Universität St.Gallen

Does the Early Bird Catch the Worm?

Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany

Patrick A. Puhani and Andrea M. Weber

01 2006 Discussion Paper no. 2006-02

Department of Economics University of St. Gallen

Editor:	Prof. Jörg Baumberger University of St. Gallen Department of Economics Bodanstr. 1 CH-9000 St. Gallen Phone +41 71 224 22 41 Fax +41 71 224 28 85 Email joerg.baumberger@unisg.ch
Publisher:	Department of Economics University of St. Gallen Bodanstrasse 8 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 22 98
Electronic Publication:	http://www.vwa.unisg.ch

Does the Early Bird Catch the Worm?

Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany

Patrick A. Puhani
Andrea M. Weber

Author's address: Prof. Dr. Patrick A. Puhani
TU Darmstadt and SIAW, Universität St. Gallen
Residenzschloss, Marktplatz 15
D-64283 Darmstadt, Deutschland
Tel. +49 6151 16-3693
Fax +49 6151 16-3897
Email puhani@vwl.tu-darmstadt.de
URL: www.tu-darmstadt.de/fb/fb1/vwl4/

Abstract

We estimate the effect of age of school entry on educational attainment using three different data sets for Germany, sampling pupils at the end of primary school, in the middle of secondary school and several years after secondary school. Results are obtained based on instrumental variable estimation exploiting the exogenous variation in month of birth. We find robust and significant positive effects on educational attainment for pupils who enter school at seven instead of six years of age: Test scores at the end of primary school increase by about 0.42 standard deviations and years of secondary schooling increase by almost half a year.

Keywords

Education, Immigration, Policy, Identification

JEL Classification

I21, I28, J24

Acknowledgement

This project has been initiated through discussions with Michael Fertig, RWI, Essen. We are also grateful to Gianno De Fraja, Peter Fredriksson, Edwin Leuven, Stephen Machin, Kjell Salvanes as well as seminar participants at the University of St. Gallen, the IZA Summer School 2005 in Buch am Ammersee and the CEPR-IFAU-Upssala Universitet Second Network Workshop 'Economics of Education and Education Policy in Europe' in Uppsala for helpful comments.
We thank Hans-Peter Hafner of the Research Data Center (*Forschungsdatenzentrum*) of the Statistical Office of the state of Hessen for help with the administrative data for Hessen.
'The Youth and Young Adult Longitudinal Survey used in this paper has been obtained from the Central Archive for Empirical Social Research (*Zentralarchiv für Empirische Sozialforschung*), University Cologne. The data of the "Youth and Young Adult Longitudinal Survey" has been collected by A. Fischer (Psydata Institute), the *"Jugendwerk der Deutschen Shell"*, R. K. Silbereisen (*Friedrich-Schiller-Universität Jena*), L. A. Vaskovics (*Otto-Friedrich-Universität Bamberg*) and J. Zinnecker (*Universität Siegen*) and prepared and documented by the *Zentralarchiv für Empirische Sozialforschung* (*ZA*). The mentioned persons and institutions are not responsible for the analysis and interpretation of the data in the present study.' All remaining errors are our own.

1 Introduction

The ideal age at which children should start school and the effectiveness of pre-school learning programs are subjects of ongoing debates among researchers and policy makers. For example, in the economic literature Currie (2001) summarises evidence on early childhood education. Age of school entry effects are estimated in Angrist and Krueger (1992) and Mayer and Knutson (1999) for the United States, Leuven *et al.* (2004) for the Netherlands, Strøm (2004) for Norway, Bedard and Dhuey (2005) for a set of industrialised countries, Fertig and Kluve (2005) for Germany and Fredriksson and Öckert (2005) for Sweden. In Germany, as in most other European countries, children are traditionally supposed to start school when they are about six years old. A look back in history reveals that starting education at the ages six or seven is not just a feature of the industrialised time. Already in Germany's mediaeval predecessor, the Holy Roman Empire, the track to knighthood began at age seven as a footboy (*Page*).[1] In post-war Germany, the changing attitude towards school entry age has been driven by debates among educationalists. In the beginning of the 1950s, Kern (1951) hypothesised that a higher school entry age could prevent children from failing in school. Subsequently, the school entry age was increased by a total of five months in 1955 and in 1964. Since that time, there has also been a trend to have children with learning problems enter school one year later than recommended by the official school entry rule. In recent years, however, debates on the long duration of the German education system have taken early school entry back on the agenda. Policy makers in Germany's decentralised education system have subsequently implemented measures to reduce the average age of school entry.[2] Therefore it seems reasonable to ask whether such policies can be expected to improve educational attainment.

[1] The period as a footboy was followed by the stages varlet (*Knappe*) at age 14 and knight (*Ritter*) at age 21. In contrast to modern times, it was not deemed important for a footboy to know how to read and write.

[2] In many schools, pupils may now enter school when they are five years old. For example, starting in the school year of 2005/2006, Berlin is changing the school entry regulation (*Schulgesetz*), so that the compulsory school entry age is half a year earlier than it used to be. At the same time the possibility to start school at a later than at the prescribed age has been abolished in Berlin.

In this paper, we estimate the causal effect of varying the age of school entry in Germany between six and seven years by an instrumental variable strategy using the exogenous variation of month of birth as an instrument for the age of school entry. The variation between ages six and seven is both a major variation observed internationally for the school starting age and a major issue of discussion in the national German debates. Using three different data sets, we measure the effect of age of school entry at three different stages: at the end of primary school, in the middle of the secondary school track and several years after the end of secondary schooling. Our outcome measures are a test score for primary school pupils and the school track attended or accomplished in the latter two data sets, respectively. To the best of our knowledge, ours is the second study investigating the effect of age of school entry by instrumental variable estimation for Germany. Because the previous study by Fertig and Kluve (2005) uses data sampled on the outcome variable (potentially leading to biases), we use two further data sets, one of which contains *all* pupils in a German state.

The influence of school entry age on educational outcomes is a well-discussed topic, especially in the U.S. and British empirical educationalist literature.[3] However, these studies do not sufficiently account for the endogeneity of the age of school entry: In Germany, as well as in many other countries, school entry age is not only determined by some exogenous rule, but depends on the child's intellectual or physical development or the parents' will, too. In

[3] Stipek (2002) provides a thorough review of this literature. One type of existing studies considers the effects of academic red-shirting (*i.e.* the delay of school entry) and early grade retention (*e.g.* May *et al.*, 1995; Jimerson *et al.*, 1997; Zill *et al.*, 1997; Graue and DiPerna, 2000) or of early school admission of selected children (*cf.* Proctor *et al.*, 1986, for a review). However, these studies do not appropriately take the endogeneity problem in measuring entry age effects into account and the mixed findings are therefore hard to interpret (*cf.* Stipek, 2002; Angrist, 2004). A second stream of literature examines the effect of entry age induced through season of birth on educational and social outcomes or mental development (*e.g.* Kinard and Reinherz, 1986; Morrison *et al.*, 1997; Hutchison and Sharp, 1999; Stipek and Byler, 2001). The results mostly indicate that there are no long-lasting effects while there is evidence of positive effects of a higher school entry age in the short run. Since outcomes are separately analysed by season of birth, which is taken as exogenous, the applied methods solve the endogeneity problem by producing reduced form estimates (without however explicitly discussing it). None of the mentioned studies uses an IV approach as in the recent economic literature.

several countries (*e.g.* the U.S.) some schools even use standardised tests in order to assess potential first graders' or kindergartners' school readiness.

A key institutional difference between Germany on the one hand and the U.S. or the U.K. on the other is that in Germany each child independently of date of birth has to complete at least nine years of compulsory full-time schooling[4]. In the U.S. and the U.K., length of mandatory schooling varies with date of birth, as children are allowed to leave school once they have reached a certain age (*cf.* Angrist and Krueger, 1992, for the U.S. and Del Bono and Galindo-Rueda, 2004, for the U.K.).[5] Hence, in these Anglo-Saxon countries compulsory schooling length is shorter for pupils having entered school at an older age. In Germany, however, all pupils at least have to wait until their ninth school year has finished before they may leave full-time education. Consequently, the German institutional setup allows identification of age of school entry effects independently of compulsory schooling, which is not possible in the U.S. or the U.K.

A further feature that makes the German case interesting to examine is that the German education system is highly selective. Unlike in most other countries, the child's performance in primary school is crucial for the educational career of a person because at the end of primary school (at age ten; primary school usually lasts for four years) children are selected into one of three educational tracks: the most academic is *Gynmasium*, usually consisting of nine further years of schooling, followed by *Realschule* (six years) and *Hauptschule* (five years and the most vocational track). As track selection is supposed to be based on the pupil's primary school performance, the German track system may aggravate age of school entry effects by perpetuating inequalities arising at early stages of the education

[4] The exact rule depends on the state. The nine or ten years of compulsory full-time education are followed by either at least one additional year of full-time education or by several years of part-time education in a vocational school (*Berufsschule*) within the German apprenticeship system.

[5] To be more precise, in England and Wales children could traditionally (between 1962-1997) leave school at the beginning of the Easter holiday in the school year in which they attained the relevant leaving age if they were born between September and the end of January. Children born between February and the end of August could not leave before the end of May.

system (*cf*. Hanushek and Wößmann, 2005). Hence, age of school entry may have larger and more lasting effects in Germany than in countries with a comprehensive school system.

The paper is structured as follows. Section 2 outlines age of school entry regulations for the cohorts we observe in our data and sketches main features of the German school system. The three data sets we use are described in Section 3. First, for primary school test scores we rely on the 'Progress in International Reading Literacy Study' of 2001 (PIRLS). Second, for the school track during secondary schooling we use newly available administrative data for the state of *Hessen* including *all* pupils in general education in the school year 2004/2005. Third, for schooling accomplishment several years after compulsory schooling we draw on data from the German Youth and Young Adult Longitudinal Survey of 1991. Section 4 argues that our empirical approach to identify the effect of age of school entry on educational outcomes is justified. We show that the instruments are effectively uncorrelated with the observed variables used as regressors and that first-stage regressions do not exhibit a weak instrument problem. The estimation results are presented and discussed in Section 5. We find robust evidence that increasing the age of school entry from six to seven years raises primary school test scores by more than two fifths of a standard deviation and increases the amount of secondary schooling by almost half a year (about five months). Only results based on the German Youth and Young Adult Longitudinal Survey are not robust. However, we place less weight on results from this survey as it is sampled on the basis of our outcome variable, which is likely to lead to biased estimates. Section 6 concludes and reports results from a small-size survey of headmasters and headmistresses, which we carried out in order to discuss potential explanations for our empirical estimates.

2 Age Of School Entry And The German Education System

In international comparison, the German compulsory school starting age of six years is equal to the median and mode of the distribution displayed in Table 1. Before the age of six, German children usually attend kindergarten, which is a playgroup rather than a pre-school. Projects where children learn how to read and write in kindergarten are recent and rare. Therefore, entering primary school for a German child traditionally has meant moving from a playgroup to an educational regime of teaching from eight o'clock in the morning to 12 o'clock in the afternoon with only short breaks (there is some variation on these times by state).

Although the exact school entry age is regulated by law in Germany, personal and school discretion is high. The school laws (*Schulgesetze*) of the states (*Länder*) are traditionally based on the so-called Hamburg Accord (*Hamburger Abkommen*) which was in place in Western Germany between 1964 and 1997. The Hamburg Accord states that children whose sixth birthday is before the end of June of a given calendar year enter school at the beginning of the corresponding school year (normally in August). Children born later are supposed to start school in the following calendar year (again around August). Thus, the *theoretical school entry age* $I^1\left(b_i, s_i\right)$ (as recommended by the Hamburg Accord) is related to a child's month of birth b_i and the month the school year starts s_i in the following way:

$$I^1(b_i, s_i) = \begin{cases} \dfrac{(72+s_i)-b_i}{12} & if\ 1 \le b_i \le 6 \\ \dfrac{(84+s_i)-b_i}{12} & if\ 6 < b_i \le 12 \end{cases} \qquad (1)$$

where the theoretical school entry age $I^1\left(b_i, s_i\right)$ is measured in years (in decimals up to the month). The indicator for the month of birth b_i ranges from one to twelve, whereas the variation in s_i is between seven and nine, as school always started in July, August or

September for the cohorts considered in our samples. If b_i and s_i are exogenous (*cf.* Sections 3 and 4), the theoretical school entry age $I^1(b_i, s_i)$ is exogenous and can be used as an instrument for the actual age of school entry. Note that the start of the school year s_i varies over calendar year and state as shown in Table A1 in the Appendix (whereas August 1st is the official nationwide school starting date, the actual starting dates vary by calendar year and state in order to avoid traffic jams on the motorways during vacation times).

Contrary to the Hamburg Accord, children born between the official cut-off date 'end of June' and the school year starting date s_i are often admitted to school in the calendar year when they turn six years of age. This practice provides an alternative instrument, which is exogenous under the same conditions as $I^1(b_i, s_i)$:

$$I^2(b_i, s_i) = \begin{cases} \dfrac{(72 + s_i) - b_i}{12} & \text{if } 1 \leq b_i \leq s_i \\ \dfrac{(84 + s_i) - b_i}{12} & \text{if } s_i < b_i \leq 12 \end{cases} \quad (2)$$

As – despite of the Hamburg Accord – the decisions on when to admit children to school are *de facto* taken at the school or parent level, we consider alternative instruments based on other cut-off dates in order to see which instrument works best in the first-stage regression. In addition to the end of June, we define instruments with the end of July until the end of September as cut-off dates c:[6]

$$I^3(b_i, c, s_i) = \begin{cases} \dfrac{(72 + s_i) - b_i}{12} & \text{if } 1 \leq b_i \leq c \\ \dfrac{(84 + s_i) - b_i}{12} & \text{if } c < b_i \leq 12 \end{cases} \quad (3)$$

Although the variation in the three instruments just introduced is mainly driven by variation in month of birth b_i rather than the school starting month s_i, the latter may be a

[6] We also tried later cut-off dates up to December, but these did not explain school entry age behaviour well.

source of endogeneity rendering the instruments invalid. One potential cause for the endogeneity of s_i is that s_i depends on the calendar year and state and both these factors may be correlated with unobserved characteristics relevant to educational outcomes. In addition to that, we observe pupils several years after the start of primary school, such that a pupil may have entered primary school in a different state than the state where he or she is observed in the data. In this case, we do not know the exact school starting date, which may lead to endogeneity of s_i by a correlation of the state we wrongly allocate a mover to and unobserved characteristics of that person.

To avoid these potential endogeneities, we build the following alternative instrument, which does not depend on s_i:

$$I^4(b_i, c) = \begin{cases} \dfrac{(72+8)-b_i}{12} & if \;\; 1 \leq b_i \leq c \\ \dfrac{(84+8)-b_i}{12} & if \;\; c < b_i \leq 12 \end{cases} \tag{4}$$

where we substitute '8' for the school starting date s_i, as August is the official start of the school year in Germany around which actual school starting dates vary.

The Hamburg Accord was made less binding in 1997, when the Council of the Ministers of Education encouraged the states to deviate from the traditional school entry cut-off date of end of June and to allow later cut-off dates (up to the end of September). This increased even further the discretion that schools and parents already had *de facto*. For example, in the state of *Hessen*, for which we use recent administrative data, the current official school entry age policy is to generally recommend application of the Hamburg Accord but additionally allow for early entry of children born several months later. In how far actual school starting ages comply with the regulations outlined here will be exhibited in Section 4, when we discuss the validity of the instruments introduced above.

Apart from the school entry regulations, tracking is another feature of the German education system important to the analyses in this paper. After four years in primary school, pupils usually change to one of three secondary school tracks.[7] The most vocational and least academic level of secondary schooling is called *Hauptschule* (grades five to nine), the intermediate level *Realschule* (grades five to ten) and the most academic level *Gymnasium* (grammar school, grades five to thirteen).[8] Track selection is important, as only graduation from *Gymnasium* directly qualifies for university or polytechnic tertiary education. *Hauptschule* and *Realschule* are supposed to be followed by vocational training within the German apprenticeship system. The distribution of pupils across the three tracks varies by state, but for Germany as a whole it is about equal. Although there are ways to enter the *Gymnasium* track after *Hauptschule*, *Realschule* or apprenticeship training, the track selection after primary school is a key decision for the economic and social life of a person in Germany (Dustmann, 2004). Note that Germany also has comprehensive schools (*Gesamtschulen*) as well as schools for children with special needs, mostly due to physical or mental disabilities (*Sonderschulen*). There are also so-called Waldorf schools that follow a special pedagogy which does not give marks to pupils, for example. In the year 2003, only 17 percent of graduates came from schools outside of the standard tracking system (eleven percent were in comprehensive schools, six percent in special schools and one percent in Waldorf schools), as Figure 1 shows.

[7] In the East German States of *Berlin* and *Brandenburg*, primary school goes up to grade six, so that the selection into school tracks starts two years later there than in the rest of Germany.

[8] In the East German states of *Sachsen* and *Thüringen, Gymnasium* ends after grade twelve. In the small West German state of *Rheinland-Pfalz*, *Gymnasium* nowadays ends after twelve and a half years of schooling. Most states are currently planning to have *Gymnasium* end after grade twelve, but this is not relevant for our samples.

3 Data

We use three different data sets measuring educational outcomes at three stages of pupils' or young adults' lives. First, the 'Progress in International Reading Literacy Study' (PIRLS) of 2001 provides us with internationally standardised test scores and other relevant information for 6,591 German pupils in the fourth grade of primary school. Second, we use administrative data on all pupils from the state of *Hessen* in the school year 2004/2005 who entered primary school between 1997 and 1999 and were attending secondary school at the time of observation. The observed cohorts overlap with those tested in the PIRLS study.[9] Our estimation sample thus contains 182,676 observations. Finally, the Youth and Young Adult Longitudinal Survey of 1991 provides data on secondary educational achievement for a sample of 1,199 persons aged between 22 and 29 at the date of interview. The surveyed cohorts are thus between about 20 and 30 years older than the persons sampled in the first two data sets. More detail is given in the following subsections.

3.1 The Progress In International Reading Literacy Study (PIRLS)

The PIRLS data has been collected by the International Association for the Evaluation of Educational Achievement (IEA) and includes test scores of an internationally conducted standardised reading literacy test as well as background information on pupils and parents. For Germany, 7,633 pupils at the end of fourth grade in 211 primary schools are sampled.[10] Because we lack information on the age of school entry (to the month) for more than one thousand observations, our effective sample size is reduced to 6,591.[11]

[9] We also tried to obtain administrative pupil statistics from other German states, but were either denied access or told that an essential variable for our analysis is missing.

[10] Because the sampling units are schools rather than pupils, all of our results presented in the following sections use standard errors adjusted for clustering. We also use the sampling weights provided in the data set.

[11] The age of school entry is unfortunately not missing at random: immigrants and pupils whose parents have a comparatively low level of education are overrepresented among the missing observations. If age of school entry is also missing systematically for pupils with unobserved characteristics that are relevant to educational attainment, our estimates based on the selected sample might be biased. However, as we control for parental

As we are interested in estimating the effect of age of school entry on educational attainment, we might like to sample a birth or school entry cohort and estimate the effect of interest after four years of schooling, no matter which grade pupils have achieved by then. The other possibility is to measure educational attainment at the end of primary school irrespective of how long it took the pupil to reach grade four. The advantage of the latter approach is that the pupil's performance at grade four of primary school is what matters in the end for the secondary school track recommendation he or she receives. As the PIRLS data samples pupils in grade four, we can only identify the parameter associated with the latter approach, except that it is not an entry cohort, but an exit cohort (fourth graders at the end of primary school) that is sampled. In our data, 86 percent of pupils have entered school in 1997, whereas eleven and two percent have entered in 1996 (grade repeaters) and 1998 (grade skippers), respectively. Hence, we observe pupils once they have reached grade four, even if they have spent only three or even five years in school. If grade repetition and skipping behaviour has not changed significantly between these neighbouring cohorts, our results should be roughly representative for the 1997 school entrants.

The instruments we can build with the PIRLS data are limited to the $I^4(b_i, c)$-type, because the data does not contain information on the state a pupil lives in (*cf.* Section 2). Therefore, the types of instruments using the school year starting month s_i, which depends on the state, cannot be constructed with the PIRLS data.

3.2 Administrative Data On All Pupils In The State Of *Hessen*

The second data source we use is 'Pupil-Level Data of the Statistics of General Schools for the State of *Hessen*' (*Hessische Schülereinzeldaten der Statistik an allgemein bildenden Schulen*). It covers *all* pupils in general education in the school year 2004/2005 and is

background and immigrant status, which is likely to be correlated with these characteristics, we hope to reduce this potential bias markedly.

collected on behalf of the state Ministry of Education. To our knowledge, this is the first research paper using this individual-level administrative data.

The original data set contains 694,523 observations from 1,869 schools. As it does not contain any school marks or test scores, we use the track attended in 2004/2005 by pupils having entered school between 1997 and 1999 as the outcome variable. This leaves us with 182,676 observations, 93 percent of them in grades six to eight. Tracks are coded according to the years of schooling they imply: 13 for *Gymnasium* (grammar school), 10 for *Realschule* and 9 for *Hauptschule*. Pupils at comprehensive schools (*Gesamtschule*) are frequently allocated to an internal track that corresponds to *Gymnasium*, *Realschule* or *Hauptschule*, as well. In this case, the administrative data codes them as if they were in these schools. If no such information is given, we code them as 10, *i.e.* equivalent to *Realschule*. Pupils in special schools (*Sonderschule*) are allocated code 7.[12]

Although the administrative data for the state of *Hessen* allows the construction of all four types of instruments introduced in Section 2, *i.e.* $I^1(b_i,s_i)$, $I^2(b_i,s_i)$, $I^3(b_i,c,s_i)$ and $I^4(b_i,c)$, we have a preference for $I^4(b_i,c)$-type instruments. The reason is that we do not know whether pupils entered schools in the state of *Hessen*, so that the exact school starting month s_i has to be proxied by assuming that pupils entered primary school in the same state where they attend secondary school. However, if we make that assumption, $I^1(b_i,s_i)$ and $I^4(b_i,c)$ as well as $I^3(b_i,c,s_i)$ and $I^4(b_i,c)$ will accidentally be identical, as August was the theoretical school starting date s_i in *Hessen* for the cohorts considered here. If both the cut-off

[12] 0.86 percent of pupils in the original sample are still in primary school when we observe them: they are excluded from the sample in the reported estimates since we do not know which track they will be assigned to. To check in how far these pupils affect our results, we carry out a rather extreme robustness check by allocating code 4 to individuals still in primary school, which indicates the fact that they failed to move to secondary school in time. We carry out a further sensitivity check by excluding pupils in comprehensive and special schools. Pupils in Waldorf schools are not separately identified: they are like comprehensive schools. Note that private schools are included in our sample: 10,709 pupils are in private schools, about 76 percent of whom attend grammar school (*Gymnasium*).

and the school-starting dates are set to August ($c = s_i = 8$), then $I^2(b_i,s_i)$, $I^3(b_i,c,s_i)$ and $I^4(b_i,c)$ are identical.

3.3 The Youth And Young Adult Longitudinal Survey

The final data set used in this study is the German Youth and Young Adult Longitudinal Survey *(Jugend- und Junge-Erwachsenen-Längsschnitt).* This data is an extension of the so-called Y*outh 92* (*Jugend 92*) survey conducted by the German Shell Company (*Deutsche Shell AG*). We use the 1991 cross section of this survey because it is – to our knowledge – the only German data set informing on school entry age and educational attainment later in life. However, we also have doubts about the adequacy of this survey for the analysis of age of school entry effects on educational outcomes. The problem is the stratified sampling by gender, region, town size, birth year *and* achieved secondary school track. Stratification on the outcome variable 'school track' is likely to lead to biases. We could not obtain more information on the stratification procedure and therefore have to carry out the analysis with this *caveat* in mind.

Since we want to consider individuals with completed secondary education, we restrict the sample to persons who are between 22 and 29 years old at the time of interview (29 is the maximum age in the survey; our sample thus covers birth cohorts 1961 to 1969). We only consider West Germans, as the school system in Eastern Germany was quite distinct from the one in the West. This leaves us with 1,199 observations. As the data set does not contain weights, we prefer regression specifications where we control for gender, region, town size and birth year, which are all variables influencing the sampling design.

The coding of educational attainment is similar to the one of the administrative data for the state of *Hessen*, *i.e.* 13 years of schooling for *Gymnasium* (grammar school), 10 for *Realschule* and 9 for *Hauptschule*. Persons without any of these degrees are coded as 7.[13]

As to the construction of the instrumental variables, we do not know whether a person went to primary school in a different state from the one where he or she was interviewed in 1991. Therefore, the instruments $I^1(b_i,s_i)$, $I^2(b_i,s_i)$ and $I^3(b_i,c,s_i)$ might be endogenous through a correlation of unobserved skills with state of residence in the survey year 1991, which would translate into a correlation of unobserved skills with the assumed theoretical school starting date s_i and hence the instruments. Note, however, that the variation in s_i is small (only three months) compared to the variation in month of birth b_i (twelve months), which is the main factor driving the variation in the instrument. Therefore, potential biases might be small. Nevertheless, we mainly rely on $I^4(b_i,c)$-type instruments in order to avoid these potential problems.

In the following section, we provide more detail on theoretical and actual age of school entry in our data and further discuss the validity of the instruments.

4 The Exogeneity Of Month Of Birth And First Stage Regressions

4.1 The Endogeneity Of Age Of School Entry

Regressing educational outcomes on age of school entry by ordinary least squares regression (OLS) must be expected to yield biased estimates rather than the causal effect of age of school entry on educational results. The reason is that the school entry decision is influenced not just by regulations like the Hamburg Accord, but also by the child's development as well as the parents' and the school's judgements. Thus, ambitious parents may want to push for an early

[13] As comprehensive schools (*Gesamtschulen*) were mostly introduced in the 1980s, the birth cohorts 1961 to 1969 did not attend them.

school entry (at age 5) of their child or children with learning problems might be recommended to enter school one year later (at age 7) than prescribed by official regulations. These mechanisms suggest that on average, less able pupils will enter school at a later age and thus OLS estimates of age of school entry effects on educational attainment should exhibit a downward bias.

Figure 2 displays the distributions of the actually observed school entry age and the theoretical entry age according to the 'Hamburg Accord' (for PIRLS, we exhibit the instrument $I^4(b_i, c = 6)$ with the end of June as cut-off date; for the administrative data for *Hessen* as well as the Youth and Young Adult Longitudinal Survey, we display the school entry age according to the Hamburg Accord $I^1(b_i, s_i)$, where we calculate s_i on the assumption that a pupil has not changed state. The prescribed school entry age varies between 6.17 and 7.08 years in the first two data sets (PILRS and administrative data for *Hessen*), whereas it varies between 6.17 and 7.22 years in the Youth and Young Adult Longitudinal Survey. The larger variation in the latter data set is explained by the variation of school starting months s_i across German states.

It is clearly visible that the actual distribution of age of school entry is far more dispersed and skewed to the right than the distribution prescribed by the Hamburg Accord (the skewness is positive and ranges from 0.33 to 0.50 in the three graphs). This is because many parents/schools have children start school one year later than suggested by the regulations. However, a few children also start school one year earlier at about age five. Despite of that, the large majority of pupils start school at the prescribed age.

A further graphical illustration of the degree of compliance with the age of school entry rules discussed in Section 2 is provided in Figure 3. The first panel displays the actual age of school entry by month of birth together with three different instruments. The instruments are a $I^4(b_i,c)$-type version of the 'Hamburg Accord' using the end of June as cut-off date without knowing the actual school entry month, $I^4(b_i,c=6)$, and two further versions of $I^4(b_i,c)$, one with the end of July and one with the end of August as cut-off date c. Visual inspection suggests a significant correlation between the instruments and the actual age of school entry. However, children born from October to June enter school a little older on average than prescribed by the Hamburg Accord. This is consistent with the graphs in Figure 2 showing that late entry is more frequent than early entry. However, for those born between July and September, the average age of school entry is lower than prescribed by the Hamburg Accord illustrating the fact that close to the cut-off point, many parents decide for their children to enter school early. This suggests using instruments with later cut-off dates.

A similar picture concerning non-compliance with the cut-off date of the Hamburg Accord arises in the last two panels of Figure 3. In the administrative data for *Hessen*, pupils born just after the cut-off date 'end of June' enter school earlier on average than demanded by the Hamburg Accord. Therefore we also consider instruments of type $I^3(b_i,c,s_i=8)=I^4(b_i,c)$ using the end of July and the end of August as cut-off dates, as in the PIRLS data.[14] Note that as school always started in August for the cohorts we analyse with the administrative data from *Hessen*, $I^2(b_i,s_i=8)=I^3(b_i,c=8,s_i=8)=I^4(b_i,c=8)$ if August is chosen as the cut-off date.

In the Youth and Young Adult Longitudinal Survey (third panel of Figure 3), the compliance with the Hamburg Accord, $I^1(b_i,s_i)$, seems weakest of all analysed data. The instruments using the start of the school year in the respective state, $I^2(b_i,s_i)$, as well as

[14] In the following analyses using the administrative data for *Hessen* we only report results based on instruments with June and July as cut-off dates. We also tried instruments based on later cut-off dates but there is no relevant compliance with these cut-off dates.

$I^4(b_i, c)$ with August or September ($c = 8; c = 9$) as cut-off dates appear to describe school entry behaviour better.

4.2 The Exogeneity Of The Instruments

In order for $I^1(b_i, s_i)$ to $I^4(b_i, c)$ to be valid instruments, they have to be both correlated with the actual age of school entry and uncorrelated with unobserved factors influencing educational performance in a prospective regression equation. In order to gauge whether the instruments are truly exogenous variables, *i.e.* uncorrelated with any *unobserved* factors that might influence educational performance, an assumption we cannot test directly, we test whether the instruments are correlated with *observed* variables that we believe might influence educational performance. In addition, we rely on 'discontinuity samples' where the included observations are limited to pupils born in the two adjacent months around the cut-off dates.

Table 2 lists the groups of regressors that we include in our two-stage least squares (2SLS) instrumental variable estimation models. Note that the regressors enter both in the first-stage (as discussed below in this section) and in the second-stage regressions (as discussed in Section 5). The set of variables is partly determined by data availability in the respective data sets. In the first set of regressions ('exo1') we include no regressors in the model except age of school entry as the variable to be instrumented. The justification for this procedure is that if the instrument (mainly driven by month of birth, *cf.* Section 2) is completely random and therefore exogenous, no other control variables are required in order to estimate the causal effect of age of school entry on educational attainment consistently in a 2SLS estimation procedure. Nevertheless, control variables that influence educational attainment may reduce the standard errors of the estimates.

As a first extension of the set of regressors (‘exo2’), we therefore include gender and regional indicators (if available). In the administrative data for *Hessen*, we also control for the school entry cohort among ‘exo2’. In the Youth and Young Adult Longitudinal Survey, year of birth is included among this set of regressors, as it is also one of the variables on which the sample is stratified. The third set of regressors (‘exo3’) adds cultural background to the set of regressors, measured either by an immigration or nationality indicator, as in PILRS and the administrative data for *Hessen*, or by religion, as in the Youth and Young Adult Longitudinal Survey, which does not contain information on country of origin. The fourth extension (‘exo4’) adds parental education, which is available in the PIRLS data and the Youth and Young Adult Longitudinal Survey, but not in the administrative data for *Hessen*. The fifth addition (‘exo5’) is only feasible for the Youth and Young Adult Longitudinal Survey and consists of variables referring to the first school day, *e.g.* an indicator whether the child received a gift from the parents, which is a tradition in Germany. Finally, we add family background variables like books at home or number of siblings to obtain the last set of regressors (‘exo6’). This is only possible for the PIRLS data and the Youth and Young Adult Longitudinal Survey, because the administrative data for *Hessen* does not contain this information. We consider the control variables added in ‘exo5’ and ‘exo6’ as potentially problematic, as they might be an outcome of pupils’ (potential) performance and hence be endogenous: For example, parents might be more likely to give presents to children or to buy them books if they are not (expected to be) performing well in school. Hence, controlling for these sets of variables may take out some of the effect that age of school entry has on educational attainment.

Although low correlations between the instrument and observable variables are supportive of the instrument’s exogeneity, they do not provide a guarantee. Therefore, we additionally consider so-called ‘discontinuity samples’, where only students born in two adjacent months around the respective school entry cut-off points are included. The reason for examining these discontinuity samples are potential direct effects of month or season of birth on health and educational outcomes, as briefly surveyed in Bound, Jaeger and Baker (1995). By restricting our samples to persons with two adjacent months of birth, we hope to eliminate any potential seasonal effects which might affect the validity of the instruments. Furthermore, any differences in parental attitudes potentially reflected in planned timing of births should be minimised for children born in two adjacent months, as it is hard to assure for a child to be born in a very specific month.

In Table 3 to Table 5 we display the simple correlations between a selected set of instruments and the full set of our control variables for all three data sets. Correlations significant at the 10 or 5 percent level are marked with one or two asterisks, respectively. As Table 3 shows, the maximum correlation for the full-sample of the PIRLS data equals 0.02 in absolute value, which is very small. Hence, the few correlations of instruments with regressors that are significantly different from zero are very close to zero. This finding is even more striking in the full sample of the large administrative data set for *Hessen* in Table 4: No correlation is larger than 0.01 in absolute value. Our instruments (mainly driven by month of birth) thus seem unrelated to gender, the district of residence and the country of origin. Table 3 also shows that our instruments are virtually unrelated to parental education, the number of siblings and the number of books in the household. In the discontinuity samples, there are two statistically significant correlations of 0.05 in the PIRLS data (*cf.* Table 3). In the administrative data for *Hessen* (*cf.* Table 4), which is larger in sample size, the maximum correlation remains 0.01 in the discontinuity samples. We interpret the findings of no or extremely small correlations of the instruments with the observed characteristics as indications that the instruments are plausibly exogenous and thus not correlated with unobserved variables either.

A *caveat* applies in this respect when considering the findings based on the Youth and Young Adult Longitudinal Survey in Table 5. Although by far the largest part of the correlations of the instruments with the regressors is insignificant and small in absolute size, some correlations are as large as 0.06 to 0.08 in the full sample and up to 0.14 in the (small) discontinuity samples in absolute value, respectively. This may be explained by the stratified sampling procedure. As Table 5 demonstrates, parental education is one of the variables slightly correlated with the instruments (the correlation is 0.05 in absolute value in the full sample). This is potentially reflecting the fact that sampling is based on the outcome variable 'school track'. Because educational levels of parents and children are known to be correlated (Dustmann, 2004), sampling on the outcome variable must be expected to generate biases. Therefore, we believe that instrumental variable estimates based on the Youth and Young Adult Longitudinal Survey should be regarded with caution. Nevertheless, we will still report results based on this survey, also to compare our findings with those of Fertig and Kluve (2005), who use this data.

4.3 First-Stage Regressions

Having discussed the exogeneity of our instruments, we now check the second condition for a valid instrument, namely the (partial) correlation with the variable to be instrumented (age of school entry). Table 6 to Table 8 report coefficients of the instruments together with the *F*-statistics of the tests for significance of the instruments in the first-stage regressions of the 2SLS estimation procedure. A rule of thumb states that an *F*-statistic below about 10 is indicative of a weak instrument problem (Staiger and Stock, 1997; Stock, Wright and Yogo, 2002).[15] The tables therefore display the *F*-statistics for various specifications distinguished by

[15] If instruments are weak, the 2SLS estimator has a high standard error and inference using asymptotic approximations for the standard errors is not reliable. Furthermore, already a very small correlation between the instrument and the error term of the outcome equation may lead to significant inconsistencies if instruments are weak (Bound, Jaeger, and Baker 1995). In other words, 2SLS with weak instruments is generally not appropriate.

both the choice of instrument and the choice of regressors (‘exo1’ to ‘exo6’) as outlined in Section 4.2.

Table 6 to Table 8 clearly show that, in all three data sets, we have instruments with *F*-statistics largely above the threshold value of 10. The degree of compliance with the rules built into the various instruments can be seen from the coefficients reported in the tables. Note that using the full samples, the degree of compliance is influenced by the behaviour of both individuals born around the cut-off date and persons born in months like January, which are quite distant from the alleged cut-off dates. In the PIRLS data (Table 6), the coefficients of the full sample vary between 0.31 and 0.49, which means that increasing the prescribed age of school entry by one year raises the actual age of school entry by between a third and a half of a year on average. The ‘Hamburg Accord’ shows the highest compliance in the PIRLS data.

Using the discontinuity samples of persons born in the two months adjacent to the respective cut-off date also reveals that the compliance with the Hamburg Accord is strongest with a coefficient of 0.40. The cut-off date July renders a weak instrument (with an *F*-statistic of around 5). The compliance is stronger again when August is used as cut-off, with a first-stage coefficient of 0.27 and an *F*-statistic of around 70. Although this makes both the Hamburg Accord (June as cut-off date) and August as cut-off date the relevant instruments (later cut-offs do not render useful instruments), one has to keep in mind that 2SLS estimation identifies the causal effect of age of school entry for ‘compliers’, *i.e.* those persons who react to variations in the instrument (Imbens and Angrist, 1994). Although the 2SLS model implicitly assumes that the effect of age of school entry is homogeneous across the population, the estimate is an equivalent of the local average treatment effect (LATE) as introduced in Imbens and Angrist (1994) for binary instruments. Persons reacting to June (the Hamburg Accord) as cut-off might consequently be more representative for the average pupil, unlike those reacting to August as the cut-off. It is plausible that the group of pupils born in

August and entering school at the age of just about six (younger than prescribed by the Hamburg Accord) are above-average achievers and hence distinct from the representative pupil. If virtually all 'compliers' born in August and September are high achievers, it may be that the 'compliers' for the instrument $I^4(b_i,c)$ with August as the cut-off date are affected differently by the variation in the age of school entry than compliers with the official rule of the Hamburg Accord.[16]

In the administrative data for *Hessen* the degree of compliance is also half a year for the Hamburg Accord in the full sample, but a third of a year for July as the cut-off date (*cf.* Table 7). An investigation of the discontinuity samples around the cut-off points reveals that compliance is only sufficiently strong with the Hamburg Accord (June as cut-off) with 0.41 as the first-stage coefficient. The first-stage coefficient for July as the cut-off date is very low at 0.04 (so is the coefficient for August as the cut-off date, which is not shown here). The full-sample first-stage results with the July cut-off date thus seem to stem from a broader concept of compliance than the specific cut-off date 'end of July'. This is to say they are explained by the fact that months of birth in spring (*e.g.* May) lead to lower ages of school entry than months of birth in autumn (*e.g.* October), which is true no matter whether June or July are chosen as hypothetical cut-off date. Hence, what we observe in the full-sample with July as the cut-off date is in fact a reflection of the Hamburg Accord (the June cut-off). Therefore, in the administrative data for *Hessen* just as in the PIRLS data, the Hamburg Accord is the appropriate instrument.

In the Youth and Young Adult Longitudinal Survey, the degree of compliance is highest if the school starting months or September are used as cut-off dates (the coefficients vary between 0.74 and 0.79 in these cases, meaning that an increase in the prescribed school entry age by one year increases the average age of school entry by about three quarters of a

[16] We do not know who is a complier: the counterfactual had a pupil been born in a different month is not

year). The coefficient in the full sample reduces to 0.58 or 0.59 if August is used as the cut-off month. However, the discontinuity samples reveal compliance only for September as the cut-off date. The displayed coefficients vary between 0.66 and 0.72. First-stage coefficients of other cut-off dates are not significant and not shown here.

In sum, the estimated first-stage coefficients and their *F*-statistics confirm the picture given in Figure 2 that compliance with the school entry rules is considerable, but not perfect. As mentioned above in this section, 2SLS estimates the effect of age of school entry on educational attainment only for the group of persons complying with the regulations. Therefore, we have to keep in mind that the results discussed in the following section may not be representative for the pupil population as a whole (*cf.* Imbens and Angrist, 1994, on local average treatment effects). Non-compliers are likely to be particularly weak pupils who enter school later than prescribed or strong performers who enter school earlier than suggested by the rules, or children of parents who have strong views on the age at which their child should enter school and consequently would not respond to cut-off dates.

Having justified our instruments in terms of exogeneity and (partial) correlation with the age of school entry, we present the results of the second stage of the 2SLS estimates in the following section.

5 The Effect Of Age Of School Entry On Educational Outcomes

5.1 Ordinary Least Squares Results

Table 9 to Table 11 report the estimated effects of age of school entry on educational attainment from regressions with different sets of control variables ('exo1' in the first line indicating no control variables, and the last line indicating the full set of control variables as listed in Table 2 to Table 5).

observable. Thus we cannot test whether the compliers born in August or September are high achievers.

The columns headed '(0)' of Table 9 to Table 11 report OLS regression coefficients for the full samples. In all data sets, the regression coefficient is negative and significantly different from zero if no control variables are included (specifications 'exo1'). This means that educational attainment and age of school entry are negatively correlated: Pupils who enter school at a later age achieve less than their peers entering at a younger age. However, as we include more and more control variables into the regressions, the OLS-coefficients decrease in absolute value in all data sets indicating that actual age of school entry is influenced by factors relevant to educational performance. This is highly suggestive of age of school entry being an endogenous variable, which warrants instrumental variable estimation.

5.2 Two-Stage Least Squares Results

What happens to the estimated effect of age of school entry on educational attainment if we apply 2SLS estimation with the instruments tested in Section 4? A glance at Table 9 to Table 11 reveals first that instrumental variable estimation switches the sign of the estimated effect from negative to positive in all data sets in all specifications with statistically significant coefficients. Second, in the PILRS data and in the administrative data for *Hessen*, the 2SLS estimates with the Hamburg Accord, the strongest instrument, are all positive and significantly different from zero. Third, the differences of the point estimates between the full and the discontinuity samples are smaller than a discontinuity-sample standard deviation in both the PIRLS and administrative data for *Hessen* if the Hamburg Accord (June as the cut-off date) is chosen as instrument (the specification with control variables 'exo2' in the data for *Hessen* is the only exception where this difference is slightly larger). Fourth, although the size of the estimated effects varies by the choice of instrument (as can be expected from the first-stage results discussed in Section 4.3), they hardly vary by the choice of control variables (*i.e.* between specifications 'exo1' to 'exo6'/'exo3') in Table 9 and Table 10: Indeed, the variation of the 2SLS estimates within a column is virtually always less than any estimated standard error of a coefficient in that column. In the Youth and Young Adult Longitudinal Survey (Table 11) this is only true in column (4a), where $I^4(b_i,c)$ is used as instrument with September as cut-off date, the appropriate cut-off date for this data.

In the following, we will discuss the 2SLS results in detail by data set. The PIRLS data do not contain information on the state a pupil lives in, so that we can only use instruments of type $I^4(b_i,c)$ for this data. Column (1a) in Table 9 sets $c = 6$ in the full sample, which is the same cut-off date as in the Hamburg Accord. The results for the corresponding discontinuity sample are shown in column (1b). Columns (2a), (3a) and (2b), (3b) shift the cut-off date to July and August for the full and the discontinuity samples, respectively. As reasoned in Section 4.2, the inclusion of more control variables in the 2SLS regressions mostly reduces the standard error of the estimated coefficient on age of school entry (as we move from 'exo1' to 'exo6').

The main finding in Table 9 is that the estimated effect of age at school entry on educational outcomes varies from 25.8 to 30.7 test scores in columns (1a) and (1b) when June is used as the cut-off date. The ranges of the estimated effects overlap between the full sample in column (1a), 26.8 to 30.7 test scores, and the discontinuity sample in column (1b), 25.8 to 29.0 test scores. The estimates for the strongest instrument are therefore robust across the full and the discontinuity samples.

In the discontinuity samples, we find no significant effect when July or August are used as cut-off dates. Given the weak instrument for July as cut-off, this is not surprising. However, the instrument with August as the cut-off is not weak, which has been shown in Table 6. As argued in Section 4.3, the 'compliers' reacting to June and August as cut-off dates may be rather different groups, with the latter plausibly comprising more talented pupils (the latter group consists of pupils who enter school early even if they are born two months after the official cut-off date). Because 2SLS estimates a local average treatment effect, the effect for the group of compliers with respect to the August cut-off may be different from the effect for the group of compliers with respect to the Hamburg Accord (June cut-off). Given the results based on the discontinuity samples, the statistically significant full-sample results in columns (2a) and (3a) in Table 9 are likely to stem from compliance of individuals born in months distant from the respective cut-off dates and hence seem to be driven by the cut-off date June, *i.e.* the Hamburg Accord (*cf.* our discussion in Section 4.3). We thus regard the results based on the Hamburg Accord as our most relevant estimates. Nevertheless, the findings on the compliers reacting to the August cut-off indicate that age of school entry effects are heterogeneous across the pupil population.

How can the results be interpreted? A representative estimate from the Hamburg Accord as instrument is an increase in test scores of around 27 points for raising the school entry age by one year (from about six to seven years of age). This is a bit more than two fifths of the standard deviation of test scores in PIRLS (the standard deviation is 63.61, so that the estimated effect amounts to 0.42 standard deviations). More intuition for the size of this effect is derived from a comparison of the differences in test scores between the different German school tracks in the PISA 2000 study (where ninth graders' reading literacy is tested).[17] In the PISA data for ninth graders, the differences in test scores are 0.78 standard deviations between pupils in *Gymnasium* and *Realschule* and 1.01 standard deviations between *Realschule* and *Hauptschule* (Baumert *et al.*, 2003). Therefore, our estimates imply that increasing the age of school entry from six to seven years increases reading literacy by more than half of the difference between the average *Gymnasium* track and the average *Realschule* track performance. This is quite a substantial effect and indicates that age of school entry may influence track choice, as also shown in the following paragraphs.

Table 10 presents the effects of age of school entry on track attendance in the middle of secondary school (measured by the number of school years associated with each track as outlined in Section 3.2). Results are based on administrative data for the state of *Hessen.* The 2SLS estimates with the Hamburg Accord $I^1(b_i,s_i)$ as instrument are given in columns (1a) and (1b) for the full and the discontinuity samples, respectively. The results for $I^4(b_i,c)$-type instruments with July as the cut-off dates are shown in columns (2a) and (2b), respectively. As discussed in Section 4.3, there is hardly any compliance with the July cut-off in this data, so that columns (2a) and (2b) are only displayed for illustrative purposes. We do not consider them indicative of the causal effect of school entry age on track attendance, since they are based on weak instruments.

Because the administrative data for *Hessen* is large in terms of number of observations (in fact we observe the population), the reported 'standard errors' in Table 10 all indicate significance (only the standard error in column (2b) is sizeable because compliance with the corresponding instrument is very low, *cf.* Table 7). As to the estimated effect of age of school entry on educational attainment using the Hamburg Accord as instrument, columns (1a) and (1b) yield comparable estimates in the ranges of 0.41 to 0.45 and 0.37 to 0.40 for the full and discontinuity samples, respectively, with minor variation among specifications with different sets of control variables.[18] The Hamburg Accord as the appropriate instrument thus exhibits robust positive effects of age of school entry on track attainment across full and discontinuity samples as well as across specifications with different sets of control variables: Entering school at the age of seven rather than six raises secondary schooling by almost half a year (around five months).[19] This effect is implied if a deferral of school entry by one year

[17] We do not use the PISA data for our estimations, because it does not contain the required information.

[18] The reported coefficients would be similar but somewhat higher if we did not exclude persons still in primary school from the sample. If we include primary school pupils (with code 4 as the outcome, *cf.* footnote 12), the coefficients related to columns (1a) and (1b) range between 0.46 to 0.49 and 0.43 to 0.46 in the full and discontinuity samples, respectively. Hence, early school entry seems to increase the likelihood of repeating grades in primary school. As a further robustness check we exclude pupils in comprehensive and special schools (*Gesamtschule* and *Sonderschule*). In this case the effects are only slightly different from the presented effects and range between 0.42 and 0.47 (1a) and 0.36 to 0.39 (1b) in the full and discontinuity samples, respectively.

[19] This interpretation implies the assumption that pupils will complete the track which they attend in the middle of secondary school, when we observe them.

increases the probability of attending *Gymnasium* instead of *Realschule* by about 13 percentage points.[20] Thus the results from the largest data set used in this study are qualitatively consistent with the findings from primary school reading literacy test scores.

Table 11 presents the estimated coefficients based on the third and smallest data set, the Youth and Young Adult Longitudinal Survey. As we have outlined above, the Youth and Young Adult Longitudinal Survey is sampled on the outcome variable (educational achievement), which may generate biases. This fact or simply the lower sample size in this data, and hence the relatively large standard errors, may explain the comparatively wide variation of the estimated effects of age of school entry on educational attainment across different specifications concerning the set of control variables.[21] Due to these data deficiencies, the results from this data set can only be taken with a grain of salt.

In contrast to the findings from the first two data sets, expanding the set of regressors in the Youth and Young Adult Longitudinal Survey changes (mostly reduces) the coefficient of age of school entry on educational attainment and all effects are insignificant in the specifications with the full set of regressors ('exo6'). It is interesting, though, that the 2SLS point estimates based on the full sample without control variables ('exo1') are in a similar range, *viz.* from 0.37 to 0.77, to the results based on the administrative data for *Hessen*, albeit with much larger standard errors. Column (1a) displays the results with the Hamburg Accord as the instrument, $I^1(b_i,s_i)$, column (2a) with the school starting dates, $I^2(b_i,s_i)$, and columns (3a) and (4a) show the estimates based on instrument $I^4(b_i,c)$ with August and September as

[20] The estimated effect is potentially driven by both increases in the probability to attend *Realschule* rather than *Hauptschule* and increases in the probability to attend *Gymnasium* rather than *Realschule*. In order to find out which of these effects drives the results, we estimated linear probability models of *Gymnasium* versus *Realschule/Hauptschule* attendance as well as of *Gymnasium/Realschule* versus *Hauptschule* attendance. Estimates were obtained by two-stage least squares (2SLS) using the same instruments and control variables as in Table 10. The results show robust increases of *Gymnasium* versus *Realschule/Hauptschule* attendance by between 11 and 13 percentage points and increases of *Gymnasium/Realschule* versus *Hauptschule* attendance of about 2 to 3 percentage points. The numbers are very robust and significant across different specifications (using the Hamburg Accord as instrument). Hence, it seems that the age of school entry matters for achieving *Gymnasium* attendance, which is the step towards university education and high labour market returns.

[21] For the discontinuity sample, the point estimates are negative, but insignificant due to large standard errors.

cut-off dates, respectively. Estimation results for the discontinuity sample with September as the cut-off date are shown in column (4b). As we do not know whether a person has changed the state of residence since entering primary school (an event 15 to 23 years before the survey date), the instruments $I^1(b_i,s_i)$ and $I^2(b_i,s_i)$ might be affected by measurement error as discussed in Section 3.3. Thus focusing on the $I^4(b_i,c)$-type instruments, we observe a remarkable fall in the point estimates in columns (3a), (4a) and also (4b) when including parental education among the set of regressors, *i.e.* when moving from 'exo3' to 'exo4'. Considering the findings with 'exo4' as the control variables (*i.e.* without first school day and family background variables, which might be endogenous as argued in Section 4.2), we find an estimated effect of 0.47 years in column (3a), which is significant at the 10 percent level. This estimate is in line with the results from the administrative data from *Hessen* and provides very tentative evidence that age of school entry might have long-run effects on educational achievement.

5.3 Results For Subgroups

Having established robust evidence from PIRLS and the administrative data from *Hessen* that an older age of school entry raises educational attainment, we carry out a subgroup analysis in Table 12 and Table 13 for these two data sets. Due to the smaller sample size and the reservations we have concerning the quality of the Youth and Young Adult Longitudinal Survey, we do not consider this data set for a subgroup analysis.

Table 12 displays first-stage coefficients and *F*-Statistics as well as second-stage estimation results for native males, native females, immigrant males, immigrant females and for pupils with parents with and without an academic degree, respectively. The estimates are exhibited both for the full and the discontinuity samples with the Hamburg Accord as the instrument and refer to the specification with all control variables ('exo6').

The main results from the subgroup analysis based on the PIRLS data are that German males benefit more than German females from later school entry: Coefficients are 42.9 (standard error 8.6) versus 16.2 (standard error 8.4) in the full samples, respectively. Due to smaller sample sizes and large standard errors (the latter ranging from 5.9 to 62.1 test scores), the subgroup estimates, especially in the discontinuity samples, are generally harder to pin down. Potentially for the same reasons, some estimated effects for male immigrants (full sample), female immigrants (full and discontinuity sample), for female natives (discontinuity sample) and for pupils with parents holding an academic degree (discontinuity sample) are not significantly different from zero.

Note that only the effects for the group of persons who comply with the instruments in the respective subgroup are identified by 2SLS. Therefore, the estimated 'local average treatment effects' do not have to be representative for the subgroups in general (for example, if most immigrant males enter school at the age of seven anyway, the compliers will be a small and unrepresentative group). However, first-stage coefficients show that the degree of compliance is similar for most subgroups, especially in the full sample. First-stage coefficients in the full sample mainly range between 0.44 and 0.56. Exceptions are immigrant females and pupils whose parents have attained an academic degree, for whom compliance is somewhat lower (the full-sample first-stage coefficients for these two groups are 0.38 and 0.35, respectively).

As in Table 12 for the PIRLS data, the estimates in Table 13 are shown both for the full and for the discontinuity samples with the Hamburg Accord as instrument and refer to the specification with all control variables ('exo3' in this case). The subgroup results for the administrative data for the state of *Hessen* do not confirm that German males benefit more from later school entry than German females. However, the different results from these two data sets need not contradict as PIRLS measures only reading literacy, whereas the secondary school track in the data for *Hessen* is a more general indicator for educational attainment.

In the administrative data for *Hessen*, we can distinguish between different groups of nationalities (German, Turkish, predominantly Muslim countries without Turkey, Italy/Grecce and former Yugoslavia). As sample sizes for all subgroups except Germans and Turks are below 1,600 (full samples) or 300 (discontinuity samples), the standard errors of the second-stage estimates range between 0.26 and 0.76, so that second-stage coefficients for these nationality groups are hard to pin down. We therefore ignored other nationality groups with even smaller sample sizes.

The first-stage coefficients for almost all subgroups are close to those of the sample as a whole, exceptions being both males and females from predominantly Muslim countries without Turkey, where compliance is lower (full-sample first-stage coefficients range between 0.35 and 0.36 for these groups compared to between 0.45 and 0.52 for the rest). Although there is some indication based on the first-stage F-statistics that the instruments for these two groups are not that strong, the marginally (in)significant point estimates for females from predominantly Muslim countries without Turkey tentatively suggest that they benefit more than natives from a later age of school entry. However, the large standard errors associated with these estimates make this interpretation somewhat speculative as the difference in the estimated effects is not statistically significant. The somewhat smaller point estimates for Turkish than native pupils are also associated with a sizeable standard error making this difference statistically insignificant. We cannot detect any significant effects of age of school entry for male or female pupils from Italy and Greece or for males from former Yugoslavia. However, at least in the full sample, the estimated effect for females from former Yugoslavia is significant and the largest of all groups (0.95), albeit with a sizeable standard error (0.41).

6 Conclusions

Based on instrumental variable estimation, we recover positive and statistically significant effects on educational outcomes for entering school at a relatively higher age in the current German school system. In the fourth grade of primary school, we find a large effect of about 0.42 standard deviations improvement in the PIRLS test score if the pupil enters at about the age of seven rather than six (*i.e.* a year later according to the school entry rule). This amounts to more than half of the difference in the average *Gymnasium* versus *Realschule* test scores in the OECD PISA study. Administrative data for the state of *Hessen* suggest that the effect of age of school entry persists into secondary school by prolonging average years of schooling by almost half a year (about five months). As the statutory length of *Gymnasium* is mostly 13 years versus ten years in *Realschule*, our estimates of the age of school entry effect are smaller in secondary than in primary school. Yet they remain sizeable.

Compared to Fredriksson and Öckert (2005) and Bedard and Dhuey (2005), who apply an instrumental variable strategy similar to ours to Swedish administrative and the international TIMSS data, respectively, the results for Germany are comparable in size: Fredriksson and Öckert (2005) report that entering school a year later increases ninth graders' grade point average by about 0.2 standard deviations. Similarly, the effects reported in Bedard and Dhuey (2005) range from 0.1 to 0.35 standard deviations for fourth and eighth graders in the countries investigated, most of which are in Europe. Strøm (2004) estimates an effect of 0.17 standard deviations for 15-16 year olds in the Norwegian PISA study, arguing that age of school entry is exogenously driven by regulations in Norway.[22] However, these and our

[22] Our estimates based on the PIRLS data (0.42 standard deviations) are on the high end of the range of results from other countries. However, in relation to the first-stage coefficients reported for eleven countries in Table 5 of Bedard and Dhuey (2005) as well as those in Fredriksson and Öckert (2005) for Sweden, the degree of compliance with the instrument in Germany is at the very low end in international comparison. As we can only estimate a local average treatment effect, the compliers in Germany might be less representative of the average pupil in Germany than in Sweden, for example, were compliance is higher. This might be one reason – apart from differences in school systems, data collection and other factors – why point estimates differ across countries. Indeed, correlating first- and second-stage coefficients for the eleven countries analysed in Table 5 of Bedard and Dhuey (2005) provides a correlation of -0.20 for science and -0.01 for maths test scores

estimates differ from those of Angrist and Krueger (1992) and Mayer and Knutson (1999) for the United States, where either no or negative effects for late school entry are reported. The findings for the U.S. can only be partly explained by the fact that quarter of birth in the U.S., unlike in Germany, affects the duration of compulsory schooling: No and negative effects of later school entry are found for persons having obtained post-compulsory schooling in Angrist and Krueger (1992) and Mayer and Knutson (1999), respectively.

Like Fertig and Kluve (2005), we can hardly detect effects of age of school entry on the educational achievement of young adults based on the German Youth and Young Adult Longitudinal Survey. However, we believe that the sampling procedure in this survey may lead to biased estimates and therefore we place little weight on estimates from this data set.

Given the current trend in Germany to have pupils start school earlier, we interviewed 25 primary school headmasters or headmistresses in the state of *Hessen* by telephone. We asked them about their views on our finding that late school entry improves educational performance.[23] Of the 25 schools, two were operating under a special regime where pupils enter school at the age of five, but with extra logopedic, German language and nursery teacher support. In these schools, five-year olds do not enter grade one, but 'grade zero', which is a mixture between a kindergarten and a school regime. Both schools are satisfied with this regime, as they are able to correct deficits some children have through the extra teaching and nursery resources they have (one of these schools stated that they have a 75 percent immigrant share). In a third school, we were not able to communicate the substance of our question. However, in the remaining 22 'standard' primary schools, 95 percent of headmasters or headmistresses (21 out of 22) said they found our results 'plausible'. We then went on to ask

in the TIMSS study. Hence, at least for maths, estimates based on a larger degree of compliance seem to be associated with a lower average treatment effect. The fact that the compliers with the August cut-off in the German PISA data exhibit no significant effect of age of school entry on test scores is consistent with these arguments. We thank Peter Fredriksson for pointing this issue out to us.

[23] We drew 30 telephone numbers of primary schools from the school registry of *Hessen* until we managed to talk to 25 of them (three schools refused to be interviewed by telephone and in two of the schools we could not reach a contact person after several trials).

them what they believed could be the reasons for these findings. All 95 percent (21 schools) made statements along the lines that older pupils are more mature, are more able to concentrate when having to keep still in the classroom for long periods of time, are more able to organise themselves (like keeping their belongings together), are less distracted by play and find it easier to overcome frustration. Only 18 percent of schools (four out of 22) felt that relative age effects matter, too. The other schools, however, explicitly denied the importance of relative age effects and stressed that it is personal maturity that matters.[24] Most 'standard' primary schools were opposed to early school entry in the current 'standard' educational regime, but supported the idea of early school entry if the school system changed to a situation similar to the special regime schools, which have extra support for pupils with learning, language or social problems and a 'grade zero' which combines learning with kindergarten elements.

Interpreting the results of our statistical analysis and those of the school survey, we thus find consistent evidence against early school entry into the current German school regime. However, our results should not be interpreted as evidence against early learning. Early learning, organised differently from the standard German school system, might be promising. Which type of early learning works best is an interesting research agenda for the future, once state governments decide to collect and make available appropriate data in this respect. The research presented in this paper suggests that simply reducing the age of school entry in Germany without adapting the style of teaching in early grades is not an optimal strategy. Although it is true that negative effects of early school entry have to be weighed against the economic gains of entering the labour market earlier, new data on earnings, age of school entry and educational outcomes for a representative sample would be needed to carry out an appropriate cost-benefit analysis on this issue.

[24] This is consistent with the findings of Fredriksson and Öckert (2005) that absolute age effects dominate relative age effects in Sweden. Since our pupil data do not allow us to distinguish between absolute and relative age effects, we can only present our telephone survey results on this matter.

References

Angrist, J.D. (2004): American Education Research Changes Tack, *Oxford Review of Economic Policy* 20, 198-212.

Angrist, J.D. and A.B. Krueger (1992): The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples, *Journal of the American Statistical Association* 87: 328-335.

Baumert, J., U. Trautwein and C. Artelt (2003): Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens, in Deutsches PISA-Konsortium (eds): *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*, Verlag Leske + Budrich, Opladen.

Bedard, K. and E. Dhuey (2005): *The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects*, Working Paper: Department of Economics University of California, Santa Barbara.

Bertram, T. and C. Pascal (2002): *Early Years Education: An International Perspective*, Qualifications and Curriculum Authority, London.

Bound, J., D.A. Jaeger and R.M. Baker (1995): Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variables is Weak, *Journal of the American Statistical Association* 90: 443-450.

Currie, J. (2001): Early Childhood Education Programs, *Journal of Economic Perspectives* 15: 213-238.

Del Bono, E. and F. Galindo-Rueda (2004): *Do a Few Months of Compulsory Schooling Matter? The Education and Labour Market Impact of School Leaving Rules*, IZA Discussion Paper No. 1233.

Dustmann, C. (2004): Parental Background, Secondary School Track Choice, and Wages, *Oxford Economic Papers* 56: 209–230.

Fertig. M and J. Kluve (2005): *The Effect of Age at School Entry on Educational Attainment in Germany*, IZA Discussion Paper No. 1507.

Fredriksson, P. and B. Öckert (2005): *Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labour Market Performance*, IZA Discussion Paper No. 1659.

Graue, M.E. and J. DiPerna (2000): Redshirting and Early Retention: Who Gets the “Gift of Time” and What Are Its Outcomes?, *American Educational Research Journal* 37: 509-534.

Hanushek, E.A. and L. Wößmann (2005): *Does Educational Tracking Affect Performance and inequality?: Differences-in-Differences Evidence across Countries*, NBER working paper series 11124.

Hutchison, D. and C. Sharp (1999): *A Lasting Legacy? The Persistence of Season of Birth Effects*, NFER Conference Paper, presented at the British Educational Research Association Conference, University of Brighton, September.

Imbens, G.W. and J.D. Angrist (1994): Identification and Estimation of Local Average Treatment Effects, *Econometrica* 62: 467-475.

Jimerson, Sh., E. Carlson, M. Rotert, B. Egeland and L.A. Sroufe (1997): A Prospective, Longitudinal Study of the Correlates and Consequences of Early Grade Retention, *Journal of School Psychology* 35: 3-25.

Kern, A. (1951): *Sitzenbleiberelend und Schulreife*, Verlag Herder, Freiburg.

Kinard, E.M. and H. Reinherz (1986): Birthdate Effects on School Performance and Adjustment: A Longitudinal Study, *Journal of Educational Research* 79: 366-372.

Leuven, E., M. Lindahl, H. Oosterbeek and D. Webbink (2004): *New evidence on the effect of time in school on early achievement*, HEW 0410001, Economics Working Paper Archive at WUSTL.

May, D.C., D.K. Kundert, and D. Brent (1995): Does Delayed School Entry Reduce Later Grade Retentions and Use of Special Education Services?, *Remedial and Special Education* 16: 288-294.

Mayer, S. and D. Knutson (1999): Does the Timing of School Affect How Much Children Learn?, in S. Mayer and P. Peterson (eds): *Earning and Learning: How Schools Matter*, Brookings Institution Press, Washington D.C.

Morrison, F.J., E.M. Griffith and D.M. Alberts (1997): Nature-Nurture in the Classroom: Entrance Age, School Readiness, and Learning in Children, *Development Psychology* 33: 254-262.

Proctor, T.B., K.N. Black and J.F. Feldhusen (1986): Early Admission of Selected Children to Elementary School: A Review of the Research Literature, *Journal of Education Research* 80, 70-76.

Sharp, C. (2002): *School Starting Age: European Policy and Recent Research*, NFER Conference Paper, presented at the LGA seminar 'When Should Our Children Start School?', LGA Conference Centre, London.

Staiger, D. and J.H. Stock (1997): Instrumental Variables Regression with Weak Instruments, *Econometrica* 65: 557-586.

Stipek, D (2002): At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents, *Social Policy Report* 16, 3-16.

Stipek, D. and P. Byler (2001): Academic achievement and social behaviors associated with age of entry into kindergarten, *Applied Developmental Psychology* 22: 175-189.

Stock, J.H., J.H. Wright and M. Yogo (2002): Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics* 4: 518-529.

Strøm, B. (2004): *Student achievement and birthday effects*, Mimeo, Norwegian University of Science and Technology.

Zill, N., L.S. Loomis and J. West (1997): *National Household Education Survey. The Elementary School Performance and Adjustment of Children Who Enter Kindergarten Late or Repeat Kindergarten: Findings from National Surveys*, NCES Statistical Analysis Report 98-097, Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Table 1: Compulsory School Starting Age by Country

Age 4	Age 5	Age 6	Age 7
Northern Ireland	Australia (Tasmania)	Austria	Bulgaria
Netherlands (from 8/02)	England	Australia*	Canada
	Malta	Belgium	Denmark
	Netherlands (until 8/02)	Cyprus	Estonia
	New Zealand	Czech Republic	Finland
	Scotland	France	Latvia
	Wales	Germany	Poland
		Greece	Romania
		Hong-Kong	Singapore
		Hungary	Sweden
		Iceland	*Switzerland*
		Republic of Ireland	
		Italy	
		Japan	
		Korea	
		Liechtenstein	
		Lithuania	
		Luxembourg	
		Norway	
		Portugal	
		Slovakia	
		Slovenia	
		Spain	
		Switzerland	
		U.S.A.	

Note: Based on information from 2002. * Except the state of Tasmania. In *Switzerland* entry age differs by region.
Sources: Sharp (2002) and Bertram and Pascal (2002).

Table 2: Variables Included in the Regression Models

Group of Regressors	PIRLS 2001	Administrative Data for *Hessen*	Youth and Young Adult Longitudinal Survey
Exo1	Entry age only	Entry age only	Entry age only
Exo2	Exo 1 + Gender	Exo 1 + Gender + Entry cohorts + County indicators	Exo1 + Sampling variables [a)]
Exo3	Exo2 + Cultural variables (immigrant [b)])	Exo2 + Cultural variables (country of origin)	Exo2 + Cultural variables (religion [c)])
Exo4	Exo3 + Parental education [d)]		Exo3 + Parental education [d)]
Exo5			Exo4 + First school day variables [e)]
Exo6	Exo 4 + Family background [f)]		Exo5 + Family background [g)]

Note: [a)] Sampling variables include: dummy variables for gender, year of birth, region, and city size. [b)] Immigrant background is controlled for by a dummy variable indicating whether the student or his/her parents were born abroad or if the student often speaks a foreign language at home. [c)] We use information on whether individuals characterize themselves as being (1) Christian, (2) believing in a non-Christian religion or (3) not religious. [d)] Three categories of parental education are defined: (1) academic education, (2) non-academic degree, (3) no vocational degree. [e)] Includes dummy variables on whether the children received a gift at the first school day and whether the parents attended the school entrance ceremony. [f)] Includes the number of siblings and its square and the number of books at home. [g)] Includes the number of siblings and its square and home/parental background variables (see Table 5 for details).

Table 3: Simple Correlations Between Instruments and Observables (PIRLS)

Sample	Full Sample: born January-December			Discontinuity Samples: born June/July	born July/Aug.	born Aug./Sept.
Instruments / Observable Characteristics	*I4* (cut-off: June)	*I4* (cut-off: July)	*I4* (cut-off: August)	*I4* (cut-off: June)	*I4* (cut-off: July)	*I4* (cut-off: August)
Added in Exo2: Gender (Reference = Female):						
Male	0.00	0.00	0.01	0.03	-0.02	0.04
Added in Exo3: Immigration (Reference = No immigrant background):						
Immigrant	0.00	-0.01	0.00	0.04	-0.02	-0.00
Missing: Immigrant	-0.02	0.01	0.00	-0.03	0.05**	-0.03
Added in Exo4: Parental Education (Reference = No vocational degree)						
Father: Academic degree	0.00	0.00	-0.02	0.00	0.03	-0.03
Mother: Academic degree	0.00	0.01	0.01	-0.02	0.01	-0.00
Father: Non-academic degree	0.01	-0.02	-0.02**	0.03	-0.03	-0.03
Mother: Non-academic degree	0.00	-0.02*	-0.02	0.02	-0.03	0.00
Missing: Education of Father	0.00	0.01	0.02*	-0.03	0.00	0.03
Missing: Education of Mother	0.00	0.01	0.01	-0.01	0.01	-0.01
Added in Exo6: Family Background						
Number of siblings	0.01	0.01	-0.01	-0.01*	0.03	-0.02
Missing: Number of siblings	-0.02**	0.00	-0.01	-0.05	0.05*	-0.01
Log number of books at home	0.01	0.00	-0.02	0.02	0.01	-0.02
Missing: Log number of books	-0.01	0.01	0.02	-0.03	0.02	0.00
Number of observations	*6,591*	*6,591*	*6,591*	*1,123*	*1,226*	*1,262*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2. Exo1 includes only the age of school entry.

Source: PIRLS 2001. Own calculations.

.

Table 4: Simple Correlations Between Instruments and Observables (Administrative Data for *Hessen*)

Sample	Full Sample: born January-December		Discontinuity Samples: born June/July	Discontinuity Samples: born July/Aug.
Instruments / Observable Characteristics	*I1* = I4 (Hamb. Accord)	*I3=I4* (cut-off: July)	*I1* = I4 (Hamb. Accord)	*I3=I4* (cut-off: July)
Added in Exo2: Gender (Reference = Female), Entry Cohort (Refer. = 1997) and County Indicators:				
Gender dummy variable (Male = 1)	0.00	0.00*	0.00	0.01
School entry in 1998	0.01**	0.01**	0.00	0.01
School entry in 1999	0.00*	0.00	0.01*	-0.01
County indicator 1 (Darmstadt)	0.00	0.00	0.00	0.00
County indicator 2 (Frankfurt)	0.00	0.00	0.01	0.00
County indicator 3 (Offenbach Stadt)	0.00	0.00*	0.00	0.01
County indicator 4 (Wiesbaden)	-0.01**	-0.01**	0.00	-0.01*
County indicator 5 (Bergstraße / Odenwald)	0.01**	0.00	0.01	-0.01
County indicator 6 (Darmstadt-Dieburg)	0.00	0.00	-0.01	0.00
County indicator 7 (Groß-Gerau)	-0.01**	0.00	-0.01**	0.01**
County indicator 8 (Hochtaunus)	0.00	0.00	0.00	0.00
County indicator 9 (Main-Kinzig)	0.00	0.00	0.00	0.00
County indicator 10 (Offenbach)	0.00**	0.00	0.00	0.00
County indicator 11 (Rheingau-Taunus)	0.00	0.00	0.00	0.00
County indicator 12 (Offenbach)	0.00	0.00	0.00	0.00
County indicator 13 (Wetterau)	0.00	0.00	0.00	0.00
County indicator 14 (Gießen)	0.00	0.00	0.00	0.00
County indicator 15 (Lahn-Dill)	0.00	0.00	0.00	0.00
County indicator 16 (Limburg-Weilburg)	0.00**	0.00	0.01**	0.01*
County indicator 17 (Marburg-Bied./Vogelsb.)	0.00	0.00	-0.01	0.01*
County indicator 18 (Kassel Stadt)	0.00**	0.01**	0.00	0.01
County indicator 19 (Fulda / Hersfeld-Rotenb.)	0.00	0.00	-0.01	0.00
County indicator 20 (Kassel/Werra-Meißner)	0.00	0.00	0.00	0.00
County indicator 21 (Schwalm-Ed./Waldeck-F.)	0.00	0.00	0.00	0.00
Added in Exo3: Country of origin:				
Country 1 (German speaking countries)	0.01**	0.01**	0.00	0.00
Country 2 (Turkey)	-0.01**	-0.01**	0.00	-0.01
Country 3 (Italy and Greece)	-0.01**	0.00	-0.01**	0.01
Country 4 (Former Yugoslavian states)	0.00	0.00	0.01	0.00
Country 5 (Remaining „Western" countries)	0.00	0.00	-0.01	0.00
Country 6 (Eastern Europe; former Soviet Un.)	0.00	0.00	0.00	0.00
Country 7 (Remaining Muslim countries)	0.00**	0.00	0.00	0.00
Country 8 (Remaining Asia)	0.00	0.00	0.00	0.00
Country 9 (Remaining countries)	0.00	0.00	0.00	0.00
Number of observations	*182,676*	*182,676*	*32,059*	*32,760*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo3) are explained in Table 2. Exo1 includes only the age of school entry.

Source: Student-Level Data of the Statistics of General Schools for the State of *Hessen* 2004/2005 provided by the State Statistical Office (*Hessisches Statistisches Landesamt*). Own calculations.

Table 5: Simple Correlations Between Instruments and Observables (Youth and Young Adult Longitudinal Survey Data)

Sample	Full Sample: born January-December			Born Sept./Oct.
Instruments / Observable Characteristics	*I1* (Hamburg Accord)	*I2* (School Year Starting Dates)	*I4* (Cut-off: September)	*I4* (Cut-off: September)
Added in Exo 2: Sampling Variables:				
Male (Reference = Female)	0.01	-0.07**	-0.06**	-0.06
Year of birth 1962	0.02	-0.03	-0.03	-0.02
Year of birth 1963	0.02	-0.02	-0.02	-0.09
Year of birth 1964	0.03	-0.01	0.00	0.09
Year of birth 1965	0.06**	0.01	-0.03	-0.09
Year of birth 1966	0.05	0.02	-0.03	-0.09
Year of birth 1967	-0.04	0.02	0.02	0.11
Year of birth 1968	-0.03	0.01	0.03	-0.03
Region 1 (Schleswig-Holstein)	0.03	0.04	0.02	0.03
Region 2 (Bremen)	-0.02	0.03	0.03	0.02
Region 3 (Hamburg)	0.02	0.04	0.02	0.03
Region 4 (Niedersachsen)	-0.04	-0.01	0.02	0.05
Region 5 (Nordrhein-Westfalen)	-0.07**	0.00	0.01	0.04
Region 6 (Hessen)	0.01	0.05	0.04	0.00
Region 7 (Rheinland-Pfalz)	0.02	-0.01	0.01	-0.01
Region 8 (Saarland)	0.02	0.03	0.05*	0.02
Region 9 (Baden-Württemberg)	0.01	-0.03	-0.03	0.02
Region 10 (Bavaria)	0.06**	-0.04	-0.07**	-0.14*
Region 11 (West Berlin)	-0.03	-0.01	0.00	-0.02
< 2,000 residents	0.02	0.04	0.01	-0.05
2,000 - 4,999 residents	-0.01	-0.08**	-0.06**	-0.13*
5,000 - 19,999 residents	0.04	0.00	-0.02	0.01
20000-49,999 residents	-0.02	0.01	0.02	0.11
50,000 - 99,999 residents	0.00	0.00	0.03	0.08
100,000 - 499,999 residents	-0.03	0.00	0.00	0.11
> 499,999 residents	0.00	0.02	0.03	0.07
Added in Exo 3: Cultural Variables (Reference = Not religious):				
Christian religion	-0.02	0.00	-0.01	-0.06
Non-Christian religion	0.00	-0.02	0.00	0.00
Added in Exo 4: : Parental Education (Reference = No vocational degree):				
Father: Academic degree	0.00	0.04	0.05*	0.09
Mother: Academic degree	-0.05*	0.02	0.04	0.06
Father: Non-academic degree	0.01	-0.01	-0.03	-0.04
Mother: Non-academic degree	0.02	0.01	-0.01	-0.09
Added in Exo 5: First School Day Variables:				
Gift received at first school day	0.00	-0.03	-0.01	-0.03
Missing information: Gift received	-0.01	0.04	0.02	0.03
Parents attended entry ceremony	-0.05*	0.01	0.04	0.10
Added in Exo 6: Family Background:				
Number of siblings	0.02	-0.05	-0.05*	-0.06
Parental attitude dummy [a)]	0.00	0.03	0.02	-0.04
Home resources indicator [b)]	-0.02	0.08**	0.07**	0.06
Joint activities dummy [c)]	0.00	0.01	-0.02	-0.14*
Number of observations	*1,199*	*1,199*	*1,199*	*173*

Note: [a)] Indicating whether parents had ambitious perceptions concerning child's future when aged 6-12. [b)] Taking the value of one if person read books/magazines at home when aged 6-12. [c)] Indicating whether parents undertook joint leisure time activities (e.g. music, sports) together with child when she was 6-12 years old. * Significant at ten percent level. ** Significant at five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2. Exo1 includes only age of school entry.
Sources: Data of the German Youth and Young Adult Longitudinal Survey, data on school starting dates. Own calculations.

Table 6: First-Stage Results (PIRLS)

Sample	Full Sample: born January-December			Discontinuity Samples: born June/July	born July/August	born August/Sept.
Instruments/ Specifications	(1a) *I4* ("Hamburg Accord")	(2a) *I4* (cut-off: July)	(3a) *I4* (cut-off: August)	(1b) *I4* ("Hamburg Accord")	(2b) *I4* (cut-off: July)	(3b) *I4* (cut-off: August)
Exo1	0.49**	0.35**	0.32**	0.40**	-0,10**	0.27**
(*F*-statistic)	(433.1)	(209.3)	(176.4)	(86.7)	(5.3)	(73.3)
Exo2	0.49**	0.35**	0.32**	0.40**	-0,10**	0.27**
(*F*-statistic)	(427.1)	(213.5)	(177.4)	(89.1)	(5.5)	(71.2)
Exo3	0.49**	0.35**	0.32**	0.40**	-0,10**	0.27**
(*F*-statistic)	(426.5)	(216.2)	(175.5)	(90.6)	(5.4)	(70.3)
Exo4	0.49**	0.35**	0.32**	0.40**	-0,10**	0.27**
(*F*-statistic)	(440.8)	(218.9)	(174.0)	(94.6)	(5.8)	(70.4)
Exo6	0.49**	0.35**	0.31**	0.40**	-0.09**	0.26**
(*F*-statistic)	(428.6)	(210.9)	(170.6)	(95.1)	(5.1)	(68.0)
Observations	*6,591*	*6,591*	*6,591*	*1,123*	*1,226*	*1,262*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2.

Source: PIRLS 2001. Own calculations.

Table 7: First-Stage Results (Administrative Data for *Hessen*)

Sample	Full Sample: born January-December		Discontinuity Samples: born June/July	born July/August
Instruments/ Specifications	(1a) *I1*=I4 (Hamburg Accord)	(2a) *I3=I4* (cut-off: July)	(1b) *I1=I4* (Hamburg Accord)	(2b) *I3=I4* (cut-off: July)
Exo1	0.49**	0.32**	0.41**	0.04**
(*F*-statistic)	(8196.0)	(3456.7)	(2277.1)	(21.6)
Exo2	0.49**	0.32**	0.41**	0.04**
(*F*-statistic)	(8189.0)	(3443.0)	(2306.4)	(19.6)
Exo3	0.49**	0.32**	0.41**	0.04**
(*F*-statistic)	(8321.2)	(3499.6)	(2325.5)	(20.3)
Observations	*182,676*	*182,676*	*32,059*	*32,760*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo3) are explained in Table 2.

Source: Student-Level Data of the Statistics of General Schools for the State of *Hessen* 2004/2005 provided by the State Statistical Office (*Hessisches Statistisches Landesamt*). Own calculations.

Table 8: First-Stage Results (Youth and Young Adult Longitudinal Survey Data)

Sample	Full Sample: born January-December				Discontinuity Sample: born Sept./Oct.
Instruments/ Specifications	(1a) *I1* (Hamburg Accord)	(2a) *I2* (School year starting dates)	(3a) *I4* (cut-off: August)	(4a) *I4* (cut-off: September)	(4b) *I4* (cut-off: September)
Exo1	0.08	0.79**	0.59**	0.75**	0.72**
(*F*-statistic)	(1.8)	(150.0)	(80.1)	(139.2)	(35.7)
Exo2	0.04	0.78**	0.58**	0.74**	0.66**
(*F*-statistic)	(0.6)	(147.5)	(75.5)	(136.8)	(35.8)
Exo3	0.04	0.78**	0.58**	0.74**	0.68**
(*F*-statistic)	(0.6)	(148.2)	(75.3)	(136.9)	(35.6)
Exo4	0.04	0.78**	0.59**	0.74**	0.70**
(*F*-statistic)	(0.6)	(153.5)	(78.3)	(141.6)	(38.1)
Exo5	0.05	0.78**	0.59**	0.74**	0.70**
(*F*-statistic)	(0.7)	(153.1)	(78.1)	(139.7)	(37.8)
Exo6	0.05	0.78**	0.59**	0.74**	0.66**
(*F*-statistic)	(0.8)	(152.3)	(79.3)	(139.0)	(26.8)
Observations	*1,199*	*1,199*	*1,199*	*1,199*	*173*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2.

Sources: Data of the German Youth and Young Adult Longitudinal Survey, data on school starting dates. Own calculations.

Table 9: OLS and Second-Stage Results (PIRLS)

Sample	FullSample: born January-December				Discontinuity Samples: born June/July	born July/Aug.	born Aug./Sept.
	(0) OLS	(1a) *I4* ("Hamburg Accord")	(2a) *I4* (cut-off: July)	(3a) *I4* (cut-off: August)	(1b) *I4* ("Hamburg Accord")	(2b) *I4* (cut-off: July)	(3b) *I4* (cut-off: August)
Exo1	-12.80**	30.74**	29.43**	4.34	28.17**	-42.16	-7.53
(s.e.)	(3.0)	(6.2)	(8.5)	(11.5)	(13.2)	(52.5)	(20.9)
Exo2	-11.49**	30.64**	30.20**	6.02	28.18**	-42.15	-5.90
(s.e.)	(3.0)	(6.3)	(8.5)	(11.6)	(13.1)	(52.4)	(21.3)
Exo3	-8.65**	27.14**	27.52**	4.20	28.98**	-46.50	-9.31
(s.e.)	(2.7)	(6.2)	(7.9)	(10.5)	(12.6)	(50.7)	(20.1)
Exo4	-4.57**	27.37**	31.29**	10.12	26.41**	-54.00	-3.10
(s.e.)	(2.3)	(5.8)	(7.8)	(10.2)	(11.5)	(48.8)	(18.8)
Exo6	-1.24	26.77**	30.14**	11.55	25.83**	-43.70	0.34
(s.e.)	(2.2)	(5.6)	(7.2)	(9.9)	(11.2)	(50.2)	(19.5)
Obs.	*6,591*	*6,591*	*6,591*	*6,591*	*1,123*	*1,226*	*1,262*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2.

Source: PIRLS 2001. Own calculations.

Table 10: OLS and Second-Stage Results (Administrative Data for *Hessen*)

Sample	Full Sample: born January-December			Discontinuity Sample: born June/July	born July/August
	(0) OLS	(1a) *I1=I4* (Hamburg Accord)	(2a) *I3=I4* (cut-off: July)	(1b) *I1=I4* (Hamburg Accord)	(2b) *I3=I4* (cut-off: July)
Exo1	-0.37**	0.45**	0.48**	0.40**	1.37**
(s.e.)	(0.01)	(0.03)	(0.04)	(0.05)	(0.66)
Exo2	-0.36**	0.44**	0.49**	0.38**	1.57**
(s.e.)	(0.01)	(0.03)	(0.04)	(0.05)	(0.71)
Exo3	-0.31**	0.41**	0.45**	0.37**	1.49**
(s.e.)	(0.01)	(0.03)	(0.04)	(0.05)	(0.67)
Obs.	*182,676*	*182,676*	*182,676*	*32,059*	*32,760*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo3) are explained in Table 2.

Source: Student-Level Data of the Statistics of General Schools for the State of *Hessen* 2004/2005 provided by the State Statistical Office (*Hessisches Statistisches Landesamt*). Own calculations.

Table 11: OLS and Second-Stage Results (Youth and Young Adult Longitudinal Survey Data)

Sample	Full Sample: born January-December					Discontinuity Sample: born Sept./Oct.
	(0) OLS	(1a) *I1* (Hamburg Accord)	(2a) *I2* (School year starting dates)	(3a) *I4* (cut-off: August	(4a) *I4* (cut-off: September)	(4b) *I4* (cut-off: September)
Exo1	-0.21**	0.47	0.37	0.77**	0.47*	-0.01
(s.e.)	(0.08)	(2.04)	(0.23)	(0.32)	(0.25)	(0.39)
Exo2	-0.21**	0.40	0.27	0.68**	0.46*	-0.16
(s.e.)	(0.09)	(3.33)	(0.23)	(0.32)	(0.24)	(0.44)
Exo3	-0.21**	0.43	0.27	0.68**	0.46*	-0.23
(s.e.)	(0.09)	(3.33)	(0.23)	(0.32)	(0.25)	(0.42)
Exo4	-0.17**	0.88	0.12	0.47*	0.27	-0.46
(s.e.)	(0.08)	(3.22)	(0.21)	(0.28)	(0.22)	(0.39)
Exo5	-0.16**	0.50	0.14	0.50*	0.30	-0.40
(s.e.)	(0.08)	(2.90)	(0.21)	(0.28)	(0.22)	(0.39)
Exo6	-0.14*	0.22	0.09	0.39	0.29	-0.31
(s.e.)	(0.08)	(2.65)	(0.20)	(0.27)	(0.22)	(0.44)
Obs.	*1,199*	*1,199*	*1,199*	*1,199*	*1,199*	*173*

Note: * Significant at the ten percent level. ** Significant at the five percent level. The different specifications (Exo1 – Exo6) are explained in Table 2.
Sources: Data of the German Youth and Young Adult Longitudinal Survey, data on school starting dates. Own calculations.

Table 12: Subgroup Results for the PIRLS Data

	First Stage		Second Stage	
Male – Native	Full sample	0.45**	Full sample	42.86**
	(*F*)	(138.9)	(s.e.)	(8.6)
(Full sample: 2,642 observations;	born June/July	0.30**	born June/July	59.83**
born June/July: 447 observations)	(*F*)	(21.6)	(s.e.)	(22.5)
Female - Native	Full sample	0.56**	Full sample	16.23**
	(*F*)	(244.7)	(s.e.)	(8.4)
(Full sample: 2,717 observations;	born June/July	0.52**	born June/July	7.25
born June/July: 469 observations)	(*F*)	(104.5)	(s.e.)	(12.8)
Male – Immigrant	Full sample	0.44**	Full sample	20.50
	(*F*)	(33.4)	(s.e.)	(20.2)
(Full sample: 668 observations;	born June/July	0.43**	born June/July	67.38*
born June/July: 109 observations)	(*F*)	(17.7)	(s.e.)	(36.2)
Female - Immigrant	Full sample	0.38**	Full sample	37.65
	(*F*)	(10.8)	(s.e.)	(30.0)
(Full sample: 564 observations;	born June/July	0.30**	born June/July	-4.06
born June/July: 98 observations)	(*F*)	(4.6)	(s.e.)	(62.1)
Parents: Academic Degree	Full sample	0.35**	Full sample	29.36*
	(*F*)	(45.2)	(s.e.)	(17.0)
(Full sample: 1,330 observations;	born June/July	0.29**	born June/July	32.11
born June/July: 223 observations)	(*F*)	(10.1)	(s.e.)	(30.5)
Parents: No Academic Degree	Full sample	0.53**	Full sample	25.71**
	(*F*)	(438.6)	(s.e.)	(5.9)
(Full sample: 5,261 observations;	born June/July	0.43**	born June/July	24.14**
born June/July: 900 observations)	(*F*)	(97.1)	(s.e.)	(11.6)

Note: Effects for the full specifications (Exo6) using the Hamburg Accord based instrument. * Significant at the ten percent level. ** Significant at the five percent level. *F* refers to the *F*-statistics of joint significance of the instruments in the first-stage regressions.
Source: PIRLS 2001. Own calculations.

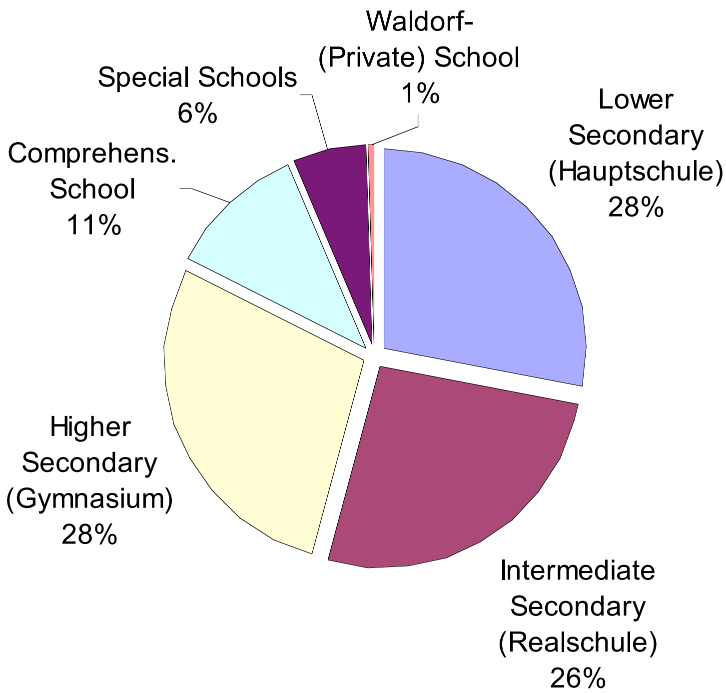
Table 13: Subgroup Results for the Administrative Data from the State of Hessen

	First Stage		Second Stage	
Male – Native (German speaking	Full sample	0.50**	Full sample	0.41**
countries)	(*F*)	(3885.8)	(s.e.)	(0.04)
(Full sample: 79,400 observations;	born June/July	0.41**	born June/July	0.35**
born June/July: 13,898 observations)	(*F*)	(1025.0)	(s.e.)	(0.08)
Female – Native (German speaking	Full sample	0.50**	Full sample	0.45**
countries)	(*F*)	(3845.2)	(s.e.)	(0.04)
(Full sample: 77,106 observations;	born June/July	0.41**	born June/July	0.39**
born June/July: 13,555 observations)	(F)	(1039.2)	(s.e.)	(0.08)
Male – Turkish	Full sample	0.46**	Full sample	0.21
	(*F*)	(221.0)	(s.e.)	(0.14)
(Full sample: 5,772 observations;	born June/July	0.42**	born June/July	0.33
born June/July: 1,009 observations)	(F)	(62.5)	(s.e.)	(0.23)
Female - Turkish	Full sample	0.49**	Full sample	0.32**
	(*F*)	(255.5)	(s.e.)	(0.13)
(Full sample: 5,647 observations;	born June/July	0.45**	born June/July	0.32
born June/July: 1,045 observations)	(F)	(88.3)	(s.e.)	(0.22)
Male – Predominantly Muslim	Full sample	0.36**	Full sample	0.37
Countries (without Turkey)	(*F*)	(25.0)	(s.e.)	(0.41)
(Full sample: 1,539 observations;	born June/July	0.31**	born June/July	-0.24
born June/July: 247 observations)	(F)	(6.2)	(s.e.)	(0.72)
Female - Predominantly Muslim	Full sample	0.35**	Full sample	0.55
Countries (without Turkey)	(*F*)	(26.3)	(s.e.)	(0.40)
(Full sample: 1,474 observations;	born June/July	0.43**	born June/July	1.00*
born June/July: 248 observations)	(F)	(16.0)	(s.e.)	(0.55)
Male - Italy/Greece	Full sample	0.52**	Full sample	-0.16
	(*F*)	(86.9)	(s.e.)	(0.26)
(Full sample: 1,462 observations;	born June/July	0.37**	born June/July	0.34
born June/July: 271 observations)	(F)	(22.5)	(s.e.)	(0.61)
Female – Italy/Greece	Full sample	0.51**	Full sample	-0.07
	(*F*)	(67.1)	(s.e.)	(0.27)
(Full sample: 1,419 observations;	born June/July	0.50**	born June/July	-0.57
born June/July: 244 observations)	(F)	(31.3)	(s.e.)	(0.44)
Male - Former Yugoslavia	Full sample	0.46**	Full sample	0.04
	(*F*)	(48.9)	(s.e.)	(0.34)
(Full sample: 1,217 observations;	born June/July	0.51**	born June/July	0.01
born June/July: 213 observations)	(F)	(20.1)	(s.e.)	(0.51)
Female - Former Yugoslavia	Full sample	0.45**	Full sample	0.95**
	(*F*)	(46.2)	(s.e.)	(0.41)
(Full sample: 1,190 observations;	born June/July	0.38**	born June/July	1.09
born June/July: 221 observations)	(*F*)	(15.7)	(s.e.)	(0.76)

Note: Effects for the full specifications (Exo3) using the Hamburg Accord based instrument. * Significant at the ten percent level. ** Significant at the five percent level. *F* refers to the *F*-statistics of joint significance of the instruments in the first-stage regressions.

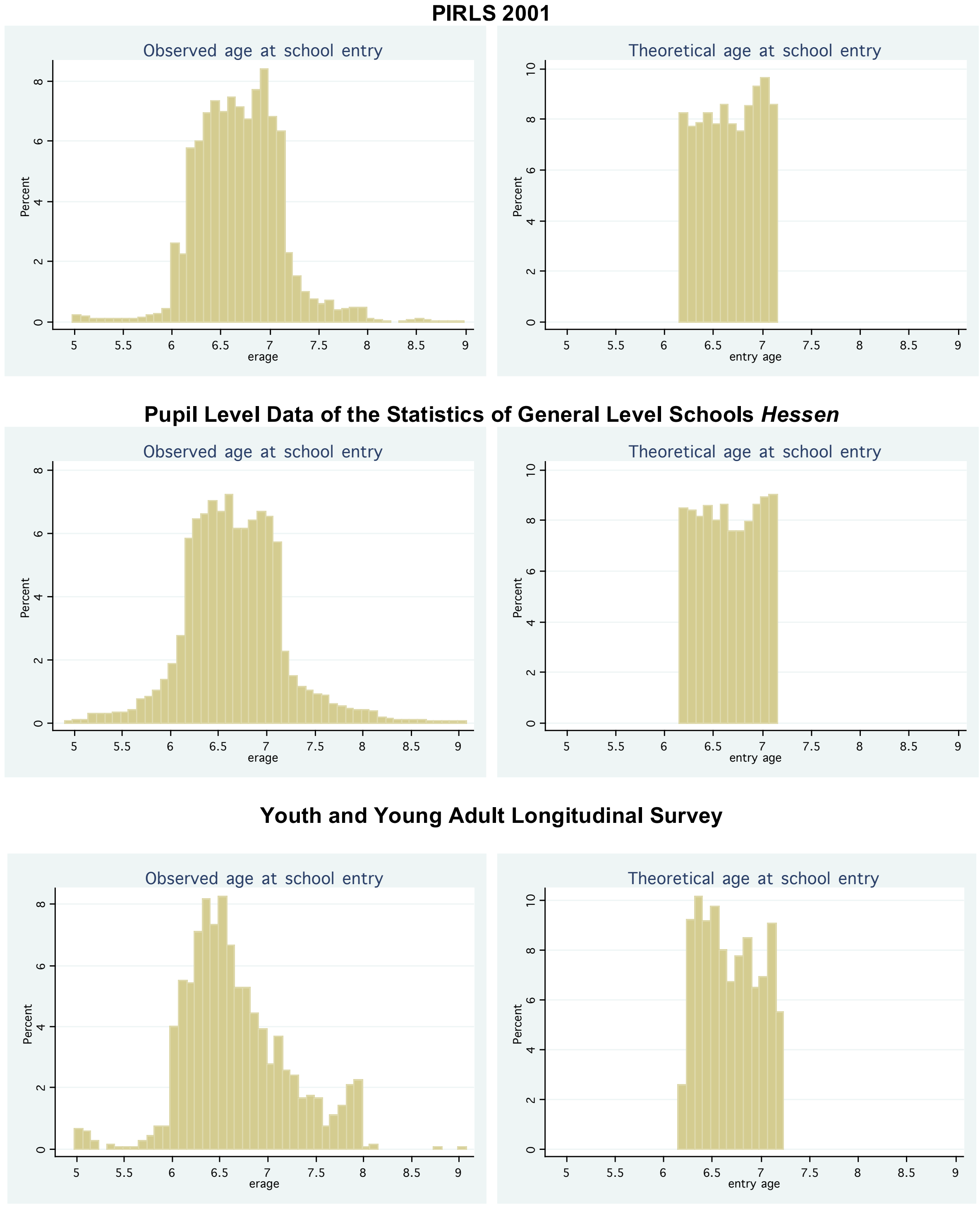
Source: Student-Level Data of the Statistics of General Schools for the State of *Hessen* 2004/2005 provided by the State Statistical Office (*Hessisches Statistisches Landesamt*), data on school starting dates. Own calculations.

Figure 1: The German Tracking System: Graduates in 2003



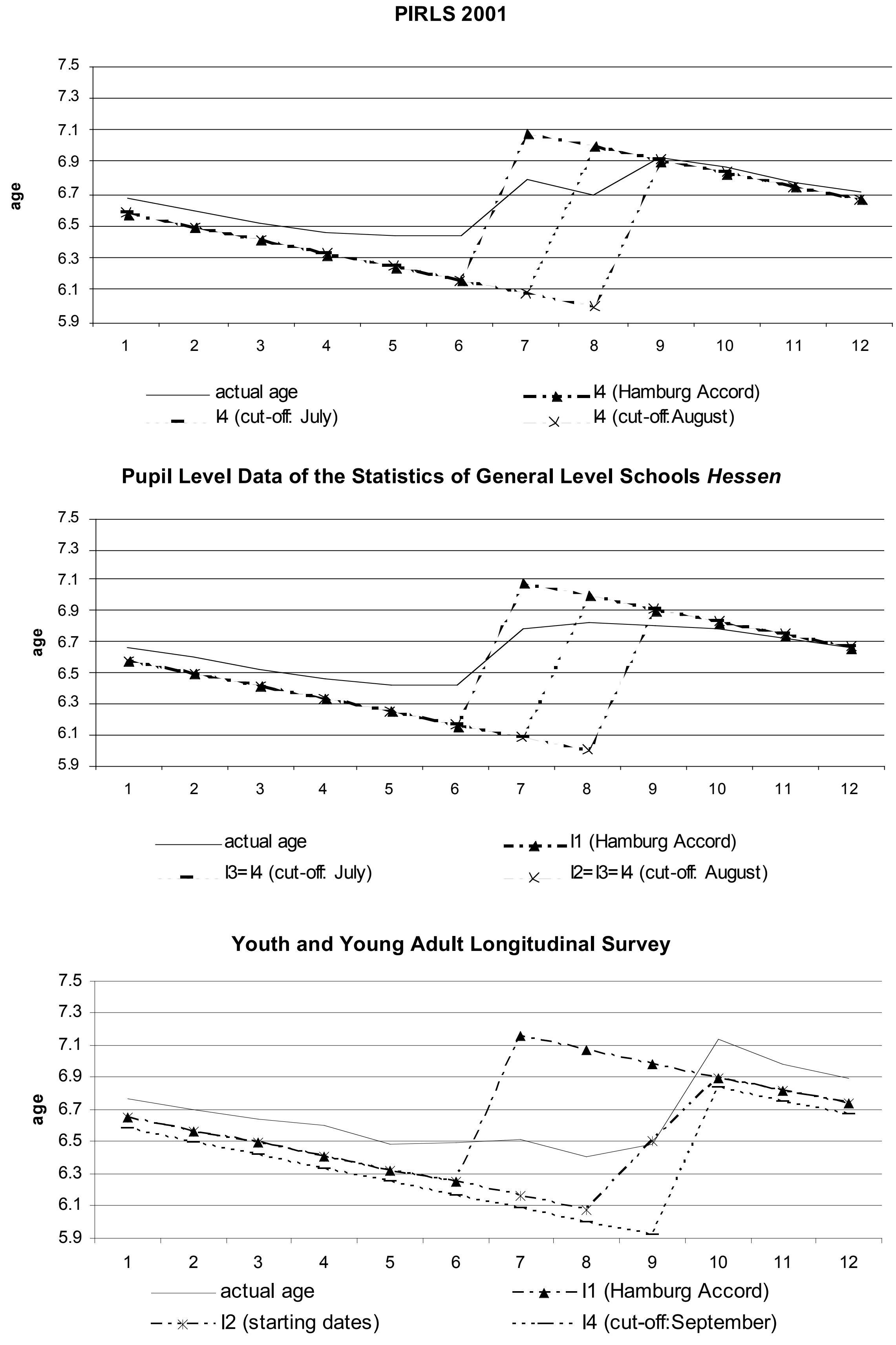
Source: German Federal Statistical Office (2004): *Fachserie 11 / Reihe 1: Bildung und Kultur, Schuljahr 2003/04*, Wiesbaden.

Figure 2: Observed and Theoretical Age at School Entry



Note: Theoretical age at school entry according to the 'Hamburg Accord' (end of June as the cut-off date).
Sources: PIRLS 2001. Pupil-Level Data of the Statistics of General Schools for the State of *Hessen* provided by the State Statistical Office (*Hessisches Statistisches Landesamt*). Youth and Young Adult Longitudinal Survey. Data on school starting dates. Own computations.

Figure 3: Observed and Theoretical Entry Ages by Birth Month



Note: Since there are no state identifiers in the PIRLS data we can only calculate $I^4(b_i, c)$ in this case.

Sources: PIRLS 2001. Pupil-Level Data of the Statistics of General Schools for the State of *Hessen*. Youth and Young Adult Longitudinal Survey. Data on school starting dates. Own computations.

Appendix

Table A1: Start Dates of the School Years in the West German States 1966-1976

	Baden-Wurttemberg	Bavaria	Berlin	Bremen	Hamburg	Hessen	Niedersachsen	Nordrhein-Westfalen	Rheinland-Pfalz	Saarland	Schleswig-Holstein
1966	3-Sep-66	5-Sep-66	20-Aug-66	23-Aug-66	15-Aug-66	16-Aug-66	10-Aug-66	7-Sep-66	23-Aug-66	1-Sep-66	9-Aug-66
1967	6-Sep-67	6-Sep-67	31-Aug-67	26-Aug-67	13-Aug-67	31-Aug-67	14-Aug-67	6-Sep-67	22-Aug-67	2-Sep-67	29-Aug-67
1968	7-Sep-68	10-Sep-68	24-Aug-68	22-Aug-68	24-Aug-68	3-Sep-68	26-Aug-68	8-Aug-68	27-Aug-68	31-Aug-68	27-Aug-68
1969	6-Sep-69	10-Sep-69	23-Aug-69	9-Aug-69	9-Aug-69	3-Sep-69	11-Aug-69	3-Aug-69	27-Aug-69	30-Aug-69	9-Aug-69
1970	8-Sep-70	8-Sep-70	22-Aug-70	8-Aug-70	22-Aug-70	27-Aug-70	5-Aug-70	5-Sep-70	26-Aug-70	20-Aug-70	20-Aug-70
1971	13-Sep-71	6-Sep-71	21-Aug-71	4-Sep-71	21-Aug-71	25-Aug-71	8-Sep-71	14-Aug-71	25-Aug-71	1-Sep-71	21-Aug-71
1972	12-Sep-72	18-Sep-72	19-Aug-72	2-Sep-72	26-Aug-72	26-Aug-72	30-Aug-72	5-Aug-72	16-Aug-72	19-Aug-72	26-Aug-72
1973	3-Sep-73	12-Sep-73	25-Aug-73	25-Aug-73	18-Aug-73	18-Aug-73	22-Aug-73	28-Jul-73	8-Aug-73	11-Aug-73	18-Aug-73
1974	24-Aug-74	16-Sep-74	17-Aug-74	17-Aug-74	10-Aug-74	10-Aug-74	14-Aug-74	7-Sep-74	31-Jul-74	3-Aug-74	10-Aug-74
1975	16-Aug-75	10-Sep-75	9-Aug-75	9-Aug-75	2-Aug-75	2-Aug-75	6-Aug-75	30-Aug-75	31-Jul-75	31-Jul-75	2-Aug-75
1976	14-Aug-76	15-Sep-76	7-Aug-76	7-Aug-76	31-Jul-76	28-Jul-76	4-Aug-76	28-Aug-76	8-Sep-76	11-Sep-76	31-Jul-76

Source: Council of the Education Ministers (*Kultusministerkonferenz*)