Does the early bird catch the worm?

Instrumental variable estimates of early educational effects of age of school entry in Germany

Patrick A. Puhani · Andrea M. Weber

Abstract We estimate the effect of age of school entry on educational outcomes using two different data sets for Germany, sampling pupils at the end of primary school and in the middle of 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 outcomes for pupils who enter school at 7 instead of 6 years of age: test scores at the end of primary school increase by about 0.40 standard deviations and the probability to attend the highest secondary schooling track (*Gymnasium*) increases by about 12% points.

P. A. Puhani (✉)
University of Hannover, Institut für Arbeitsökonomik,
Königsworther Platz 1, 30167 Hannover, Germany
e-mail: puhani@aoek.uni-hannover.de

P. A. Puhani
SIAW, University of St. Gallen, St. Gallen, Switzerland

P. A. Puhani
IZA, Bonn, Germany

A. M. Weber
Darmstadt University of Technology, Fachbereich 1,
Residenzschloss S313/136, Marktplatz 15, 64283 Darmstadt, Germany
e-mail: weber@vwl.tu-darmstadt.de

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. From a theoretical point of view skill formation can be modelled as a process characterised by multiple stages in which early investments are crucial for later investments (cf. Cunha et al. 2006, on life cycle skill formation). In the empirical literature, 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 (2006) 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 6 years old. A look back in history reveals that starting education at the ages 6 or 7 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 7 as a footboy (*Page*). 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 5 months in 1955 and in 1964. Since that time, there has also been a trend to have children with learning problems enter school 1 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 (see Sect 2). Therefore, it seems reasonable to ask whether such policies can be expected to improve educational attainment.

In this article, we estimate the causal effect of varying the age of school entry in Germany between 6 and 7 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 6 and 7 is both a major variation observed internationally for the school starting age and a major issue of discussion in the national German debates. Using two different data sets, we measure the effect of age of school entry at the end of primary school and in the middle of secondary school. Our outcome measures are a test score for primary school pupils and the school track attended, 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. We do not show results based on the same data as used in the previous study by Fertig and Kluve (2005) since we cast doubt on the quality of this data for our purposes (cf. the discussion paper version, Puhani and Weber 2005).

The influence of school entry age on educational outcomes is a well-discussed topic, especially in the US and British empirical educationalist literature.[1] 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 several countries (e.g. the US) 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 US or the UK on the other is that in Germany each child independently of date of birth has to complete at least 9 years of compulsory full-time schooling.[2] In the US and the UK, 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 US and Del Bono and Galindo-Rueda 2004, for the UK).[3] 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 US or the UK.

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 10; primary school usually lasts for 4 years) children are selected into one of three educational tracks: the most academic is *Gynmasium*, usually consisting of 9 further years of schooling, followed by *Realschule* (6 years) and *Hauptschule* (5 years and

---

[1] 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.

[2] The exact rule depends on the state. The 9 or 10 years of compulsory full-time education are followed by either at least 1 additional year of full-time education or by several years of part-time education in a vocational school (*Berufsschule*) within the German apprenticeship system.

[3] To be more precise, in England and Wales children could traditionally (between 1962 and 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.

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 system (cf. Hanushek and Wößmann 2006). Hence, age of school entry may have larger and more lasting effects in Germany than in countries with a comprehensive school system.

The article 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 data sets we use are described in Sect 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. 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 instrument is 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 Sect. 5. We find robust evidence that entering the current German school system at the age of 7 instead of 6 years raises primary school test scores by two fifths of a standard deviation and increases the probability to attend the highest school track (*Gymnasium*) by about 12% points. If we assume that the school track attended will be completed as we observe it in the data, the amount of secondary schooling is increased by almost half a year (about 5 months) on average by entering school 1 year older. Section 6 concludes and reports results from a small-size survey of headmasters and head-mistresses, 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 6 years is equal to the median and mode of the distribution displayed in Table 1. Before the age of 6, 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 8 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

Table 1 Compulsory school starting age by country

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

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)

corresponding school year (normally in August). Children born later are supposed to start school in the following calendar year (again around August).[4] Deviation from the Hamburg Accord may be caused by parents and school principals considering a child (not) mature enough to start school at an early age. Traditionally, the school laws allow for such leeway. In practice, this yields a situation where children born between the official cut-off date 'end of June' and the school year starting date are often admitted to school in the calendar year when they turn 6 years of age. Formally, the Hamburg Accord with its June cut-off date is (by law) the relevant regulation in all German states during the time period referred to in our data sets.

Only after 1997, the Hamburg Accord was made less binding: the Council of the Ministers of Education encouraged the states to deviate from the traditional school entry cut-off date of end of June to allow later cut-off dates (usually up to the end of September). This increased even further the discretion that schools

---

[4] Note that the real start of the school year slightly varies over calendar year and state: 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.

and parents already had de facto. However, today most state laws are still referring to June as the cut-off date while *explicitly* allowing for discretion. Some states (Baden-Württemberg, Bayern, Berlin, Brandenburg, Thüringen) have recently chosen later cut-off dates.

Apart from the school entry regulations, tracking is another feature of the German education system important to the analyses in this article. After 4 years in primary school, pupils usually change to one of three secondary school tracks.[5] The most vocational and least academic level of secondary schooling is called *Hauptschule* (grades 5–9), the intermediate level *Realschule* (grades 5–10) and the most academic level *Gymnasium* (grammar school, grades 5–13).[6] 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% of graduates came from schools outside of the standard tracking system (11% were in comprehensive schools, 6% in special schools and 1% in Waldorf schools), as Fig. 1 shows.

3 Data

We use two different data sets measuring educational outcomes at two stages of pupils' lives. First, the 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. The observed cohorts overlap with those tested in the PIRLS study.[7] Our estimation sample thus contains 182,676 observations. More detail is given in the following subsections.

---

[5] In the East German States of *Berlin* and *Brandenburg*, primary school goes up to grade 6, so that the selection into school tracks starts 2 years later there than in the rest of Germany.

[6] In the East German states of *Sachsen* and *Thüringen*, *Gymnasium* ends after grade 12. 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 12, but this is not relevant for our samples.

[7] 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.
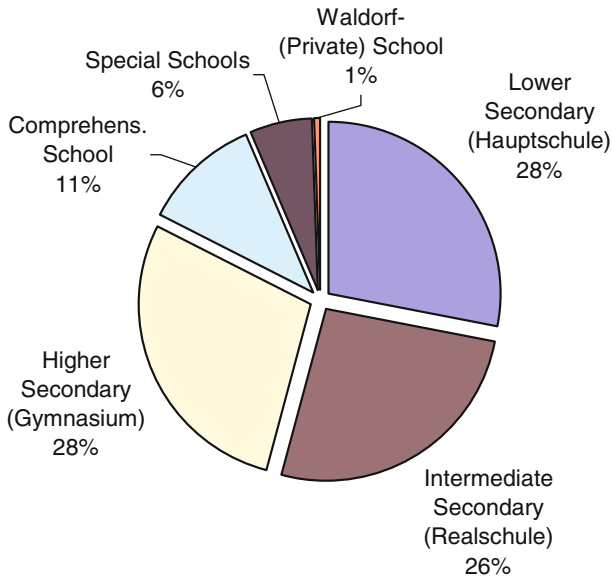
Source: German Federal Statistical Office (2004): Fachserie 11 / Reihe 1: Bildung und Kultur, Schuljahr 2003/04, Wiesbaden.

Fig. 1   The German tracking system: graduates in 2003

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. The underlying reading literacy tests refer to basic competences which are crucial in key situations of daily life and skills required in order to be able to succeed in future education, vocational training and professional life (cf. Bos et al. 2003). More specifically, reading achievement is assessed by different items covering four defined 'reading processes'. These different aspects of reading literacy relate to the ability to 'focus on and retrieve explicitly stated information', 'make straightforward inferences', 'interpret and integrate ideas and information' and 'examine and evaluate content, language and textual elements' (Gonzalez and Kennedy 2003). Each child answers two out of eight 'blocks' of the entire test and individual achievement is scaled using item response theory methods (the scaling methodology is explained in Gonzalez and Kennedy 2003). In order to conduct international comparisons, these test scores have been standardised so that the international mean is 500 and the standard deviation equals 100. For Germany the mean equals 539 and the standard deviation is 67.

Overall, 7,633 pupils at the end of fourth grade in 211 primary schools are sampled in the German PIRLS data. 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. As we lack information on the age of school entry (to the month) for more than 1,000 observations, our effective sample size is reduced to 6,591.[8] As we are interested in estimating the effect of age of school entry on educational outcomes, we might like to sample a birth or school entry cohort and estimate the effect of interest after 4 years of schooling, no matter which grade pupils have achieved by then. The other possibility is to measure educational outcomes at the end of primary school irrespective of how long it took the pupil to reach grade 4. The advantage of the latter approach is that the pupil's performance at grade 4 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 4, 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% of pupils have entered school in 1997, whereas eleven and 2% have entered in 1996 (grade repeaters) and 1998 (grade skippers), respectively. Hence, we observe pupils once they have reached grade 4, even if they have spent only three or even 5 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.

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 collected on behalf of the state Ministry of Education. To our knowledge, this is the first research article 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% of them in grades 6–8. 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

---

[8] 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 outcomes, our estimates based on the selected sample might be biased. However, as we control for parental background and immigrant status, which is likely to be correlated with these characteristics, we hope to reduce this potential bias markedly.

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.[9]

   In the following section, we provide more detail on theoretical and actual age of school entry in our data and suggest an instrumental variable strategy for estimating the effect of age of school entry on educational outcomes.

4 The exogeneity of month of birth and first stage regressions

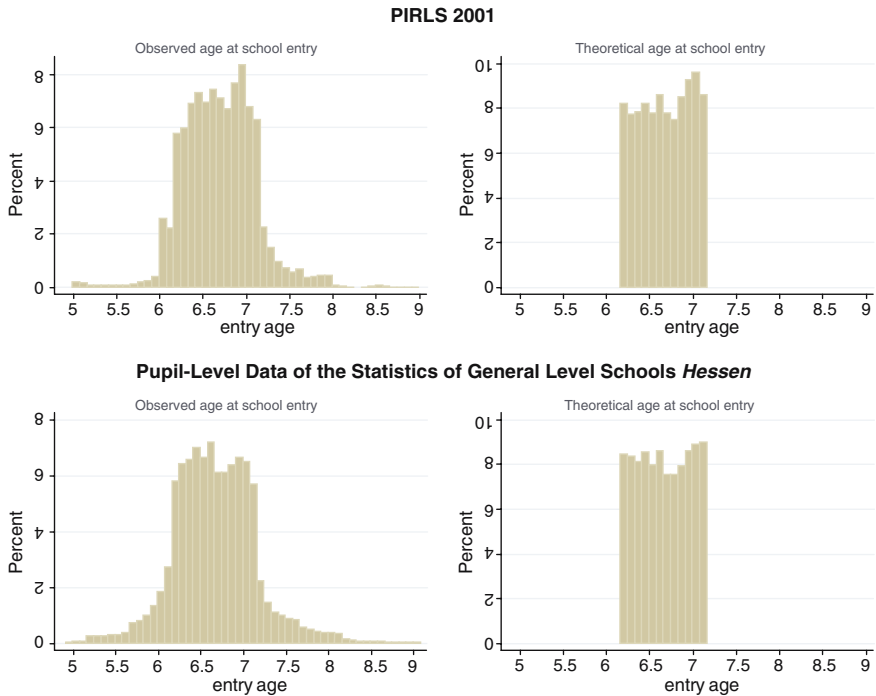### 4.1 The endogeneity of age of school entry

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 (cf. Sect. 2). Thus, ambitious parents may want to push for an early school entry (at age 5) of their child or children with learning problems might be recommended to enter school 1 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 outcomes 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'. The *theoretical school entry age* $I(b_i, s_i)$ is related to a child's month of birth b_i and the month the school year starts s_i in the following way:

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

where the theoretical school entry age $I(b_i, s_i)$ is measured in years (in decimals up to the month). The indicator for the month of birth b_i ranges from 1 to 12, whereas the variation in s_i is generally between the end of July, August, or the beginning of September. If b_i and s_i are exogenous, the theoretical school entry age $I(b_i, s_i)$ 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

---

[9] About 0.86% 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% of whom attend grammar school (*Gymnasium*).

PIRLS 2001

Observed age at school entry — Theoretical age at school entry

Pupil-Level Data of the Statistics of General Level Schools *Hessen*

Observed age at school entry — Theoretical age at school entry

Note: Theoretical age at school entry according to the 'Hamburg Accord'.
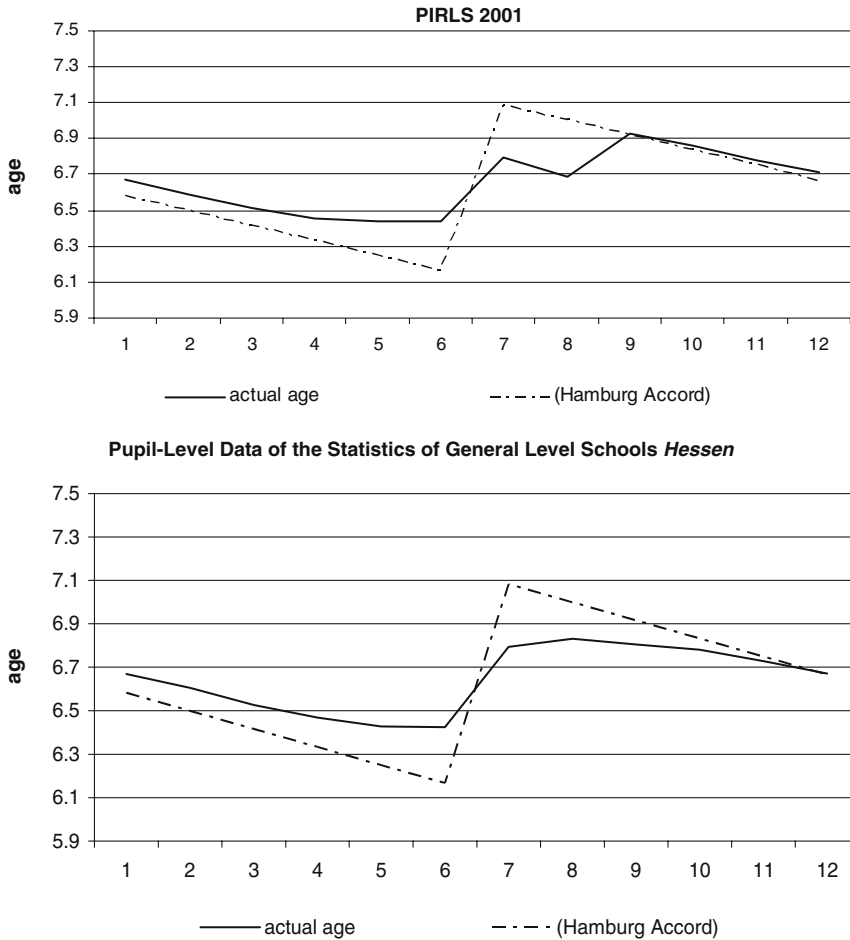
Sources: PIRLS 2001. Pupil-Level Data of the Statistics of General Schools for the State of *Hessen* provided by the State Statistical Office (*Hessisches Statistisches Landesamt*).

Fig. 2 Observed and theoretical age at school entry

year and state. Since we do not have a state identifier in the PIRLS data we assume that August 1st, which is the official nationwide school starting date, is the actual starting date. For the cohorts we observe in the state of *Hessen*, the first year of primary school always started in August.

From Fig. 2, 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). This is because many parents/schools have children start school 1 year later than suggested by the regulations. However, a few children also start school 1 year earlier at about age 5. 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 rule discussed in Sect. 2 is provided in Fig. 3. The first panel displays the actual age of school entry by month of birth in the PIRLS data together with the theoretical age according to the Hamburg Accord. Visual inspection suggests a significant correlation between the theoretical 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 Fig. 2 showing that late entry is more frequent than early

PIRLS 2001

                          actual age   (Hamburg Accord)

Pupil-Level Data of the Statistics of General Level Schools *Hessen*

                          actual age   (Hamburg Accord)

Sources: PIRLS 2001. Pupil-Level Data of the Statistics of General Schools for the State of *Hessen*. Own computation

Fig. 3 Observed and theoretical entry ages by birth month

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. A similar picture concerning non-compliance with the cut-off date of the Hamburg Accord arises in the second panel of Fig. 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.

4.2 Identification strategy

In order to estimate the causal effect of age of school entry on educational outcomes, we adopt an instrumental variable identification strategy (two-stage least

Table 2 Variables included in the regression models

Group of regressors	PIRLS 2001	Administrative data for *Hessen*
Specification 1	Entry age only	Entry age only
Specification 2	Specification 1 + gender	Specification 1 + gender + entry cohorts + county indicators
Specification 3	Specification 2 + cultural variables (immigrant[a])	Specification 2 + cultural variables (country of origin)
Specification 4	Specification 3 + parental education[b]	
Specification 5	Specification 4 + family background[c]	

[a] 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
[b] Three categories of parental education are defined: (1) academic education, (2) non-academic degree, (3) no vocational degree
[c] Includes the number of siblings and its square and the number of books at home

squares, 2SLS). The instrument for the endogenous age of school entry is the theoretical age of school entry as prescribed by the Hamburg Accord, where the school starting month is set to August as explained in the previous subsection: $I(b_i, s_i = 8)$. In order for the instrument to be valid, it has 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 instrument is truly exogenous, *i.e.* uncorrelated with any *unobserved* factors that might influence educational performance, an assumption we cannot test directly, we test whether it is correlated with *observed* variables that we believe might influence educational performance.

Table 2 lists the groups of regressors that we include in the 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 Sect. 5). The set of variables is partly determined by data availability in the respective data sets. In the first set of regressions (specification 1) 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 (driven by variation in month of birth) 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 outcomes consistently in a 2SLS estimation procedure. Nevertheless, control variables that influence educational outcomes may reduce the standard errors of the estimates.

As a first extension of the set of regressors (specification 2), we therefore include gender and regional indicators (the latter are only available in the data for the state of *Hessen*). In the administrative data for *Hessen*, we also control for the school entry cohort among 'specification 2'. The third set of regressors (specification 3) adds cultural background, measured either by an immigration or nationality indicator. The fourth extension (specification 4) adds parental education, which is available in the PIRLS data but not in the administrative data for *Hessen*. The fifth addition (specification 5) adds family background variables, i.e. the number of books at home and the number of siblings, which is

again only possible for the PIRLS data. We consider the control variables added in 'specification 5' 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 buy books if their children are (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 outcomes.

Although low correlations between the instrument and observable variables are supportive of the instrument's exogeneity, they do not provide a guarantee. Recent evidence from medical studies suggests that birth month, which drives our instrument, might exert some direct effect on physical and psychological health (e.g. Willer et al. 2005). Furthermore, our instrument might be endogenous if parents plan the month in which a child is born or if, for example, better educated parents prefer certain birth months over others (cf. the discussion in Bound et al. 1995).

Therefore, we do not exclusively rely on a 'traditional' instrumental variable approach. Drawing on a 'fuzzy regression discontinuity design' (cf. Hahn et al. 2001), our main results relate to a narrow sampling window where only students born in the 2 months adjacent to the respective school entry cut-off point are included in the 2SLS regressions. By restricting the samples to persons born just in June and July, we hope to eliminate any potential direct seasonal effects which might affect the validity of the instrument. Furthermore, any differences in parental attitudes reflected in planned timing of births should be minimised for children born in 2 adjacent months, as it is hard to assure for a child to be born in a very specific month.

In Tables 3 and 4 we display the simple correlations between the instrument and the full set of control variables for different sampling windows. Correlations significant at the 10 or 5% level are marked with one or two asterisks, respectively. As Table 3 shows, the maximum correlation for the PIRLS data equals 0.02 in absolute value, which is very small. Hence, the few correlations of the instrument with regressors that are significantly different from zero are very close to zero. This finding is even more striking for the large administrative data set for *Hessen* in Table 4: no correlation is larger than 0.01 in absolute value. Our instrument (driven by month of birth) thus seems unrelated to gender, the district of residence and the country of origin. Table 3 also shows that the instrument is virtually unrelated to parental education, the number of siblings and the number of books in the household.

4.3 First-stage regressions

Having discussed the exogeneity of our instrument and the use of different sampling windows, we now check the second condition for a valid instrument, namely the (partial) correlation with the variable to be instrumented (age of school entry). Tables 5 and 6 report coefficients of the instrument together with the F-statistics of the tests for significance of the instrument in the first-stage regressions of the 2SLS estimation procedure. A rule of thumb states that an

Table 3 Simple correlations between instrument and observables (PIRLS)

Sampling window/ observable characteristics	June/July	June–September	January–December
Added in specification 2: gender (reference = female)			
Male	0.03	0.02	0.00
Added in specification 3: immigration (reference = no immigrant background)			
Immigrant	0.04	0.02	0.00
Missing: Immigrant	−0.03	0.00	−0.02
Added in Specification 4: parental education (reference = no vocational degree)			
Father: academic degree	0.00	0.01	0.00
Mother: academic degree	−0.02	−0.01	0.00
Father: non-academic degree	0.03	0.01	0.01
Mother: non-academic degree	0.02	0.00	0.00
Missing: education of father	−0.03	−0.01	0.00
Missing: education of mother	−0.01	0.00	0.00
Added in specification 5: family background			
Number of siblings	−0.01*	0.00	0.01
Missing: Number of siblings	−0.05	−0.02	−0.02**
Log number of books at home	0.02	0.02	0.01
Missing: Log number of books	−0.03	−0.02	−0.01
Number of observations	1,123	2,943	6,591

Note: *Significant at the 10% level. **Significant at the 5% level. The different specifications (specifications 1–5) are explained in Table 2. Specification 1 includes only the age of school entry
Source: PIRLS 2001. Own calculations

F-statistic below about ten is indicative of a weak instrument problem (Staiger and Stock 1997; Stock et al. 2002).[10] The tables therefore display the F-statistics for various specifications (specification 1–5) as outlined in Sect. 4.2.

Tables 5 and 6 clearly show that, in both data sets, we have an instrument with an F-statistic largely above the threshold value of ten. The degree of compliance with the rule can be seen from the coefficients reported in the tables. Using the narrowest sampling window of persons born in the 2 months adjacent to the respective cut-off date reveals that the compliance with the Hamburg Accord is significant with a coefficient of 0.40 in the PIRLS data (Table 5) and 0.41 in the *Hessen* data (Table 6). In the discontinuity sample, this means that the share of compliers is about 40%. The coefficient is slightly higher if we widen our sampling window to include pupils born until the end of September. Note that using the full samples of pupils born in any month (January–December sampling window) the degree of compliance is also influenced by the compliance with the assigned variation in school entry age between individuals born in months like January or April, i.e. born in months distant from the official cut-off dates. We expect that non-compliance is lower for persons born further away from the cut-off date which is confirmed by Fig. 3. Indeed, the coefficients

---

[10] 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 et al. 1995). In other words, 2SLS with weak instruments is generally not appropriate.

Table 4 Simple correlations between instrument and observables (administrative data for *Hessen*)

Sampling window	June/July	June–September	January–December
Added in specification 2: gender (reference = female), entry cohort (refer. = 1997) and county indicators			
Gender dummy variable (male = 1)	0.00	0.00	0.00
School entry in 1998	0.00	0.00	0.01**
School entry in 1999	0.01*	0.01	0.00*
County indicator 1 (Darmstadt)	0.00	0.00	0.00
County indicator 2 (Frankfurt)	0.01	0.00	0.00
County indicator 3 (Offenbach Stadt)	0.00	0.00	0.00
County indicator 4 (Wiesbaden)	0.00	0.00	−0.01**
County indicator 5 (Bergstraße/Odenwald)	0.01	0.01	0.01**
County indicator 6 (Darmstadt-Dieburg)	−0.01	−0.01	0.00
County indicator 7 (Groß-Gerau)	−0.01**	−0.01*	−0.01**
County indicator 8 (Hochtaunus)	0.00	0.00	0.00
County indicator 9 (Main-Kinzig)	0.00	0.00	0.00
County indicator 10 (Offenbach)	0.00	0.00	0.00**
County indicator 11 (Rheingau-Taunus)	0.00	0.00	0.00
County indicator 12 (Offenbach)	0.00	0.00	0.00
County indicator 13 (Wetterau)	0.00	0.00	0.00
County indicator 14 (Gießen)	0.00	0.00	0.00
County indicator 15 (Lahn-Dill)	0.00	0.00	0.00
County indicator 16 (Limburg-Weilburg)	0.01**	0.01	0.00**
County indicator 17 (Marburg-Bied./Vogelsb.)	−0.01	0.00	0.00
County indicator 18 (Kassel Stadt)	0.00	0.00	0.00**
County indicator 19 (Fulda/Hersfeld-Rotenb.)	−0.01	0.00	0.00
County indicator 20 (Kassel/Werra-Meißner)	0.00	0.00	0.00
County indicator 21 (Schwalm-Ed./Waldeck-F.)	0.00	0.00	0.00
Added in specification 3: country of origin			
Country 1 (German speaking countries)	0.00	0.01	0.01**
Country 2 (Turkey)	0.00	−0.01	−0.01**
Country 3 (Italy and Greece)	−0.01**	−0.01**	−0.01**
Country 4 (Former Yugoslavian states)	0.01	0.00	0.00
Country 5 (Remaining "Western" countries)	−0.01	0.00	0.00
Country 6 (Eastern Europe; former Soviet Union)	0.00	0.00	0.00
Country 7 (Remaining Muslim countries)	0.00	0.00	0.00**
Country 8 (Remaining Asia)	0.00	0.00	0.00
Country 9 (Remaining countries)	0.00	0.00	0.00
Number of observations	32,059	64,072	182,676

Note: *Significant at the 10% level. **Significant at the 5% level. The different specifications (specifications 1–3) are explained in Table 2. Specification 1 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

of the full sample amount to 0.49 and are thus somewhat higher than in the smaller sampling windows. In sum, the estimated first-stage coefficients and their *F*-statistics confirm the picture given in Fig. 2 that compliance with the school entry rules is considerable, but not perfect.

One has to keep in mind that 2SLS estimation identifies the causal effect of age of school entry using only the exogenous variation in the age of school entry generated by 'compliers', i.e. those persons who react to variations in

Table 5 First-stage results (PIRLS)

Sampling window/specification	June/July	June–September	January–December
Specification 1	0.40**	0.42**	0.49**
(*F*-statistic)	(86.7)	(147.2)	(433.1)
Specification 2	0.40**	0.42**	0.49**
(*F*-statistic)	(89.1)	(147.8)	(427.1)
Specification 3	0.40**	0.42**	0.49**
(*F*-statistic)	(90.6)	(147.4)	(426.5)
Specification 4	0.40**	0.42**	0.49**
(*F*-statistic)	(94.6)	(150.9)	(440.8)
Specification 5	0.40**	0.42**	0.49**
(*F*-statistic)	(95.1)	(150.6)	(428.6)
Observations	1,123	2,943	6,591

Note: *Significant at the 10% level. **Significant at the 5% level. The different specifications (specifications 1–5) are explained in Table 2
Source: PIRLS 2001. Own calculations

Table 6 First-stage results (administrative data for *Hessen*)

Sampling window/specification	June/July	June–September	January–December
Specification 1	0.41**	0.45**	0.49**
(*F*-statistic)	(2277.1)	(3504.3)	(8196.0)
Specification 2	0.41**	0.45**	0.49**
(*F*-statistic)	(2306.4)	(3524.6)	(8189.0)
Specification 3	0.41**	0.45**	0.49**
(*F*-statistic)	(2325.5)	(3567.7)	(8321.2)
Observations	32,059	64,072	182,676

Note: *Significant at the 10% level. **Significant at the 5% level. The different specifications (specifications 1–3) 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

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.[11] Therefore, the results discussed in the following section may not be representative for the pupil population as a whole. Non-compliers are likely to be either particularly weak pupils who enter school later than prescribed or strong

---

[11] We also tried further instruments based on other cut-off dates (results are reported in the discussion paper version, Puhani and Weber 2005). We assume that persons reacting to the end of June (the Hamburg Accord) as cut-off are more representative for the average pupil, unlike those reacting to alternative rules. For example 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 an instrument based on the end of 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. This hypothesis is confirmed in the discussion paper version (Puhani and Weber 2005).

Table 7 OLS and second-stage results (PIRLS)

Sampling window/ specifications	January–December	June/July	June–September	January–December
Estimate	OLS	2SLS	2SLS	2SLS
Specification 1	−12.80**	28.17**	32.87**	30.74**
(s.e.)	(3.0)	(13.2)	(11.3)	(6.2)
Specification 2	−11.49**	28.18**	33.24**	30.64**
(s.e.)	(3.0)	(13.1)	(11.3)	(6.3)
Specification 3	−8.65**	28.98**	34.29**	27.14**
(s.e.)	(2.7)	(12.6)	(11.0)	(6.2)
Specification 4	−4.57**	26.41**	33.20**	27.37**
(s.e.)	(2.3)	(11.5)	(10.2)	(5.8)
Specification 5	−1.24	25.83**	31.67**	26.77**
(s.e.)	(2.2)	(11.2)	(9.7)	(5.6)
Observations	6,591	1,123	2,943	6,591

Note: *Significant at the 10% level. **Significant at the 5% level. The different specifications (specifications 1–5) are explained in Table 2
Source: PIRLS 2001. Own calculations

performers who enter school earlier than suggested, or they might be 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 the instrument 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

Tables 7 and 8 report the estimated effects of age of school entry on educational outcomes from regressions with different sets of control variables ('specification 1' in the first line indicating no control variables, and the last line indicating the full set of control variables as listed in Table 2). Note, that while in the PIRLS data set the outcome measure is the fourth grade reading test score, in the *Hessen* data the outcome relates to the secondary school track which is coded by years of education necessary for the completion of the degree corresponding to the track (2SLS estimation). Alternatively, we define a binary response variable for attendance of the highest secondary track (*Gymnasium*) in the administrative data for Hessen and estimate a probit instrumental variable model instead of 2SLS.

The first columns of Tables 7 and 8 show the OLS regression coefficients for the full samples (pupils born in January to December). In both data sets, the regression coefficient is negative and significantly different from zero if no control variables are included (specification 1). This means that educational outcomes 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.

Table 8 OLS and second-stage results (administrative data for *Hessen*)

Sampling window/ specification	January–December	June/July		June–September		January–December	
Estimate:	OLS	2SLS	Probit-IV	2SLS	Probit-IV	2SLS	Probit-IV
Specification 1	−0.37**	0.40**	0.12**	0.45**	0.12**	0.45**	0.11**
(s.e.)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)
Specification 2	−0.36**	0.38**	0.12**	0.44**	0.12**	0.44**	0.11**
(s.e.)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)
Specification 3	−0.31**	0.37**	0.12**	0.42**	0.12**	0.41**	0.10**
(s.e.)	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)
Observations	182,676	32,059		64,072		182,676	

Note: 2SLS coefficients indicate the marginal effect of higher age at school entry on years of education according to the current track. An effect of 0.40 years of schooling corresponds to a 12% increase in the probability to attend the higher level school versus the lower level schools. Probit instrument variable estimates report the estimated change in the probability to attend the highest level secondary school (*Gymnasium*) if school entry is at age 7 compared to age 6, where control variables are set to their mean. Estimates were obtained using the statistical software 'Stata'. The standard errors of estimated effects reported in the Probit-IV columns are calculated using the 'delta method'. *Significance at the 10% level, **significance at the 5% level. The different specifications (specifications 1–3) 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

However, as we include more and more control variables into the regressions (specifications 2ff.), the OLS coefficients decrease in absolute value in both 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 outcomes if we apply 2SLS estimation with the instrument discussed in Sect. 4? A glance at Tables 7 and 8 reveals first that instrumental variable estimation switches the sign of the estimated effect from negative to positive in both data sets. Second, the 2SLS estimates are all positive and significantly different from zero. Third, the differences between the point estimates of different sampling windows are smaller than a standard deviation of the narrowest sampling window. Fourth, the size of the estimated effects hardly varies by the choice of control variables (i.e. between 'specification 1' and 'specification 5'/ 'specification 3' in Tables 7 and 8, respectively): 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 following, we will discuss the 2SLS results in detail by data set.

As reasoned in Sect. 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 'specification 1' to 'specification 5') in the PIRLS data set (Table 7). The main finding in Table 7 is that the estimated effect of age at school entry on educational outcomes varies from 25.8 to 29.0 test scores in the narrowest sampling window and is rather robust with estimates ranging from 26.8 to 34.3 when using wider sampling windows.

How can the results be interpreted? A representative estimate based on the narrowest sampling window (discontinuity sample) is an increase in test scores of around 27 points for entering school 1 year older (being about 7 instead of 6 years old). This is about two fifths of the standard deviation of test scores in PIRLS. 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).[12] 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 entering school 1 year older 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 8 presents the effects of age of school entry on track attendance in the middle of secondary school. Results are based on administrative data for the state of *Hessen*. The outcome is measured by the number of school years associated with each track as outlined in Sect. 3.2. Alternatively, we show effects of probit instrumental variable estimations indicating the change in the probability to attend the higher level secondary school (*Gymnasium*) which is due to school entry at 7 instead of 6 years while the control variables are set to their mean.

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 8 all indicate significance. As to the estimated effect of age of school entry on educational outcomes using the Hamburg Accord as instrument, the 2SLS estimation for different sampling windows yields comparable estimates in the ranges of 0.37–0.40 for the narrowest sampling window and 0.41–0.45 for the wider sampling windows. There is only minor variation among specifications with different sets of control variables.[13] Entering school at the age of 7 rather than 6 raises secondary schooling by almost half a year, around 5 months

---

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

[13] 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 9), the coefficients related to the narrowest (widest) sampling window range between 0.43 and 0.46 (0.46 and 0.49). 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 0.36 and 0.39 (0.42 and 0.47).

(assuming pupils will complete the track which they attend in the middle of secondary school, when we observe them). This effect is implied if a deferral of school entry by 1 year increases the probability of attending *Gymnasium* instead of *Realschule* by about 13% points.

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 first estimate linear probability models of *Gymnasium* versus *Realschule/Hauptschule* attendance as well as of *Gymnasium/Realschule* versus *Hauptschule* attendance. Estimates were obtained by 2SLS using the same instrument and control variables as in Table 8. The results show increases of *Gymnasium* versus *Realschule/Hauptschule* attendance by between 11 and 13% points and increases of *Gymnasium/Realschule* versus *Hauptschule* attendance of about 2–3% points. The numbers are very robust and significant across different specifications. 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.

Subsequently, we estimate probit instrumental variable models of the probability to attend *Gymnasium* rather than *Realschule/Hauptschule*. The estimated effect of entering school 1 year older (evaluated at the mean of the control variables) is 12% points using the first two sampling windows and between 10 and 11% points using the full sample. Hence, all our estimation procedures (2SLS with school track coded according to the years needed to complete the track, 2SLS linear probability models and probit instrumental variable models) lead to virtually the same conclusions regarding *Gymnasium versus Realschule/Hauptschule* attendance.

Note, however, that we do not have statistics on the percentage of pupils having attended *Gymnasium* in grade 6 who complete *Gymnasium* by obtaining the *Abitur* degree (equivalent to British A-levels). Back-of-the-envelope calculations based on administrative data for *Hessen* suggests that around 20% of pupils attending *Gymnasium* in grade 6 have left *Gymnasium* in grade 10 in *Hessen*. There might be further attrition in grades 11–13 (when *Gymnasium* ends). However, as pupil panel data currently does not exist to the best of our knowledge, we cannot judge at this stage to what extent our estimates exaggerate the effect of school entry age on *final* schooling achievement. However, separate estimates by school entry cohort suggest that the estimated effect shows no declining trend for older cohorts. Hence, with the data at hand, we have no indication that mobility between school tracks neutralises age of school entry effects in the middle of secondary school.

5.3 Results for subgroups

Having established robust evidence that a relatively older age of school entry improves educational outcomes, we carry out a subgroup analysis in Tables 9 and 10 for the two data sets. For the PIRLS data, Table 9 displays first-stage

Table 9 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 (specification 5). *Significant at the 10% level. **Significant at the 5% level. *F* refers to the *F*-statistics of significance of the instruments in the first-stage regressions

Source: PIRLS 2001. Own calculations

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 for two sampling windows, i.e. the full sample and the narrowest 'discontinuity sampling' window and refer to the specification with all control variables (specification 5).

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 instrument in the respective subgroup are identified by 2SLS. Therefore,

Table 10 Subgroup results for the administrative data from the state of Hessen

	First stage		Second stage	
Male – native (German speaking countries)	Full sample	0.50**	Full sample	0.41**
	(F)	(3885.8)	(s.e.)	(0.04)
(Full sample: 79,400 observations; born June/July: 13,898 observations)	Born June/July	0.41**	Born June/July	0.35**
	(F)	(1025.0)	(s.e.)	(0.08)
Female – native (German speaking countries)	Full sample	0.50**	Full sample	0.45**
	(F)	(3845.2)	(s.e.)	(0.04)
(Full sample: 77,106 observations; born June/July: 13,555 observations)	Born June/July	0.41**	Born June/July	0.39**
	(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: 1,009 observations)	Born June/July	0.42**	Born June/July	0.33
	(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: 1,045 observations)	Born June/July	0.45**	Born June/July	0.32
	(F)	(88.3)	(s.e.)	(0.22)
Male – predominantly Muslim countries (without Turkey)	Full sample	0.36**	Full sample	0.37
	(F)	(25.0)	(s.e.)	(0.41)
(Full sample: 1,539 observations; born June/July: 247 observations)	Born June/July	0.31**	Born June/July	−0.24
	(F)	(6.2)	(s.e.)	(0.72)
Female – predominantly Muslim countries (without Turkey)	Full sample	0.35**	Full sample	0.55
	(F)	(26.3)	(s.e.)	(0.40)
(Full sample: 1,474 observations; born June/July: 248 observations)	Born June/July	0.43**	Born June/July	1.00*
	(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: 271 observations)	Born June/July	0.37**	Born June/July	0.34
	(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: 244 observations)	Born June/July	0.50**	Born June/July	−0.57
	(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: 213 observations)	Born June/July	0.51**	Born June/July	0.01
	(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: 221 observations)	Born June/July	0.38**	Born June/July	1.09
	(F)	(15.7)	(s.e.)	(0.76)

Note: Effects for the full specifications (specification 3). *Significant at the 10% level. **Significant at the 5% level. *F* refers to the *F*-statistics of 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

the estimated 'LATE' do not have to be representative for the subgroups in general (for example, if most immigrant males enter school at the age of 7 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 9 for the PIRLS data, the estimates in Table 10 for *Hessen* are shown both for the full (pupils born January to December) and for the discontinuity samples (pupils born June/July) and refer to 'specification 3' with all control variables. 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/Greece 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 significant point estimate 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 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).

In order to find out whether the insignificance of many subgroup estimates can be explained by smaller sample sizes, we drew random sub-samples of native males, a group for which we found a significant effect. Results based on these random sub-samples indicate that the estimates are not robust and generally insignificant when based on less than 1,500 observations, which unfortunately effects almost all of our subsamples on foreigners (with Turkish citizens born

January to December as the exception). Hence, larger 'samples' (we already observe the population) or a higher degree of compliance would be needed to make statistically safe statements on immigrants (defined as non-citizens in the data for the state of *Hessen*).

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 0.40 standard deviations improvement in the PIRLS test score if the pupil enters at about the age of 7 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 increasing the probability of attending the most academic secondary schooling track (*Gymnasium*) by 12% points. Assuming that the attended track is completed, this amounts to prolonging the average years of schooling by almost half a year (about 5 months).

Compared to Fredriksson and Öckert (2005) and Bedard and Dhuey (2006), who apply an instrumental variable strategy similar to ours to Swedish administrative data and international TIMSS data together with additional data for the US and Canada, 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 (2006) range from 0.2 to 0.5 standard deviations for fourth graders in the countries investigated. Strøm (2004) estimates an effect of 0.2 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.[14] However, these and our 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 US can only be partly explained by the fact that quarter

---

[14] Our estimates based on the PIRLS data (0.40 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 11 countries in Table 3 of Bedard and Dhuey (2006) 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 LATE, 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 11 countries analysed in Table 3 of Bedard and Dhuey (2006) provides a correlation of −0.19 for science and −0.02 for maths test scores 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. We thank Peter Fredriksson for pointing this issue out to us.

of birth in the US, 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.

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.[15] Of the 25 schools, two were operating under a special regime where pupils enter school at the age of 5, but with extra logopedic, German language and nursery teacher support. In these schools, 5–6 olds do not enter grade 1, but 'grade 0', 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% 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% of headmasters or headmistresses (21 out of 22) said they found our results 'plausible'. 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 0' which combines learning with kindergarten elements.

In a further telephone survey of ten schools, we told the headmistresses and headmasters that we had found that early school entry was *good* for children, i.e. we told them the opposite of what we really found in the data.[16] It turned out that eight of ten schools disagreed that an early school entry into the current German school system was sensible. However, four of those eight schools would be in favour of earlier school entry if the school system would be adapted to the needs of younger children (more breaks, smaller classes and an adapted curriculum were named as suggestions).

It is important to note that our identification strategy does not allow for distinguishing between absolute and relative age effects. If our findings were solely driven by relative age effects (peer effects) this study would not provide insights concerning the merits of changing the official school entry age. All we would be able to tell from our findings would be that it is disadvantageous to be one of the youngest children in a given class. However, from our school survey we have some indication that teachers think that absolute age effects matter most: when we asked our contact persons what they believed could be the reasons for our findings 21 out of 22 school representatives made statements along the lines that older pupils are more mature, are more able

---

[15] 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).

[16] We thank Dominique Meurs for suggesting this strategy.

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% 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. Similarly, in the second telephone interview of ten schools, the lack of personal maturity (rather than the relative age) was given as the reason why early school entry was not favoured in the current system. This impression concerning the importance of absolute age effects is consistent with findings by Fredriksson and Öckert (2005).

If we hence believed that our results were driven by absolute age effects, the policy conclusions would depend on whether we observe a pure absolute maturity effect or a maturity-learning interaction effect. If a pure maturity effect drives our results, changing the age of school entry for all pupils makes no difference to the performance gap of older versus younger pupils. However, if we believed that our results were driven by maturity–learning interaction effects (i.e. pupils starting school at the age of 7 learn more in grades 1–4 than pupils starting school when they are younger) the efficiency of early education could be improved by increasing the school entry age. Yet, positive effects of later school entry would have to be weighed against the economic losses of higher labour market entry ages.

All in all, our statistical analysis cannot predict educational implications of changing the official age of school entry regulations. However, we have shown that the age of school entry matters at the individual level under the current school entry regime. In order to separate the underlying causes driving our statistical results (the relative age, the pure maturity and the maturity-learning interaction effect), different data including relative age in the assigned class and on the development of abilities over the life-cycle would need to be collected.

In any case, our results should not be interpreted as evidence against early learning per se. Early learning might generally be promising. Which type of early learning works best will be another interesting research agenda, once state governments decide to collect and make available appropriate data in this respect.

References

Angrist JD (2004) American education research changes track. Oxf Rev Econ Policy 20:198–212

Angrist JD, Krueger AB (1992) The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. J Am Stat Assoc 87:328–335

Baumert J, Trautwein U, Artelt C (2003) Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In: Deutsches PISA-Konsortium (ed) PISA 2000 . Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Verlag Leske + Budrich, Opladen, pp 261–331

Bedard K, Dhuey E (2006) The persistence of early childhood maturity: international evidence of long-run age effects. Working paper, Department of Economics University of California, Santa Barbara. Q J Econ (forthcoming)

Bertram T, Pascal C (2002) Early years education: an international perspective. Qualifications and Curriculum Authority, London

Bos W, Lankes EM, Prenzel M, Schwippert K, Walther G, Valtin R (2003) Erste Ergebnisse aus IGLU. Waxmann Verlag, Münster

Bound J, Jaeger DA, Baker RM (1995) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. J Am Stat Assoc 90:443–450

Cunha F, Heckman JJ, Lochner L, Masterov DV (2006) Interpreting the evidence on life cycle skill formation. In: Hanushek E, Welch F (eds) Handbook of the economics of education. Elsevier, North-Holland, New York (forthcoming)

Currie J (2001) Early childhood education programs. J Econ Perspect 15:213–238

Del Bono E, Galindo-Rueda F (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. Oxf Econ Pap 56:209–230

Fertig M, Kluve J (2005) The effect of age at school entry on educational attainment in Germany. IZA discussion paper no. 1507

Fredriksson P, Öckert B (2005) Is early learning really more productive? The effect of school starting age on school and labour market performance. IZA discussion paper no. 1659

Gonzalez EJ, Kennedy AM (2003) PIRLS 2001 user guide for the international database. International Study Center, Lynch School of Education, Boston College, Boston

Graue ME, DiPerna J (2000) Redshirting and early retention: who gets the "Gift of Time" and what are its outcomes? Am Educ Res J 37:509–534

Hahn J, Todd P, Van der Klaauw W (2001) Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69:201–209

Hanushek EA, Wößmann L (2006) Does educational tracking affect performance and inequality?: differences-in-differences evidence across countries. Econ J 116:63–76

Hutchison D, Sharp C (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

Imbens GW, Angrist JD (1994) Identification and estimation of local average treatment effects. Econometrica 62:467–475

Jimerson S, Carlson E, Rotert M, Egeland B, Sroufe LA (1997) A prospective, longitudinal study of the correlates and consequences of early grade retention. J Sch Psychol 35:3–25

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

Kinard EM, Reinherz H (1986) Birthdate effects on school performance and adjustment: a longitudinal study. J Educ Res 79:366–372

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

May DC, Kundert DK, Brent D (1995) Does delayed school entry reduce later grade retentions and use of special education services? Remedial Spec Educ 16:288–294

Mayer S, Knutson D (1999) Does the timing of school affect how much children learn? In: Mayer S, Peterson P (eds) Earning and learning: how schools matter. Brookings Institution Press, Washington, pp 79–102

Morrison FJ, Griffith EM, Alberts DM (1997) Nature-nurture in the classroom: entrance age, school readiness, and learning in children. Dev Psychol 33:254–262

Proctor TB, Black KN, Feldhusen JF (1986) Early admission of selected children to elementary school: a review of the research literature. J Educ Res 80:70–76

Puhani PA, Weber AM (2005) Does the early bird catch the worm? Instrumental variable estimates of educational effects of age of school enty in Germany. IZA discussion paper no. 1827

Sharp C (2002) School starting age: European policy and recent research. In: NFER conference paper, presented at the LGA seminar 'When Should Our Children Start School?'. LGA Conference Centre, London

Staiger D, Stock JH (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. Soc Policy Rep 16:3–16

Stipek D, Byler P (2001) Academic achievement and social behaviors associated with age of entry into kindergarten. Appl Dev Psychol 22:175–189

Stock JH, Wright JH, Yogo M (2002) Survey of weak instruments and weak identification in generalized method of moments. J Bus Econ Stat 4:518–529

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

Willer CW, Dyment DA, Sadovnick AD, Rothwell PM, Murray TJ, Ebers GC (2005) Timing of birth and risk of multiple sclerosis: population based study. Br Med J 330:120–123

Zill N, Loomis LS, West J (1997) National household education survey. The elementary school performance and adjustment of children who enter Kindergarten late or repeat Kindergarten: findings from national surveys. In: NCES statistical analysis report 98–097. US Department of Education, National Center for Education Statistics, Washington