. The cut-off date and the beginning of the school year have experienced several changes over time in German states. In recent decades, the school year has generally started in August, while it started in April during the 1950s and 1960s (see Table A2 in the appendix). The cut-off date experienced more changes, varying from the 31st of March to the 31st of December (see Table A3 in the appendix).

[5] The initial West German sample size of individuals participating in the competency tests is 3,766, 5,976, and 6,190, respectively. We had to drop between 1.4% and 1.7% of the observations to ensure that all variables required to construct the instrument are observable and due to data cleaning.

The cut-off is not strictly enforced, as there are options to enroll earlier or to delay school entry, but a sizeable share of parents sticks to the official cut-off. Figure 2 documents how the legally determined expected school starting age is related to the average of the actual SSA. The diagonal line illustrates the case of full compliance, i.e., when all children follow the cut-off rules without exception. The dots show the actual association between expected and average actual school starting age. Even though Figure 2 documents a clear, positive association between the expected and the actual relationship, it also shows that the actual relationship is flatter. This flatness indicates that children with relatively low expected entry ages, i.e., turning 6 years old just before the school cut-off, are, on average, older than 6 when entering school since some children delay entry. Further, children with relatively high expected entry ages, i.e., those turning 7 shortly after the cut-off, are, on average, somewhat younger than 7 when entering school because some children enroll early. This pattern is evidence that parents actually make use of the legal possibilities of delayed and early school entry. Since non-compliance with the cut-off rules occurs and might be selective, exploiting variation from legal cut-off dates to identify the causal effect of age at school entry seems important. Also, note that the range of expected age at school entry in Figure 2 exceeds one year. This is because the minimum expected age at school entry differs between the states in Germany. While some states have official rules that students should be enrolled between the ages of 5.8 and 6.7 (when complying with the rules), others allow the age range to vary from 6.4 to 7.3 years. This variation is exploited in a further analysis to find out whether the estimated differences in competencies by the age at school entry derives from differences in absolute or relative age.

Figure 2. The relationship between the expected and the average of the actual age at school entry



Note: The dots indicate the association between the expected school starting age and the average actual age at school entry. The diagonal line indicates how the association would look if everyone complied with the cut-off rules.

4. Estimation strategy

The causal effect of school starting age is estimated by the following equation:

$$Y_i^k = \beta_0^k + \beta_1^k SSA_i + X_i \delta^k + \varepsilon_i^k \ , \tag{1}$$

where Y_i^k is the standardized test score of individual i on test k (k = mathematical literacy, text comprehension, or receptive vocabulary). SSA_i is school starting age, and X_i is a set of covariates, including gender, dummies for the states of enrollment in primary school, a full set of year-of-birth dummies, and a full set of month-of-birth dummies. Estimating the parameter of interest β_1^k from Equation (1) by OLS would induce biased estimates because the early or late enrollment of children varies with children's abilities and parents' resources in a systematic manner. For instance, Dobkin and Ferreira (2010) show that children of highly educated parents have a lower probability of complying with the enrollment regulations. These identification issues are taken into account by using an instrumental variable (IV) estimator. The instrument is based on the school entry cut-off dates that vary by states and over time. Following the previous literature, the instrument is defined as the expected age at school entry $ESSA_i$. This is the age at which an individual would have entered school if school entry were determined solely by the official regulations. The IV estimator is implemented in the following first-stage equation:

$$SSA_i = \tilde{\gamma}_0 + \tilde{\pi}' ESSA_i + X_i \tilde{\theta} + \tilde{\mu}_i. \tag{2}$$

X_i includes the aforementioned set of covariates. The inclusion of month-of-birth dummies is only possible because the cut-off dates vary over time and between states (see Section 3.2). Few papers can account for month-of-birth effects, although Buckles and Hungerman (2013) find that parents' backgrounds correlate with children's month of birth, even in two consecutive months within the same season. This correlation might lead to systematic differences between those born before and those born after the cut-off. Using only cut-off dates without controlling for the month of birth absorbs these differences, biasing the estimation of the effect of school starting age. Including the year-of-birth dummies serves the purpose of controlling for changes over time that likely correlate with the outcome and the instrument, such as the educational expansion during the 1960s and 1970s. The year-of-birth dummies also capture the influence of age at test in a non-parametric way.[6]

All estimates account for clustering at the level of states. Since only ten clusters can be accounted for, as there are only ten states, the standard errors may suffer from downward bias. We follow suggestions in Cameron et al. (2008) for estimation with few clusters and present small-sample-adjusted p-values of a test against zero instead of the downward biased standard errors. As a test of robustness, we also present p-values of a wild cluster bootstrap, which was found to work well when the number of clusters is small (Cameron et al., 2008).

[6] We apply further robustness tests that include, directly, age at test measured on a monthly basis in the regression. The results remain the same.

First-stage estimates and validity of the identification strategy

Table 2 presents the first-stage estimates. Increasing the expected age at school entry by one year is associated with an average increase of actual age at school entry by 0.38 years. Since the F-test for the significance of the instruments is always considerably above 10, there is no problem of weak instruments (Staiger and Stock, 1997). These results reinforce the conclusions from Figure 2, meaning that the expected school starting age is well suited to serve as an instrument for actual school starting age.

Table 2. The effect of the instrument on school starting age (first stage)

	Sample: Mathematical literacy	Sample: Text comprehension	Sample: Receptive vocabulary
Expected school starting age (ESSA)	0.3755	0.3697	0.3761
	(0.000)	(0.000)	(0.000)
Control variables	Yes	Yes	Yes
F-statistic (excluded instrument)	81.25	133.29	148.19
Observations	3,678	5,855	6,053

Note: The dependent variable is the school starting age. The first row indicates the sample used to analyze each of the three competencies. As mentioned in Section 3.1, using different samples is necessary because each competency was tested in a different survey wave. The control variables include the year of birth, the month of birth, the state of primary school enrollment, and gender. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

To test the validity of the identification strategy, it has been suggested to analyze whether births are systematically displaced around the cut-off (McCrary, 2008). Figure A3 in the appendix shows the number of observations by distance to the cut-off for each of the three samples. No systematic pattern of bringing forward or postponement of births becomes visible.[7] Another test of the validity of the identification strategy is to show that predetermined variables, such as parental characteristics, are not correlated with the instrument. If the instrument were correlated with parental background, this could have a direct effect on competencies. Table 3 shows, separately for the three samples, that the family status at the age of 15 (i.e., whether the individual was raised by a single parent) and mother's and father's education, age at birth, and migration status are all unrelated to the instrument. This suggests that parents do not strategically plan to deliver children before or after the school cut-off date in Germany, which is also confirmed by Bahrs and Schumann (2016). In order to test the robustness of our main results, Section 5 also presents IV estimates controlling for the above-mentioned parental characteristics.

[7] For an example of systematic timing of the date of delivery with respect to a policy reform in Germany, see Tamm (2013).

Table 3. The effect of school starting age on predetermined characteristics (IV results)

	Single-parent family at age 15	Mother with college degree	Father with college degree	Mother's age at birth	Father's age at birth	Mother foreign born	Father foreign born
Sample: mathematical literacy							
IV estimate: SSA	-0.0284	0.0056	0.0383	1.8301	2.0310	-0.0156	-0.0194
	(0.436)	(0.854)	(0.526)	(0.260)	(0.253)	(0.733)	(0.562)
Observations	3,674	3,586	3,553	3,573	3,535	3,657	3,606
Sample: text comprehension							
IV estimate: SSA	0.0692	0.0095	0.0605	0.3586	0.9372	-0.0129	-0.0404
	(0.150)	(0.561)	(0.223)	(0.433)	(0.444)	(0.729)	(0.291)
Observations	5,848	5,674	5,627	5,687	5,627	5,823	5,736
Sample: receptive vocabulary							
IV estimate: SSA	0.0482	0.0220	0.0976	0.2359	1.4198	-0.0435	0.0075
	(0.129)	(0.341)	(0.161)	(0.740)	(0.287)	(0.093)	(0.734)
Observations	6,045	5,870	5,819	5,884	5,813	6,018	5,933

Note: The table provides IV estimates of the effect of school starting age on the outcomes listed in the first row. ESSA is used as the instrument. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

5. Results

Table 4 presents the main IV estimates in Panel A. The first column of Table 4 presents the results for mathematical literacy. The estimate is statistically insignificant and small in size, suggesting that SSA has no long-lasting effects on mathematical competencies that persist until adulthood. The point estimate for an increase of age at school entry by one year relates to 6% of a standard deviation, which is far below what was measured in Grade 10 (see Panel A of Figure 1). The second column of Table 4 documents the results using text comprehension as an outcome. Similar to the results for mathematical literacy, the estimate for text comprehension presented in Panel A is statistically insignificant, and it is much smaller than the effects that were found in the literature for tests taken at younger ages. Specifically, the point estimate relates to 8% of a standard deviation. The third column of Table 4 shows results using receptive vocabulary as an outcome. In contrast to the other two measures of competency, the impact of the estimate of school starting age on receptive vocabulary is significantly positive. An increase of age at school entry by one year is associated with an increase in test scores by a third of a standard deviation. This is an economically sizable effect.

Before discussing reasons for why the effect of SSA on the Peabody Picture test persists, while it vanishes for math skills and text comprehension, we present several robustness checks. Panel B of Table 4 shows that the results remain robust when controlling for parental background characteristics. This is not surprising given that Table 2 has already shown that parental background is uncorrelated with the instrument. We also find that controlling for parental background increases the precision of the estimates. This leads to lower p-values for the estimate of school starting age on receptive vocabulary but leaves our conclusion of insignificant results for math skills and text comprehension unchanged. As has already been discussed in Section 2, some recent studies have shown that controlling for age at test matters considerably when estimating the effects of SSA on competency tests (see, e.g., Black et al., 2011). In our main specification, age at test is partially controlled for by the year-of-birth dummies. Panels C and D directly control for age at test measured on a monthly basis to capture even small differences in age at test. Small age differences within a given cohort occur because the interviews were conducted in different months over the year and because some individuals participated in the text comprehension test in wave 2010/2011, while others participated in wave 2012/2013. Panels C and D present the results of a linear specification for age at test and of a squared specification, respectively. Both specifications still control for the year-of-birth dummies. Results in Panels C and D are very similar to those of the main specification.

States that implement reforms to the schooling system, such as changes in the cut-off date, might experience different trends in the outcome variable even without reforms. To rule out that such state-specific changes over time affect our estimates, Panel E controls for state-specific time trends using a linear specification, and Panel F controls for state-specific time trends using a squared specification. Controlling for time trends has only a minor impact on the estimates of SSA. Panel G of Table 4 shows results excluding the most recent cohorts. We do so because most of the changes in cut-off dates and the school start month took place during the 1950s and 1960s. In order to not rely on cohorts that are relatively far away from these reforms, the specification in Panel G is restricted to cohorts born between 1944 and 1973, dropping those born between 1974 and 1986.[8] The results for this reduced sample are very similar to those for the entire sample.

Panels H and I present results for alternative definitions of the instrument. While Panel A includes a linear specification for ESSA as an instrument for SSA, Panel H additionally includes ESSA-squared as an instrument. Furthermore, Panel I uses separate dummies for each value of ESSA as an instrument for SSA. These alternative definitions of the instrument generally confirm our main findings of statistically significantly effects of SSA on receptive vocabulary, although the point estimate is reduced by around one third (from 0.35 to 0.21) in Panel I.

Panel J shows that the results are robust to calculating the standard errors by alternative methods. As is pointed out in Section 4, all estimates account for clustering at the level of the states. Because the number of clusters is small, we follow Cameron et al. (2008) by presenting small-sample-adjusted p-values in our main specification. An alternative procedure for obtaining inference is using wild cluster bootstrapping, which is implemented in Panel J. The conclusion remains unchanged, i.e. the IV estimate for receptive vocabulary is statistically significant, while the estimates for mathematical literacy and for text comprehension remain statistically insignificant.

[8] The last cohort affected by reforms is the birth cohort of 1963. 1973 is chosen to include, at most, ten birth cohorts afterwards.

Table 4. The effect of school starting age on competencies in adulthood

	Mathematical literacy (1)	Text comprehension (2)	Receptive vocabulary (3)
Panel A Main specification			
IV estimate: SSA	0.0648	0.0818	0.3461
	(0.795)	(0.644)	(0.021)
F-statistic (excluded instrument)	81.25	133.29	148.19
Observations	3,678	5,855	6,053
Panel B Controlling for parental background characteristics			
IV estimate: SSA	0.0492	0.0994	0.3797
	(0.839)	(0.573)	(0.009)
F-statistic (excluded instrument)	78.99	121.89	126.57
Observations	3,678	5,855	6,053
Panel C Controlling for age at test in addition to the year-of-birth dummies			
IV estimate: SSA	0.0628	0.0849	0.3462
	(0.800)	(0.649)	(0.021)
F-statistic (excluded instrument)	81.30	132.70	148.14
Observations	3,678	5,855	6,053
Panel D Controlling for age at test and its square in addition to the year-of-birth dummies			
IV estimate: SSA	0.0422	0.0864	0.3668
	(0.873)	(0.653)	(0.028)
F-statistic (excluded instrument)	94.73	205.61	171.34
Observations	3,678	5,855	6,053
Panel E Controlling for state-specific time trends			
IV estimate: SSA	0.0712	0.0765	0.3141
	(0.767)	(0.643)	(0.043)
F-statistic (excluded instrument)	80.18	132.95	155.91
Observations	3,678	5,855	6,053
Panel F Controlling for state-specific time trends using a quadratic specification			
IV estimate: SSA	0.0742	0.0809	0.2710
	(0.771)	(0.612)	(0.075)
F-statistic (excluded instrument)	77.14	131.06	153.7
Observations	3,678	5,855	6,053
Panel G Sample 1944-1973			
IV estimate: SSA	0.1000	0.1015	0.2907
	(0.707)	(0.584)	(0.015)
F-statistic (excluded instrument)	71.61	106.66	127.47
Observations	3,028	4,799	4,984
Panel H Alternative definition of the instrument: ESSA + ESSA-squared (as additional instrument)			
IV estimate: SSA	0.061	0.0497	0.3091
	(0.809)	(0.772)	(0.047)
F-statistic (excluded instruments)	40.74	69.20	77.21
Observations	3,678	5,855	6,053
Panel I Alternative definition of the instrument: Separate dummies for each value of ESSA			
IV estimate: SSA	-0.0498	-0.0122	0.2110
	(0.803)	(0.934)	(0.054)
F-statistic (excluded instruments)	7.70	80.35	131.60
Observations	3,678	5,855	6,053
Panel J Wild cluster bootstrap			
IV estimate: SSA	0.0648	0.0818	0.3461
	[0.904]	[0.667]	[0.096]
F-statistic (excluded instrument)	81.25	133.29	148.19
Observations	3,678	5,855	6,053

Note: The table presents results of IV regressions that use ESSA as an instrument for SSA (see equations 1 and 2). The outcome variables are documented in the first row. Panel B controls for the parental background characteristics that are shown in Table 3. Standard errors account for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses. Panel J presents p-values using wild cluster bootstrap that are shown in brackets.

Finally, as has been discussed in Section 3.1 and shown in Figure A1 in the appendix, the distribution of the receptive vocabulary test score is highly left-skewed. Therefore, Table 5 presents results that use transformed test scores as the outcome variable. The transformed receptive vocabulary test scores have a distribution that is much more similar to a normal distribution. The IV estimates remain the same when using this alternative outcome. Similar to the case for the untransformed receptive vocabulary test scores, the effect of SSA is statistically significant, and the effect size is close to a third of a standard deviation (of the transformed test score). As an alternative to using transformed test scores, we also tested whether the results are robust to dropping individuals at the extreme lower end of the untransformed test score distribution. Results are generally not affected by this. For example, if we drop individuals with an untransformed z-score of below -4, the IV estimate is 0.315. If we drop individuals with an untransformed z-score of below -2, the IV estimate is 0.252. Both estimates are statistically significant.

Table 5. Robustness check using the transformed receptive vocabulary test score

	Language: transformed receptive vocabulary
IV estimate: SSA	0.3496
	(0.034)
F-statistic (excluded instruments)	148.19
Observations	6,053

Note: The table presents results of IV regressions that use ESSA as an instrument for SSA (see equation 1 and 2). The outcome variable is generated using a Box-Cox-transformation $y_i^{\text{transformed}} = (y_i^\lambda - 1)/\lambda$ *with Lambda equal to 4.22 and then standardized to having a mean of zero and a variance of one. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.*

Discussion of the results

Our results show that the effects of school starting age fade away in adulthood for math competencies and for text comprehension but remain significant for receptive vocabulary. This subsection discusses why the long-term impact of SSA differs between competencies. First, we look at the pathways of students in the schooling and vocational systems. This is highly important in the German case since the educational system is characterized by early tracking in school and a strong apprenticeship system. Second, we tie those results together with findings from the literature concerning how individuals acquire competencies. In particular, we discuss how acquiring competencies differs by the particularities of the German educational system.

In Germany, students are assigned to different school tracks at the age of 10 depending on their abilities and educational performance in primary school. The tracks differ in curriculum, degree of difficulty and abstraction of the material that is covered, in the highest degree attained when leaving school, and in the length of schooling. For the cohorts in our data set, attaining a school degree from the lowest track required 8 or 9 years of schooling (*Hauptschule*); from the middle track, 10 years (*Realschule*); and from the highest track, 13 years (*Gymnasium*). Using data from one out of ten West German states, Puhani and Weber (2007) show that the school starting age has a substantial impact on the assignment of tracks in Germany, meaning that a higher SSA leads to choosing a higher track. Based on the representative NEPS data, we answer this question using all West German states by estimating our main IV model with different educational outcomes as dependent variables, such as the highest school degree and the highest vocational degree attained. In order to save space, we will not provide results for each of the three samples used in the analysis. Instead, the analysis focuses on the sample for which the receptive vocabulary scores are available, which is not only the sample for which we have found significant results, but it is also the sample with the largest number of observations.

Table 6. IV estimates of school starting age on schooling

	Years of schooling (1)	Highest school degree: low (2)	Highest school degree: middle (3)	Highest school degree: high (4)
IV estimate: SSA	0.5102	-0.1488	0.0283	0.1205
	(0.020)	(0.044)	(0.539)	(0.021)
F-statistic (excluded instruments)	151.0719	151.07	151.07	151.07
Observations	5,920	5,920	5,920	5,920

Note: The table provides IV estimates of the effect of school starting age on the outcomes listed in the first row. The instrument used is ESSA. The estimation sample comprises individuals for whom a valid test score for receptive vocabulary is available. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

Table 7. IV estimates of school starting age on vocational education

	Highest vocational degree: none (1)	Highest vocational degree: apprenticeship (2)	Highest vocational degree: college (3)
IV estimate: SSA	-0.0340	0.0138	0.0202
	0.071	0.787	0.690
F-statistic (excluded instruments)	135.01	135.01	135.01
Observations	5,681	5,681	5,681

Note: The table provides IV estimates of the effect of school starting age on the outcomes listed in the first row. The instrument used is ESSA. The estimation sample comprises individuals for whom a valid test score for receptive vocabulary is available. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

Table 6 provides the results of the influence of age at school entry on years of schooling and on the highest school degree attained. Overall, they indicate that being one year older at school entry increases years spent in school by approximately half a year (Column 1). Columns 2 to 4 show that a one-year-higher SSA decreases the probability of attaining the lowest school degree by 15% and increases the probability of attaining the highest school degree by 12%. These findings are very similar to those in Dustmann et al. (2017) on school track at the age of 14.[9] Table 7 shows the results for vocational education: distinguishing no vocational degree,

[9] At first glance, our IV estimates appear to be considerably larger than the effects presented in Dustmann et al. (2017), but note that Dustmann et al. (2017) present reduced-form estimates, while we present IV estimates. Our reduced-form estimates are -5.8 ppt for attaining the lowest schooling degree and +4.7 ppt for attaining the highest schooling degree. Table 2 in Dustmann et al. (2017) compares individuals born in different months that differ, on average, in expected school entry age by 0.5 and 0.91 years, respectively. If we rescale the respective estimates to represent one-year differences, the effect sizes from Dustmann et al. (2017) range between -7.8 ppt and -3.8 ppt

having completed an apprenticeship, and obtaining a college degree. SSA only affects the probability of having obtained no vocational degree by 3%, which is much smaller in magnitude than the effect of SSA on the schooling degree. In contrast, SSA has no statistically significant impact on completing an apprenticeship or college.[10] We suggest that the reason for observing SSA effects mostly on school track choice, but not on the highest educational degree, is that SSA effects mirror maturity differences that are biggest, when children are young. When becoming older, maturity becomes less important for explaining educational success and students' real potential becomes visible. The German schooling system is flexible by allowing high ability students from middle school to upgrade their skills after leaving school (Dustmann et al. 2017). These students can even study, if they fulfill some requirements. In addition, high school graduates who have learned of their potential to be lower than expected can abstain from enrolling at university.

Given that we mainly find SSA effects on track choice, we hypothesize that the division into the different school tracks is the potential mechanism for the effects of SSA to persist into adulthood, while selection into different vocational tracks is less important. To shed further light on this hypothesis, we once again estimate the impact of SSA on the receptive vocabulary score but now control additionally for the potential educational mechanisms. Columns 1 and 2 of Table 8 control for the highest schooling degree attained and the highest vocational degree, respectively. These findings illustrate that once we control for the schooling degree, the effect of SSA on receptive vocabulary vanishes almost completely. The point estimate drops from 35% of a standard deviation in our main specification to 7% and is no longer statistically significant. In contrast, results in Column 2 show that controlling for the highest vocational degree has a considerably smaller impact. The effect size decreases only modestly; it is still sizeable at 28% of a standard deviation, and it is statistically significant at the 5% level. These results are indicative evidence that SSA affects the track assignment by which it has a long-term impact on receptive vocabulary. However, why does this conclusion not also hold for the other competencies?

Table 8. IV estimates of SSA on receptive vocabulary while controlling for mechanisms

	Language: receptive vocabulary (1)	Language: receptive vocabulary (2)
IV estimate: SSA	0.0773	0.2739
	(0.485)	(0.019)
Additional controls for the highest school degree	Yes	No
Additional controls for the highest vocational degree	No	Yes
F-statistic (excluded instruments)	152.10	134.91
Observations	5,920	5,681

Note: The table provides IV estimates of the effect of school starting age on receptive vocabulary. The instrument used is ESSA. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

for attending the lowest school track and between +4.0ppt and +6.6ppt for attending the highest school track. Our reduced form estimates fall exactly into these intervals.

[10] Dustmann et al. (2017) also find that the effects of SSA are much smaller when considering completed education.

Pischke and von Wachter (2008) state that basic math and reading/writing skills are taught in secondary school in Germany, regardless of the track choice. This means that basic skills are taught to all students. The difference between the lower and higher tracks in this regard is mainly that higher-track students learn more advanced and academic knowledge. Given the focus of the competency tests on basic skills and their application to everyday problems, advanced knowledge is not captured in the competency tests available in the NEPS data. The basic skills are also necessary for participating successfully and productively in the German labor market (Pischke and von Wachter, 2008). Thus, we consider it plausible that these skills are used and, thereby, practiced regularly after leaving school.

In contrast, receptive vocabulary is learned by exposure to oral or written language.[11] School tracks might differ in terms of vocabulary growth due to students being exposed differently to language activities. Students in the highest track have more opportunities to engage in these activities because they stay in school longer, but also because reading texts or books (including classical literature) is much more frequent in the highest track. The literature has also shown that the number of different and rare words in texts and books matters in terms of increasing individuals' vocabulary growth. The largest variety can be observed in scientific texts (Hayes and Ahrens, 1988). Due to its more academic curriculum, graduates from the highest track are likely to have been exposed to more academic words than lower- or medium-track students.

In a nutshell, school starting age affects the track assignments. The academic track puts much more weight on developing language skills than the middle and the lower school tracks, which might result in a more-refined and larger set of vocabulary used by academic track students. In contrast, there is evidence that academic-track schools are not superior to middle- and lower-track schools when it comes to generating the basic skills gauged in the mathematical literacy and text comprehension tests. The next section is devoted to answering what policy can do to counteract the long-lasting SSA effects on receptive vocabulary.

Absolute vs. relative age effects

As discussed in Cascio and Schanzenbach (2016), the effects documented in most of the literature and in our main results are a mixture of absolute age at school entry and of relative age. Separate identification is important from a policy perspective. For example, take a policy that changes the cut-off date by one month. If only absolute age is important, this policy does not have any impact on those further away from the cut-off, but only on those whose birthday falls between the old and the new cut-off. If those between the old and the new cut-off benefit from the reform, there will also be an improvement when looking at the aggregate of children because no child will be disadvantaged by the reform. In contrast, if relative age is important, this policy also has an impact on children further away from the cut-off because the policy affects the average age of classmates and, thus, the relative age of each student. Furthermore, if only relative age is important and not absolute age, this policy will only influence which child

[11] The scientific literature has not yet reached consensus on whether vocabulary growth occurs mainly incidentally through conversations and reading (Stenberg, 1987) or whether it is transmitted through explanations by teachers or parents or within texts (Bielmiller, 2001).

is the oldest and which child is the youngest in the class and, thus, who is hit by the negative effect of being (relatively) young, but it has no effect on the aggregate of children.

Focusing on receptive vocabulary, Table 9 provides separate estimates for absolute and relative age effects. Separate estimates are possible because we have variation in cut-off dates between states and over time. This allows us to define separately instruments for age at school entry and for age relative to the child that should be the youngest within a cohort, given all children follow the regulations. However, both instruments are highly correlated, with a correlation coefficient of 0.91. This correlation makes separate identification somewhat problematic, and we emphasize that the separate effects should not be over-interpreted. According to our findings, only absolute age effects are relevant for receptive vocabulary. In contrast, the point estimate of relative age has a negative sign, is much smaller and is far from being statistically significant. This finding is similar to that of Cascio and Schanzenbach (2016), who also report that absolute age is more important than relative age.

Table 9. IV estimates of absolute and relative school starting age

	Language: receptive vocabulary
IV estimate: absolute SSA	0.4148
	(0.039)
IV estimate: relative SSA	-0.0933
	(0.372)
F-statistic (excluded instruments)	72.13
Observations	6,053

Note: The table provides IV estimates of the effects of absolute and of relative school starting age on receptive vocabulary. Estimation accounts for clustering at the state level. Small-sample-adjusted p-values are shown in parentheses.

6. Conclusion

The previous literature has shown that scores of competency tests administered to school children are influenced by the age of children at school entry. While there is evidence that these effects become smaller as the children grow older, little is known about whether the effects fade away completely or remain important long after leaving school. We find no evidence that the effects of school starting age on math competencies and text comprehension are still relevant in adulthood, although they are considerable when children are in school. These results are also in line with the previous literature, which has shown no or only small long-term effects of SSA on wages and employment (e.g., Fredriksson and Öckert, 2014; Larsen and Solli, 2017; Dustmann et al., 2017). Assuming that basic competencies in math and text comprehension matter for labor market success, the absence of long-run SSA effects on these competencies could explain the absence of long-run effects of SSA on labor market success. Hence, we conclude that our results do not provide reasons for policy actions.

In contrast, for receptive vocabulary, the effect of school starting age remains large and statistically significant in the longer run. Our findings suggest that the long-run effect on receptive vocabulary is due to Germany's tracking system, which sorts children into different school tracks at an early age. Given that, this effect is due to absolute rather than to relative age, recent policies that shift the cut-off to an earlier date, making some children older by one year at school entry, should lead to an improvement in the average receptive vocabulary competencies even in adulthood. Of course, the benefits of such reforms need to be contrasted with their social costs (e.g., the cost of one more year in childcare) and private costs (e.g., of entering the labor market one year later).

Literature

Attar, I. and D. Cohen-Zada (2018), The Effect of School Entrance Age on Educational Outcomes: Evidence Using Multiple Cutoff Dates and Exact Date of Birth. Journal of Economic Behavior and Organization 153, 38-57.

Bahrs, M. and M. Schumann (2016), Unlucky to be young? The long-term effects of school starting age on smoking behavior and health. HCHE Research Paper #13.

Bedard, K. and E. Dhuey (2006), The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. Quarterly Journal of Economics 121(4), 1437-1472.

Biemiller, A. (2001), Teaching vocabulary: Early, direct and sequential. The American Educator 25(1), 24-28.

Black, S., P. Devereux and K. Salvanes (2011), Too young to leave the nest? The effects of school starting age. Review of Economics and Statistics 93(2), 455–67.

Blossfeld, H.-P., H.-G. Roßbach and J. von Maurice (eds.) (2011), Education as a Lifelong Process – The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft, Special Issue 14.

Buckles, K. and D. Hungerman (2013), Season of birth and later outcomes: Old questions, new answers. Review of Economics and Statistics 95(3), 711-724.

Cameron, C., J. Gelbach and D. Miller (2008), Bootstrap-based improvements for inference with clustered errors. Review of Economics and Statistics 90(3), 414-427.

Cascio, E. and E. Lewis (2006), Schooling and the Armed Forces Qualifying Test: Evidence from School-Entry Laws. Journal of Human Resources 41(2), 294-318

Cascio, E. and D. Schanzenbach (2016), First in the Class? Age and the Education Production Function. Education Finance and Policy 11(3), 225–250.

Cook, P. and S. Kang (2016), Birthdays, Schooling, and Crime: Regression-Discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation. American Economic Journal: Applied Economics 8(1), 33–57.

Cook, P. and S. Kang (2018), The school-entry-age rule affects redshirting patterns and resulting disparities in achievement. NBER Working Paper 24402.

Datar, A. (2006), Does delaying kindergarten entrance give children a head start? Economics of Education Review 25(1), 43–62.

Deming, D. and S. Dynarski (2008), The Lengthening of Childhood. Journal of Economic Perspectives 22(3), 71–92.

Dhuey, E., D. Figlio, K. Karbownik, and J. Roth (2017), School Starting Age and Cognitive Development. NBER Working Paper 23660.

Dobkin, C. and F. Ferreira (2010), Do school entry laws affect educational attainment and labor market outcomes? Economics of Education Review 29, 40–54.

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

Elder, T. and D. Lubotsky (2009), Kindergarten entrance age and children’s achievement: Impacts of state policies, family background, and peers. Journal of Human Resources 44(3), 641–83.

Fletcher, J. and T. Kim (2016). The effects of kindergarten entry age policies on educational achievement. Economics of Education Review 50, 45–62.

Fredriksson, P. and B. Öckert (2014), Life-cycle effects of age at school start. Economic Journal, 124, 977-1004.

Gehrer, K., S. Zimmermann, C. Artelt and S. Weinert (2012), The Assessment of Reading Competence. Bamberg: University of Bamberg, National Educational Panel Study.

Hayes, D. P. and M. Ahrens (1988), Vocabulary simplification for children: A special case of ‘motherese’. Journal of Child Language 15, 395-410.

Hurwitz, M., J. Smith, and J. Howell (2015), Student Age and the Collegiate Pathway. Journal of Policy Analysis and Management 34(1), 59–84.

Koppensteiner, M. (2018), RelativeAge, Class Assignment, and Academic Performance: Evidence from Brazilian Primary Schools. Scandinavian Journal of Economics 120(1), 296–325.

Landersø, R., H. Skyt Nielsen and M. Simonsen (2017), School Starting Age and the Crime-age Profile. Economic Journal 127, 1096-1118.

Larsen, E. and I. Solli (2017), Born to run behind? Persisting birth month effects on earnings. Labour Economics 46, 200-210.

Lefgren, L. and F. McIntyre (2006), The relationship between women’s education and marriage outcomes. Journal of Labour Economics 24(4), 787-830.

Lubotsky, D. and R. Kaestner (2016), Do ‘Skills Beget Skills’? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood. Economics of Education Review 53, 194–206.

McCrary, J. (2008), Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142, 698-714.

McCrary, J. and H. Royer (2011), The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. American Economic Review 101(1), 158-195.

McEwan, P. and J. Shapiro (2008), The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. Journal of Human Resources 43(1), 1–29.

Nam, K. (2014), Until when does the effect of age on academic achievement persist? Evidence from Korean data. Economics of Education Review 40, 106–122.

Pellizzari, M. and F. Billari (2012), The younger, the better? Age-related differences in academic performance at university. Journal of Population Economics 25(2), 697–739.

Peña, P. (2017), Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood. Economics of Education Review 56, 152–176.

Pischke, S. and T. von Wachter (2008), Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. Review of Economics and Statistics 90(3), 592-598.

Ponzo, M. and V. Scoppa (2014), The long-lasting effects of school entry age: Evidence from Italian students. Journal of Policy Modeling 36(3), 578–599.

Puhani, P. and A. Weber (2007), Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Early Educational Effects of Age of School Entry in Germany. Empirical Economics 32, 359-386.

Schnittjer, I. and C. Durchhardt (2015), Mathematical Competence: Framework and Exemplary Test Items. Bamberg: University of Bamberg, National Educational Panel Study.

Skirbekk, V., H. Kohler and A. Prskawetz (2004), Birth month, school graduation, and the timing of births and marriages. Demography 41(3), 547-568.

Smith, J. (2009), Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement. The B.E. Journal of Economic Analysis & Policy 9(1), 1–30.

Staiger, D. and J. Stock (1997), Instrumental variables regression with weak instruments. Econometrica 65(3), 557-586.

Statistisches Bundesamt (2017), Bildung und Kultur. Allgemeinbildende Schulen. Schuljahr 2015/16. Fachserie 11, Reihe 1. Wiesbaden.

Statistisches Bundesamt (2018), Bildung und Kultur. Allgemeinbildende Schulen. Schuljahr 2017/18. Fachserie 11, Reihe 1. Wiesbaden.

Sternberg, R. J. (1987), Most vocabulary is learned from context. In: M. G. McKeown and M. E. Curtis (Eds.), The nature of vocabulary acquisition, 89-106. Hillsdale, N.J.: Erlbaum.

Tamm, M. (2013), The Impact of a Large Parental Leave Benefit Reform on the Timing of Birth around the Day of Implementation. Oxford Bulletin of Economics and Statistics 75(4), 585-601.

Weinert, S., C. Artelt, M. Prenzel, M. Senkbeil and C. Carstensen (2011), Development of competencies across the life span. Zeitschrift für Erziehungswissenschaft 14(2), 67-86.

Appendix

Table A1. Studies considered in Figure 1

	Grade level										
	0	1	2	3	4	5	6	7	8	9	10
	Math test scores										
Attar and Cohen-Zada (2018)						x			x		
Bedard and Dhuey (2006)				x	11x				18x		
Black et al. (2011)											x
Cook and Kang (2018)				x							
Datar (2006)	x										
Elder and Lubotsky (2009)	2x	x		x		x			x		
Koppensteiner (2018)						x					
Lubotsky and Kaestner (2016)	2x										
McEwan and Shapiro (2008)					x				x		
Nam (2014)							x	x	4x		
Peña (2017)				x	x	x	x	x	x	x	
Ponzo and Scoppa (2014)					x				x		
Smith (2009)					x			x			x
	Reading/writing test scores										
Attar and Cohen-Zada (2018)						x			x		
Cook and Kang (2018)				x							
Datar (2006)	x										
Elder and Lubotsky (2009)	2x	x		x		x			x		
Lubotsky and Kaestner (2016)	3x										
McEwan and Shapiro (2008)					x						
Nam (2014)							x	x	x		
Peña (2017)				x	x	x	x	x	x	x	
Ponzo and Scoppa (2014)					x						
Puhani and Weber (2007)					x						
Smith (2009)					2x			2x			2x

Note: Those studies that include several estimates per grade level report either separate estimates for different countries, such as in Bedard and Dhuey (2006), or different test scores, such as separate tests for reading and writing.

Table A2. Month of school start by state

State	Month of school start
BW	1950-1951: September, 1952-1966: April, 1966-1994: August
BY	1950-1994: August
HB	1950-1966: April, 1967-1994: August
HH	1950-1966: April, 1967-1994: August
HE	1950-1966: April, 1966-1994: August
NI	1950-1966: April, 1967-1994: August
NW	1950-1966: April, 1966-1994: August
RP	1950-1966: April, 1966-1994: August
SL	1950-1958: August, 1959-1966: April, 1966-1994: August
SH	1950-1966: April, 1967-1994: August
BE	1950-1951: August, 1952-1966: April, 1967-1994: August

Source: State-specific laws and legislation determining the month of school start.
Note: In some states, the school year started two times during 1966 (short school year).

Table A3. Cut-off date by state

State	Cut-off date
BW	1950: 31.12., 1951: 31.5., 1952: 31.3., 1953-1963: 15.4., 1964-1966: 31.12., 1966-1994: 30.6.
BY	1950-1968: 30.9., 1969-1994: 30.6.
HB	1950-1965: 31.3., 1966: 31.5. & 30.11., 1967: 1.7., 1968-1994: 30.6.
HH	1950-1961: 31.3., 1962-1966: 31.12., 1967-1994: 30.6.
HE	1950-1956: 30.6., 1957-1961: 31.3., 1962-1965: 31.12., 1966: 31.3. & 30.11., 1967-1994: 30.6.
NI	1950-1955: 30.6., 1956-1966: 31.3., 1967-1994: 30.6.
NW	1950-1960: 30.6., 1961-1965: 31.3., 1966: 31.3. & 30.11., 1967-1994: 30.6.
RP	1950-1952: 30.6., 1953-1965: 31.3., 1966: 31.3. & 30.11., 1967-1994: 30.6.
SL	1950-1954: 31.12., 1955-1957: 30.9., 1958: 31.12., 1959-1956: 31.3., 1966: 31.3. & 31.12., 1967: 30.9., 1968-1994: 30.6.
SH	1950-1955: 30.6., 1956-1963: 31.3., 1964-1965: 31.12., 1966: 31.12. & 30.11., 1967-1994: 30.6.
BE	1950-1951: 31.12., 1952-1955: 30.6., 1956-1966: 31.3., 1967-1994: 30.6.

Source: State-specific laws and legislation determining the cut-off date of school entry for children.
Note: In some states, the school year started two times during 1966, resulting in two cut-off dates.

Figure A1. Distribution of the standardized test scores

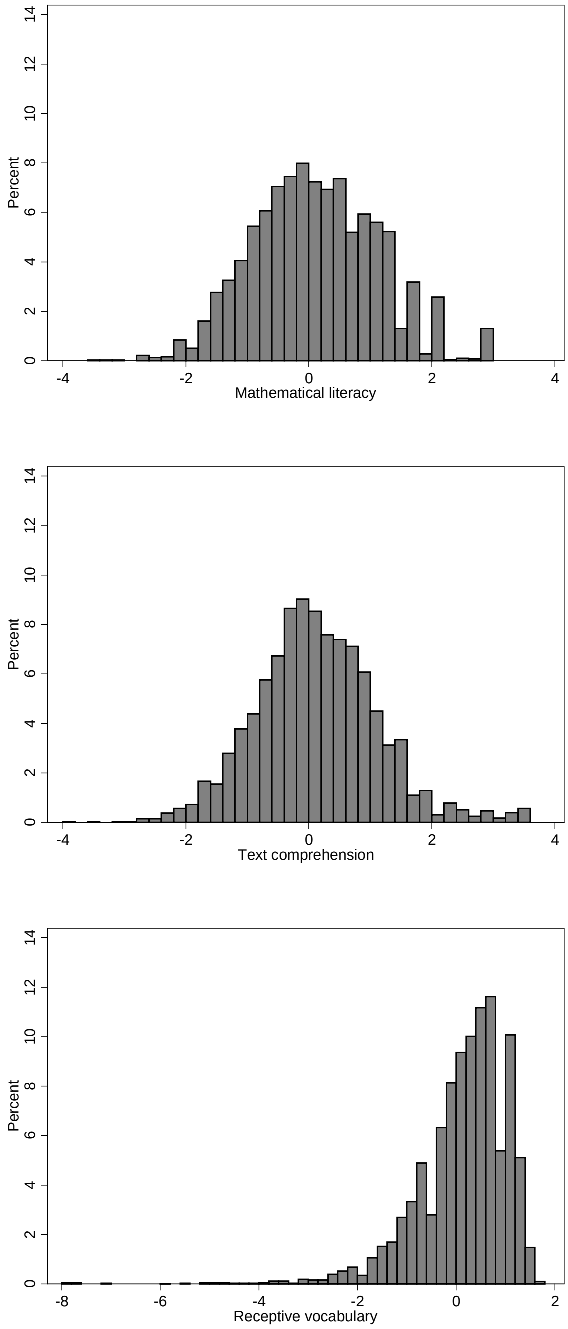
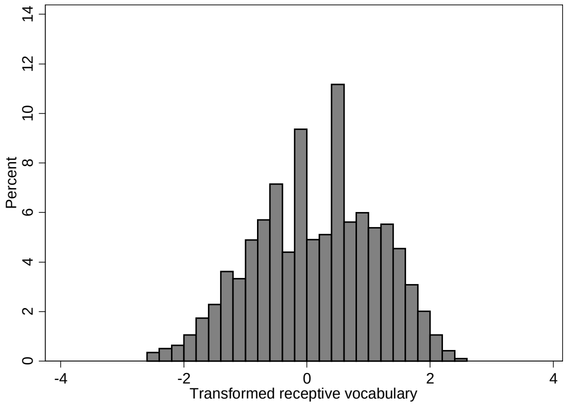
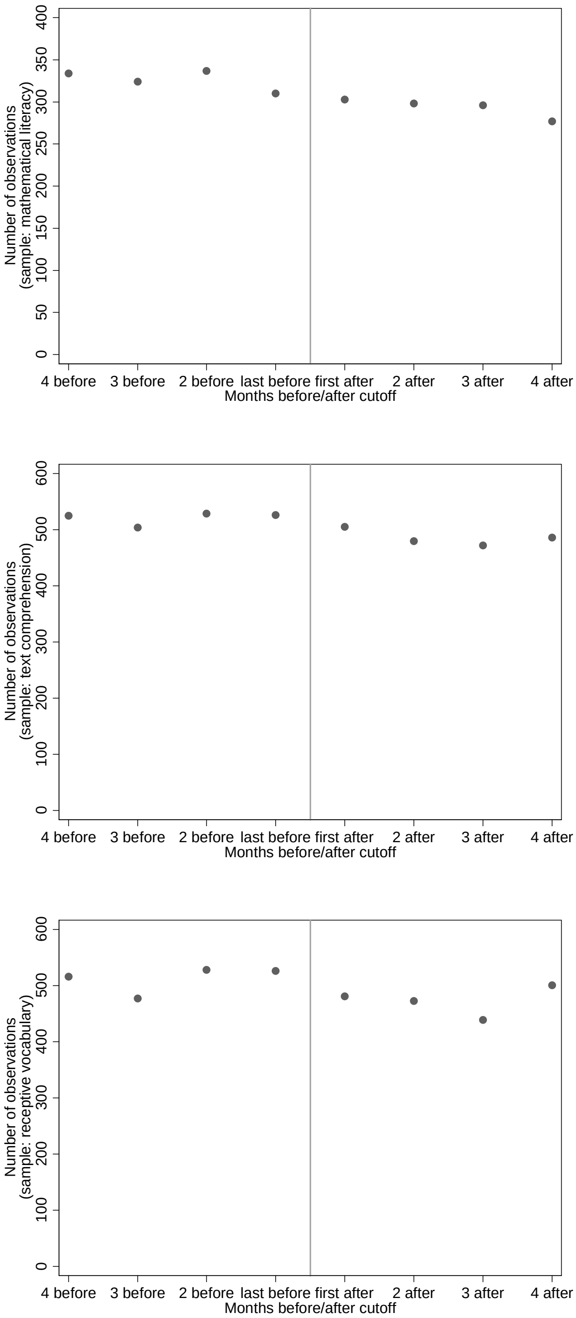


Figure A2. Distribution of the transformed receptive vocabulary test scores



Note: The transformed test scores are generated using a Box-Cox-transformation $y_i^{\text{transformed}} = (y_i^{\lambda} - 1)/\lambda$ *with Lambda equal to 4.22. The transformed test scores are then normalized to having a mean of zero and a variance of one.*

Figure A3. Number of observations by distance to cut-off



Note: Figures are for samples with information on mathematical literacy (top), text comprehension (center), and receptive vocabulary (bottom).

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business and Economics - Freie Universität Berlin

2019 erschienen:

2019/1 WANGENHEIM, Jonas von
English versus Vickrey Auctions with Loss Averse Bidders
Economics