The Long-term Effects of School Quality on Labor Market Outcomes and Educational Attainment

Christian Dustmann, Patrick A. Puhani and Uta Schönberg*

February 21, 2012

Abstract

We study the long-term causal effects of attending a "better" school - defined as one with more advanced peers, more highly paid teachers, and a more academic curriculum - on the highest degree completed, wages, occupational choice, and unemployment. We base our analysis on a regression discontinuity design, generated by a school entry age rule, that assigns students to different types of schools based on their date of birth. We find that, even though our case involves larger inter-school differences in peer quality and teaching curricula than in most previous studies, the long-term effect of school quality is very small and not significantly different from zero. This surprising finding is partly explainable by the substantial amount of student up- and downgrading between schools of varying quality at the end of middle school (age 15/16) and at the end of high school (age 18/19). This suggests that giving people a "second chance" during their education can make up for several years of schooling with a less challenging peer group and a less challenging teaching curriculum.

JEL: JEL No. I21, J10

Keywords: School quality, peer effects, regression discontinuity design

*Dustmann: Department of Economics, University College London, c.dustmann@ucl.ac.uk. Puhani: School of Economics and Management, Leibniz Universität Hannover, puhani@aoek.uni-hannover.de. Schönberg: Department of Economics, University College London and Institute for Employment Research, u.schoenberg@ucl.ac.uk. We are grateful to Josh Angrist, David Card, Michael Lechner, Dominique Meurs, Steve Pischke, Jesse Rothstein, and Guido Schwerdt for comments. Part of this research was supported by the German Research Foundation (DFG) within the project Labour Market Effects of Social Policy.

1 Introduction

Research consistently shows that parents are willing to pay substantial sums of money for their children to be educated in better quality schools (see Black and Machin 2010 for a review). Yet, so far little is known about the importance of the quality of education early in life on children's performance *in the long-term*, and the precise mechanisms by which early exposure to better quality schools may affect long-term outcomes. In this paper, we investigate the long-term effects of attending a "better" middle school (grade 5 to grade 9/10), with a focus on how the quality of the middle school affects long-term outcomes through two channels: *indirectly*, through the type and quality of education received after middle school ("secondary education"), and *directly*, holding the type and quality of secondary education constant. The indirect effect depends on the opportunities and costs of acquiring a high quality secondary education even when the individual has attended a middle school of low quality, an upgrading we label a "second chance".

We examine the direct and indirect effects of school quality within a context in which differences between middle schools are particularly striking, the German school system, which allocates fourth-graders (around age 10) to three different middle school types, here designated "high", "medium", and "low". Once this decision is made, children are locked into the chosen school type throughout middle school (and possibly beyond) for at least five years. School types differ widely with respect to peer quality and teacher quality.[1] Moreover, teaching at a more advanced school type is more abstract, more intensive, and proceeds at a faster rate than teaching at a less advanced school type.

We find that attending a middle school of higher quality has little impact on the type of secondary education received. Not only is this finding remarkable given the striking differences between middle school types, but it implies that students are able to revise their

[1] For example, the test scores of students attending the high (medium) school type are more than one (0.7) standard deviations higher than those of students attending the medium (low) school type, a much larger difference than the 0.2 standard deviations reported in Cullen, Jacob, and Levitt (2005).

initial school type selection at a later stage in their educational career. In fact, we show that there is a substantial amount of up- and downgrading between school types at the end of middle school, at age 15/16, and yet again at the end of high school, at age 18/19.

Since the type of middle school has little impact on the type of secondary education, it can affect long-term labor market outcomes only through the direct channel. Yet we find that this direct effect is also close to zero, implying that, conditional on the type of secondary education received, exposure to better peers and teachers and a more advanced teaching curriculum in middle school has little impact on long-term outcomes like wages and unemployment. Our general conclusion, therefore, is that even longer exposure to a less challenging school environment may not have any detrimental long-term effects provided that the education system offers enough opportunities for students to revise the educational choices made earlier in life. This provision of "second chances", although it has received little attention in the literature, is in our view fundamental for the assessment of any educational system, regardless of whether school choice is determined primarily by ability (as in Germany) or residential location (as in the U.S. or UK).

To identify the causal effects, we exploit school entry age cut-off rules. A large body of literature has shown that age of school entry is an important determinant of early student achievement because of the additional knowledge acquired prior to school entry (e.g., Bedard and Dhuey 2006; Frederickson and Oeckert 2006; Puhani and Weber 2007; Cascio and Schanzenbach 2009; Elder and Lubotski 2009; Black, Devereux, and Salvanes 2011). Since the decision on which middle school type to attend is made very early in Germany - at a stage when the age of school entry effect is unlikely to have dissipated - the date of birth may, through its link to school entry age, affect the type of middle school that students attend. Indeed, we show that children who are born shortly after the school entry cut-off date (and are thus older at school entry) are considerably more likely to attend a more academic middle school than children born shortly before this date (who are thus younger at school entry). This presents the first stage of our analysis. Subsequently, we investigate the reduced-form effect of the school entry cut-off on long-term outcomes, including education completed, wages, and unemployment, and assess the impact of middle school type attended on long-term outcomes by dividing the reduced-form estimates by the first-stage estimates. It is important to emphasize that our results refer to a particular group of students who were shifted to a more academic school because of birth date. These students are therefore likely to fall within a tight ability range and be close to indifferent as regards school type. This group is particularly interesting because it consists of children whose teachers and parents may be unsure of which school type they should attend after elementary school.

Our paper is closely related to the literature on school quality. Some earlier papers for the U.S. focus on the effects on educational attainment and subsequent wages of attending a Roman Catholic versus a public high school. These papers instrument school choice by either religion (e.g., Evans and Schwab 1995) or public school access (e.g., Neal 1997), both of which may have a direct impact on a student's school performance.[2] More recent papers use randomization designs, but these do not typically study the long-term effects of school quality on labor market outcomes (as examined here). Moreover, they reach conflicting conclusions. For instance, whereas Gould, Lavy, and Paserman (2004), Jackson (2010), Pop-Echeles and Urquiola (2010), and Maurin and McNally (2007) all find that attending a better school improves children's academic achievement, Cullen et al. (2005, 2006), Clark (2010) and Abdulkadiroglu, Angrist, and Pathak (2011) find no evidence that school quality improves standardized test scores.[3] A recent experimental study which, similar to us, focuses on long-term adult outcomes, is Chatty, Friedman, Hilger, Saez, Whitmore Schanzenbach

[2]To deal with the non-random selection of students into Catholic schools, Altonji, Elder, and Taber (2005) propose new estimation methods in which the amount of selection on the observed explanatory variables in a model provides a guide to the amount of selection on the unobservables.

[3]Related research (e.g., Hoxby and Murarka 2009; Abdulkadiroglu et al. 2011; Angrist et al. 2010) uses admission lotteries to estimate the impact of charter school attendance on student achievement. Charter schools differ considerably from a "high quality" school in our study in that they often target minority and low income students.

and Yagan (2011) who find that children who were randomly assigned to higher quality classrooms in kindergarten have higher earnings and college attendance rates.[4]

Our paper also adds to the studies on peer effects, many of which use idiosyncratic variation across student cohorts within schools, or across classes within a cohort and school, to isolate the role of peers (e.g., Bifulco, Fletcher, and Ross 2010; Black et al. 2010; Gould, Lavy, and Paserman 2009; Hoxby 2000; Lavy and Schlosser 2011, and Ammermüller and Pischke 2009).[5] We contribute to this literature by exploiting a very different source of exogenous variation in exposure to peers, one that results in larger differences in peer quality. In addition, whereas most earlier papers investigate short-term test score outcomes, we focus on long-term educational attainment and labor market effects.[6]

Finally, although our findings say little on whether a tracked school system is preferable to a comprehensive school system, our paper does contribute to the literature on tracking by addressing a key criticism against a tracked school system - the perpetuation of initial disadvantages driven by differences in both innate ability and social background (Figlio and Page 2002; Hanushek and Wössmann 2006; Brunello and Checchi 2007, Malamud and Pop-Echeles 2010 and Waldinger 2007).[7] While our findings confirm that such arbitrary factors as birth date, which are unrelated to innate student ability, may strongly impact the initial track choice, they also reveal that short-run effects may differ considerably from long-term effects in tracking systems that allow up- or downgrading (i.e., a second chance) once the

[4] Chatty, Friedman, and Rockoff (2011) focus on teacher quality and find that better teachers, measured in terms of their value added, raise children's college attendance rates and income.

[5] Similarly, our paper relates to the literature on teacher quality, as this is one dimension along which our school types differ. A number of papers use a value-added type approach to isolate effect teacher quality has on test scores (for instance Hanushek, Kain, and Rivkin 2005; Kinsler 2011). See Rothstein (2010) for a criticism of this approach.

[6] One exception is Black et al. (2010), who examine the role of peer composition in ninth grade on longer term outcomes like IQ scores at age 18, teenage childbearing, post-compulsory schooling, and earnings. They find that whereas the education of peers' mothers has no positive effect on these outcomes, a larger share of girls in the class improves them for both men and women, although at a different rate.

[7] In a recent paper, Duflo et al. (2011) provide experimental evidence on the overall, albeit short-run, effects of tracking, showing that students in tracked schools, including those assigned to the lower track, achieve higher test scores even after the program ends than students in non-tracked schools. Manning and Pischke (2006) evaluate a switch from a tracking to a comprehensive school system in the UK.

student's true potential is revealed. Our work therefore emphasizes that educational systems in different countries should be judged not only on when and how they segregate students but also on the extent to which they allow students to remedy initial choices that may have been based on incomplete or faulty information about student learning potential.

2 Background

2.1 General Overview of the German Education System

Figure 1 provides an overview of the German education system, which allocates students into three types of middle schools at an early point (by international comparison) in their educational career - the end of fourth grade. Education in the highest middle school type, the *Gymnasium* (comparable to the traditional British grammar school) lasts for nine years (grades 5 to 13) and prepares students for tertiary studies at such academic institutions as four-year colleges or five-year universities.

Education at lower and intermediate school types (*Haupt*- and *Realschule*), on the other hand, lasts five years (grades 5 to 9) or six years (grades 5 to 10), respectively, and is less academic, traditionally preparing students for an apprenticeship in blue-collar (e.g., crafts) and white-collar occupations (e.g., office clerk, but also medical assistant). We label these three school types high (H), medium (M), and low (L). For the men and women born between 1961 and 1976, the cohorts focused on in the empirical analysis, the shares of who attends each are roughly of equal size.

Although there is no strict rule (such as an entry exam) to determine which type of middle school children can attend, elementary school teachers do make recommendations, and in 10 out of 16 states, parents have the final word on this choice. In the remaining 6 states, if parents want to deviate from the teaching committee's recommendation, students must either have earned the required marks or pass a special test.[8] In principle, students may

[8] For more information, see http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_10_18-ueberg.pdf.

switch between school types in any grades throughout middle school; however, in practice, very few students (less than 2%) do so.[9] This point is important because it means that once students are allocated to a particular school, their exposure to different learning technologies, peers, and teachers is sustained for at least five years. Once students complete their course at an L or M school type, however (i.e., after ninth or tenth grade), students may correct their initial choice by continuing their education at a medium or high school type, respectively. In particular, at the end of ninth grade, low school type students (hereafter, L students) may switch to a medium school type (M), although some L schools also offer students the opportunity to stay on for another year to earn the same school-leaving qualification which they would have received from an M school. In addition, after tenth grade, students who graduated from an M school may either upgrade to a traditional H school or attend one targeted at former M school students. These latter, because they often provide a special focus in one discipline (e.g., agricultural, business, health, or social studies) in addition to general education are here labeled "specialized high school types" (see Figure 1). Like graduation from a general H school, graduation from a specialized H school grants access to college or university, but possibly restricts the field of study.

2.2 Differences between School Types

The first important difference between the three middle school types is that students attending an H school or M school are surrounded by academically stronger peers than students attending an L school. In fact, according to PISA data for 2003 and 2006, the type of middle school attended is strongly associated with test scores in the ninth grade: average reading and mathematics test scores at H schools are about one standard deviation higher than at M schools and about 1.7 standard deviations higher than at L schools (see Table 1). These differences in peer quality across school types are far greater than those in the quasi-experiments by Pop-Eleches and Urquiola (2010) and Cullen et al. (2007) (around 0.20 of a

[9] Own calculations based on the School Census for Bavaria and Hesse.

standard deviation in both cases). Likewise, H students are exposed to peers whose family backgrounds are far more academic than those of potential classmates at M or L schools. For instance, the parents of classmates attending an H school have parents with almost four years more education than the parents of classmates at an L school. Likewise, whereas 40 percent of the households of L students have fewer than 25 books in the home, this number decreases to 23 percent for M students and to just over 5 percent for H students. H students also come from higher income households, with only 39 percent living in households with a below-median income, a number that increases to 65 percent for M students and 76 percent for L students. As Table 1 also shows, the share of girls in L schools is about 14 percentage points lower than in H schools, which may boost the performance of boys and girls in both school (e.g., Schanzenbach 2009; Lavy and Schlosser 2011) and the labor market (Black et al. 2010). It should also be noted that school type alone can explain more than 70 percent of the overall variation in test scores in ninth grade, suggesting that schools of the same type are fairly homogeneous.

A second important difference between the three middle school types is that teachers in H schools are likely to be of higher quality than teachers in M or L schools. More specifically, as Table 1 shows, the minimum formal education for H teachers is one year higher than that for L and M teachers, and their salaries are approximately 10 percent higher.

Third, the three school types also differ with respect to teaching intensity and learning goals.[10] We summarize these details at the bottom of Table 1: in ninth grade (i.e., when attendance at L schools ends), the number of hours taught per week is 36 at H but only 32 at M or L schools. Moreover, although the number of weekly hours in core subjects like mathematics, German, and English as a foreign language is similar across school types, H schools teach more hours in the second and third foreign language and natural sciences, and

[10] The information in this paragraph was gathered from the curricula published on the webpages of the Ministries for Culture and Education of West German states. In some cases, we also contacted the ministries by telephone.

fewer hours in social sciences, physical education, and vocational subjects (e.g., the "World of Work") than L or M schools. In the core subjects of mathematics and German, school types differ with respect to the topics taught as well as in teaching intensity. For example, whereas an L school puts special emphasis in the ninth grade on such applications as writing CVs, filling out forms, and preparing for job interviews, an H school pays special attention to detailed explanations, analysis, and interpretation of various types of texts, including historical documents, and stresses creative writing. Likewise, whereas ninth-grade mathematics in an M school covers, among other things, real numbers and powers and equations with two unknowns, these topics are taught in eighth grade in an H school. Similarly, although both M and H schools introduce functions, H schools do so more intensely and cover advanced functions (e.g., exponential and broken power functions) not taught in M schools. L schools, in contrast, sometimes cover no functions at all, focusing instead on equations with rational numbers and descriptive statistics.

2.3 Mechanisms: Direct versus Indirect Effects

Given these large quality differences between middle school types, how would we expect middle school type (T^M) to affect long-term outcomes like wages (Y)? There are two channels by which the quality of the middle school can affect long-term outcomes, one direct and one indirect:[11]

$$\frac{dY}{dT^M} = \underbrace{\frac{\partial Y}{\partial T^S}\frac{\partial T^S}{\partial T^M}}_{\text{indirect}} + \underbrace{\frac{\partial Y}{\partial T^M}}_{\text{direct}} . \tag{1}$$

First, the type of middle school may have an *indirect* effect on long-term outcomes like wages by affecting the type of secondary education T^S pursued after middle school ($\frac{\partial T^S}{\partial T^M}$), which subsequently affects wages ($\frac{\partial Y}{\partial T^S}$). Middle school type may also have a *direct* effect on long-term outcomes even when secondary education type is held constant ($\frac{\partial Y}{\partial T^M}$) - which reflects the long-term effect of exposure in middle school to "better" peers, more qualified teachers,

[11] Note that for simplicity we have assumed here that middle school type is continuous.

and a more challenging curriculum. It should be noted that when middle school type has no impact on secondary education type, the indirect effect must be zero, so the impact of middle school type on long-term outcomes is equal to the direct effect only.

More generally, the magnitude of the indirect effect on long-term outcomes depends on how strongly middle school type affects secondary education type ($\frac{\partial T^S}{\partial T^M}$), which is dependent on the opportunities for switching school type. In our case, such opportunity means for instance the possibility of pursuing a college-track secondary education even when the middle school chosen at the end of elementary school is designed to prepare students for apprenticeship only. As emphasized (see Figure 1 and our discussion in Section 2.1), in principle, the German education system offers students the opportunity to switch school type at the end of middle school. However, whether students make use of this "second chance" is an open question, one whose answer is highly dependent on the costs of such switching. Therefore, we begin the empirical analysis by investigating the impact of middle school type on secondary education type. We then turn to the impact of middle school type on long-term labor market outcomes like wages and unemployment.

2.4 Outcome Differences between School Types

Does middle school type predict long-run education and labor market outcomes? In Table 2, we use data from the German Socio-Economic Panel to document a strong association between school type attended at age 14 and school type completed by age 21 (Panel A). As the table shows, 80.8 percent of individuals who attended an H school at age 14 later graduated from either a general or specialized H school, compared to only 30.8 percent and 6.6 percent of students who attended an M or L school, respectively, at that age. In addition, individuals who attended an M school at age 14 are 64 percentage points (86.6-22.6) more likely to complete at least an M school than individuals who attended an L school at that age. Hence, middle school type at age 14 is a strong predictor for school type completion.

In Table 2, Panel B, we report the correlations between labor market outcomes and school type *completed*.[12] We find that the wages of men and women who graduated from a (general or specialized) H school are roughly 19 percent higher than the wages of those who graduated from an M school, who in turn earn 10 percent higher wages than those who graduated from an L school. The incidence of unemployment also decreases with school type graduated from. Educational level, on the other hand, increases with school type completed: whereas almost a quarter of the students who completed an L school enter the labor market without any post-secondary education, only 8 percent of students who completed an M or H school do so. In addition, as expected, H graduates are much more likely to finish college or university than L or M graduates, 54 percent compared to only 3 percent and 1.6 percent, respectively. Combining these estimates with those in Panel A, we find that attending an H versus an M school at age 14 is associated with an increase in wages of 9.7 percent and an increase in the probability of completing college or university of about 27.5 percentage points. Likewise, attending an M versus an L school at age 14 is associated with an increase in wages of about 11.5 percent and a drop in the probability of entering the labor market without post-secondary education of about 9.6 percentage points.[13] These associations, however, are not causal in that they do not take into account the sorting of students into school types based on ability. Hence, in the next section, we describe a research strategy that allows us to uncover the causal effects of attending a more academic middle school for students of the same ability. We also develop a model of school type choice that facilitates precise definition of the channels through which school type affects outcomes.

[12] In the German Socio-Economic Panel, only a few individuals can be followed from age 14 (middle school) to age 30 and older (post-secondary education and entry into the labor market). We therefore present correlations between labor market outcomes and school type completed, which do not require that individuals be followed from age 14 throughout their education and labor market careers.

[13] These numbers are computed as follows: (0.808-0.308)·0.187+(0.902-0.866)·0.109; (0.808-0.308)·0.541; (0.308-0.066)·0.187+(0.866-0.226)·0.109; and (0.866-0.226)·(0.230-0.080).

3 Research Design and Interpretation

3.1 Randomization into School Types

The Natural Experiment To identify a causal effect, we require an experiment that effectively randomizes students of the same ability into different types of middle schools (see Figure 2). We begin by recognizing that, because of the enrollment cut-off rule, children whose birthdays fall before July 1 typically start school a year earlier than children whose birthdays are on or after July 1. Hence, although not every child complies with this law, children born in July are on average considerably older at school entry than children born in June. As a result, although birth month is likely to be random, July-born children (who enter school one year later) may have the temporary advantage over June-born children of greater knowledge accumulation before middle school (see e.g. Elder and Lubotsky 2009). Thus, the enrollment cut-off rule may allocate children born in June and July to different school types after elementary school, which, given the limited possibilities of switching during grades 5 through 9, locks them into a particular type throughout middle school. Hence, the first stage of our analysis addresses the difference in school type attended through grades 5 and 9, T, between children born in month Z, where Z is June or July:

$$\pi = E[T_i|Z_i = \text{July}] - E[T_i|Z_i = \text{June}] \quad (2)$$

We then investigate the impact of birth month on long-term labor market outcomes like wages and unemployment. The reduced-form of our analysis is the difference in these long-term outcomes (denoted by Y) between students born in July and those born in June, given by:

$$\eta = E[Y_i|Z_i = \text{July}] - E[Y_i|Z_i = \text{June}] \quad (3)$$

We finally obtain the impact of middle school type on long-term outcomes by dividing the reduced-form effect by the first-stage estimate, $\tau = \frac{\eta}{\pi}$. Under some assumptions, discussed below, τ identifies a weighted average of two local average treatment effects:[14]

$$\begin{aligned}\tau &= \frac{aE[Y_i^M - Y_i^L | M_i \text{ if } Z_i = \text{July}, L_i \text{ if } Z_i = \text{June}]}{a+b} \\ &+ \frac{bE[Y_i^H - Y_i^M | H_i \text{ if } Z_i = \text{July}, M_i \text{ if } Z_i = \text{June}]}{a+b} \end{aligned} \tag{4}$$

The first local average treatment effect, $E[Y_i^M - Y_i^L | M_i \text{ if } Z_i = \text{July}, L_i \text{ if } Z_i = \text{June}]$, is the impact of attending an M rather than an L middle school on the long-term outcomes for those who would have attended an M middle school had they been born in July but an L middle school had they been born in June. The second local average treatment effect, $E[Y_i^H - Y_i^M | H_i \text{ if } Z_i = \text{July}, M_i \text{ if } Z_i = \text{June}]$, is the impact of attending an H rather than an M middle school on long-term outcomes for those individuals who were shifted from an M to an H middle school because of birth month. The weights a and b represent the differences in the probability of attending an L (H) middle school between individuals born in July and those born in June; that is, $a = \Pr(L|\text{June}) - \Pr(L|\text{July})$ and $b = \Pr(H|\text{July}) - \Pr(H|\text{June})$.

Assumptions Using the ratio of the reduced-form effect to the first-stage effect to identify the causal impact of attending a more academic middle school on long-term outcomes requires two assumptions: independence and exclusion. The independence assumption stipulates that whether a child is born in June or July be random.[15] Note that since we are comparing individuals born within the same season, this independence assumption is considerably weaker here than in studies that use birth quarter as an instrument.[16] To assess

[14] Here, we make the additional assumption that no individuals are shifted from an L to an H school because of birth month.

[15] In the regression discontinuity estimates that exploit the exact date of birth (see below), the independence assumption stipulates that whether a child is born just before or just after the school entry cut-off date (i.e., on June 30 or July 1) be random.

[16] See Bound et al. (1995) and Bound and Jaeger (2000) for a criticism of using quarter of birth as an instrument. See Buckles and Hungerman (2008) for recent evidence of large seasonal effects by which children born in spring outperform children born in winter.

whether the independence assumption is being violated, in Table A2 we compare the parental characteristics of children born in July versus those born in June using German Microcensus data for 2005, the only year for which information on birth month is available. We find no significant differences between children born in July and those born June in terms of parental education and their age at the children's birth.

The exclusion assumption requires that the birth date, and hence the age of school entry, affect long-term outcomes only through middle school type attended.[17] In particular, we postulate that being older at school entry conveys initial advantages that allow children to attend a more academic middle school but that these relative age effects dissipate as children grow older. Although the recent literature agrees that the benefits of delayed school entry decline with the child's age (e.g., Cascio and Schanzenbach 2009, for the U.S.; Crawford et al. 2007, for the U.K.; Muehlenweg and Puhani 2010, for Germany), as yet, the literature has reached no consensus on the long-term impact of age at school entry.[18] If the benefits from delayed school entry persist into adulthood for reasons other than school type attendance, we must be overstating the impact of middle school type on education completed and labor market outcomes.

Empirical Implementation To implement our empirical strategy, we estimate the first-stage and reduced-form effects π and η by replacing the population means in equations (2) and (3) with their sample means, while controlling for birth year and gender effects.[19] We report heteroscedasticity-consistent (robust) standard errors, clustered at the person level, for both the first-stage and reduced-form estimates. Because the two effects are estimated

[17]In contrast to the U.S., where students are typically allowed to drop out of school on their 16th birthday (see Angrist and Krueger 1991), in Germany, children must complete at least nine years of compulsory full-time schooling. Consequently, in Germany, there is no mechanical link between age of school entry and education completed at the time of school leaving.

[18]For instance, Bedard and Dhuey (2006) report age-at-school-entry effects that last well into early adulthood, and Grenet (2010) and Solli (2011) find long-lasting effects for France and Norway, respectively. Cascio and Schanzenbach (2007), in contrast, show that for Whites, the impact of age at school entry on test scores disappears by age 16.

[19]Uur results remain almost unchanged if we exclude these control variables from the regressions.

using two different samples, we refer to $\hat{\tau} = \frac{\hat{\eta}}{\hat{\pi}}$ as the two-sample two-stage least squares estimate (TS-2SLS) of middle school type attended on long-term outcomes. We compute the standard error of the TS-2SLS estimates using the delta method.[20]

For the reduced-form effects, we also report regression discontinuity estimates that exploit the student's exact birth date.[21] Specifically, we estimate regressions of the following type:

$$Y_i = \alpha_0 + h(\mathrm{Day}_i) + \alpha_1 \mathrm{Post}_i + x_i' \alpha_2 + u_i. \tag{5}$$

Here, day_i is the student's birth date (normalized to be 0 on the school entry cut-off date, July 1), $h(.)$ is a polynomial function of birthday, Post_i is an indicator variable equal to 1 if the student was born on or after July 1, and x_i is a control variable vector that includes birth year and gender effects. The parameter of interest is α_1, the impact of being born after the school entry cut-off date (i.e., on or after July 1) on long-term outcomes. We first estimate equation (5) on a sample of all students and then restrict the sample to students born within three months of the school entry cut-off date (i.e., to students born between April and September). We approximate the function $h(\mathrm{day}_i)$ as a polynomial function of various orders. As suggested by Lee and Card (2007), we cluster standard errors at the birth date level.[22]

In Section 2.3, we emphasize that the type of middle school attended may affect long-term labor market outcomes directly, when secondary education type is held constant, or indirectly, through its impact on the type of secondary education; see equation (1). In the next section, we set up a stylized model of school type choice that clarifies how the

[20] We compute the variance of $\widehat{\tau}$ as $Var\left[\widehat{\tau}\right] = Var\left[\frac{\widehat{\eta}}{\widehat{\pi}}\right] = \frac{\widehat{\pi}^2 Var(\widehat{\eta}) + \widehat{\eta}^2 Var(\widehat{\pi})}{\widehat{\pi}^4}$, whose square root is the estimate of the standard error (when variances in the formula must be replaced by their estimates). See Inoue and Solon (2010) for a discussion, and an alternative estimator.

[21] We cannot do this for the first-stage effect because our data set provides information only on birth month or whether the pupil was born earlier or later during the year; see Section 4.1 for details.

[22] Our baseline estimate comparing individuals born in June with those born in July may also be seen as a regression discontinuity estimate in which the sample is restricted to students born within two months of the school entry cut-off date and the birthday effect ($h(.)$ in equation (5)) is assumed to be constant.

key differences between school types (peer, teacher, and curriculum quality) generate these indirect and direct effects. The model also sheds light on which students were shifted to a more academic school because of birth month. For brevity, we delegate all technical details to Appendix A and focus our discussion on intuitive interpretations. This section is not necessary to understand the empirical analysis, and can be skipped by the reader who is only interested in the empirical content of our study.

3.2 Interpretation: A Model of School Type Choice

Set-up Our model assumes three periods: periods 1 and 2, which are schooling periods, and period 3, which is the working period. The beginning of the first period corresponds to the end of elementary school (grade 4) when parents decide on which type of middle school their child should attend. The end of this period corresponds to the end of middle school (grade 9 or 10) when children have the opportunity to switch school type. The second period corresponds to the period of secondary education, and all children enter the labor market at the beginning of period 3. This setup, it should be noted, assumes that the school types are of equal length, an assumption motivated by our focus on the impact of school quality on long-term outcomes. For simplicity, we distinguish only two school types, low (L) and high (H).

In this model, children differ in both their "ability", denoted by a, and birth month, which here is limited to children born in June or July. In line with the independence and exclusion assumptions discussed in Section 3.1, at the beginning of period 1, when the initial school choice is made, children born in July outperform children born in June by Δ, but by the beginning of period 2, when the initial decision can be revised, this advantage has disappeared.

The children's ability is initially uncertain. At the beginning of period 1, parents receive a noisy signal about their child's ability, one that we can think of as school grades, which they use to update their beliefs about this ability. By the end of the first period, the child's ability is fully revealed, and parents and children can revise their school type choice accordingly. Switching from an L to an H school, however, can be academically costly. We model this cost by assuming that in H schools, the productivity of students who attended an L school in period 1 is lower, by c, than the productivity of students who attended an H school in that period. This differential reflects the risk that the less intensive and less abstract teaching method at an L school (see Section 2.2) may make it difficult for students to keep up with the more advanced learning material at the H school. Conversely, we assume that moving down from an H to an L school is costless.

We summarize the productivity or performance of June- versus July-born students in each school type and each period in Table A1. To model peer effects, we use a standard linear-in-means peer model in which a higher average peer ability increases the productivity of all students in the same way. We assume that parents take the expectations of the average ability of students in each school type as a given, thereby ignoring the possibility that they can manipulate other parents' expectations through their own school choices. Teacher effects are modeled by assuming that better teachers improve the productivity of all students by the same amount. We allow the contemporaneous effect on productivity of peers and teachers in period 1 or 2 to differ from their effect on productivity in future periods. To model the differences in teaching technology between the school types, we assume that learning technology is more sensitive to ability in an H than in an L school. Specifically, the learning technology in the two school types is linear in the student's ability a and thus is given by $\alpha^j + \beta^j a$, with $j = L, H$ and $\alpha^L > \alpha^H$ and $\beta^L < \beta^H$. These assumptions reflect the fact that the teaching technology in an L school is more adapted to children who are drawn from the lower part of the ability distribution.

To summarize, contrast the productivity in period 3 of an individual who attended an L school in period 1 and an H school in period 2 with that of an individual of the same ability who attended an H school in both periods. The difference in the productivity between these two individuals reflects the *direct* effect (i.e., conditional on school type attended in period 2) of the school type attended in period 1 on the long-term outcomes in period 3. According to our set-up, this direct effect operates through two channels: first, the individual who attended an H school in both periods has been exposed to better peers and teachers in period 1, which may boost labor market performance even when both individuals received the same type of education in period 2. Second, the individual who attended the H school in both periods has not suffered the academic cost of switching from an L to an H school.

Equilibrium At the beginning of periods 1 and 2, parents (or in period 2, students) choose the type of school that will yield the highest lifetime utility, which is the sum of the student's utility in each period. For simplicity, we assume that in each period, student utility is equal to student productivity (as described in Table A1)[23] and ignore discounting.

Period 2 Decision: First, we consider the parent's decision problem at the beginning of period 2, when the student's ability is fully known. Because in period 2, there is no difference between children born in June and those born in July, the decision problem is the same for both groups. Supposing that the student attended an L school in period 1, in Appendix A.2, we show that there exists an ability threshold a_L^* such that all students whose ability turns out to be less than a_L^* continue at an L school while all students whose ability turns out to be above a_L^* upgrade to an H school. A similar threshold a_H^* can be derived for students who attend an H school in period 1. Because of the switching cost, $a_L^* > a_H^*$.

Period 1 Decision: Next, we consider the parents' decision problem in period 1, when the student's ability is uncertain. Here, parents take into account the optimal switching behavior in period 2 and choose the L over the H school if the child's expected utility of attending an L school exceeds that of attending an H school. In Appendix A.2, we again show that a threshold exists such that all students with expected ability below the threshold attend an L school, whereas all students with expected ability above the threshold attend

[23] More generally, student utility could in each period be a positive monotonic transformation of student productivity, meaning that the transformation could differ in each period.

an H school. Because of the initial disadvantage faced by June-born students in period 1, this threshold is smaller for students born in July than for those born in June, $\widehat{a}^*_{\text{July}} < \widehat{a}^*_{\text{June}}$, which leads to the following proposition:

Proposition 1: Students who are shifted to a more academic middle school because of their birth month ("compliers") are students whose expected ability falls close to indifferent between school types, in the range $[\widehat{a}^*_{\text{July}}, \widehat{a}^*_{\text{June}}]$.

Proof: See Appendix A.2. ■

That is, intuitively, regardless of a June or July birth date, students at the top of the expected ability distribution attend an H school, whereas students at the bottom of the distribution attend an L school in period 1. Why, then, are parents of students in this ability range willing to send their children to an L school even though the initial advantage of delayed school entry fully disappears by the end of the first period? In our model, the answer is that when deciding which type of school their child should attend, parents take into account the student's utility in the first period.[24] Hence, parents trade off a higher utility in the present for a lower utility in the future. Alternatively, we could assume that parents are unaware that students born in June perform more badly on average at the beginning of period 1 than students born in July simply because they are younger. In this case, parents may base their school type decision for period 1 solely on the signal θ and ignore that a child born in June who has the same school grades as a child born in July (but is relatively younger) has a higher expected ability. Modeling school type choice in period 1 in this way has no impact on the key results for our model.

THE EFFECT OF SCHOOL TYPE ATTENDED IN PERIOD 1 ON SCHOOL TYPE ATTENDED IN PERIOD 2 AND PRODUCTIVITY IN PERIOD 3 We then ask how the school type attended in period 1 affects the school type attended in period 2. Because it is switching costs that

[24] This was one of the key arguments of Larry Summers in the "Tiger Mom" debate. Quoting from Gerard Baker's article "Larry Summers vs. the Tiger Mom" in the Wall Street Journal of January 29, 2011, "'People on average live a quarter of their lives as children. That's a lot," Mr. Summers said. "It's important that they be as happy as possible during those 18 years. That counts too."'

lead to the higher ability indifference threshold for students who attend an L school in period 1 versus those who attend an H school, the difference between the two thresholds a_L^* and a_H^* is larger, the larger the switching costs. Given the extreme case of infinite switching costs, no L student moves up to an H school, whereas given the opposite extreme of no switching costs, $a_L^* = a_H^*$, the type of school attended in period 1 has no impact on school type selection in period 2. This observation leads to the following proposition:

Proposition 2: The local average treatment effect of H attendance in period 1 on H attendance in period 2 increases in the switching costs.

Proof: See Appendix A.3. ■

As regards the impact of school type attended in period 1 on wages (i.e., productivity in period 3), we again stress that the overall impact in equation (4) can be decomposed into an indirect effect (i.e., the school type attended in period 1 affects the school type in period 2, which in turn affects productivity in period 3) and a direct effect (i.e., when the school type attended in period 2 is held constant, the school type attended in period 1 affects productivity in period 3); see also equation (1). According to Proposition 2, the school type in period 1 affects the school type in period 2 only if there are switching costs. In this case, a student who attended an H school in period 1 is on average exposed to better peers and teachers than a student of the same ability who attended an L school in that period, which leads to the following proposition:

Proposition 3: (Indirect Effect) The local average treatment effect of attending an H school in period 1 on wages in period 3, which is generated through its effect on H attendance in period 2, is larger the higher the importance of peers and teachers in period 2 (i.e., in γ_2^3 and δ_2^3 in Table A1). This result only holds, however, if the switching costs are positive.

Proof: See Appendix A.4. ■

If, in contrast, the switching costs are zero, the school type attended in period 1 has no impact on the school type attended in period 2. In this case, a productivity difference in period 3 between individuals who attended a different school type in period 1 reflects the direct effect of school type attendance on productivity only, which in turn reflects the exposure to different peers and teachers in period 1. This observation leads to our final proposition:

Proposition 4: (Direct Effect) If the switching costs are zero, the local average treatment effect of attending an H school in period 1 on wages in period 3 is equal to the impact of peers and teachers in period 1 (i.e., $\gamma_1^3(\widehat{\Pi}_1^H - \widehat{\Pi}_1^L) + \delta_1^3(q_1^H - q_1^L)$ in Table A1).

Proof: See Appendix A.4. ■

To distinguish between the direct and indirect effect, we begin our empirical analysis with the impact of middle school type on secondary education type and then turn to the impact of middle school type on wages.

4 Data

Our empirical analysis combines four main data sources, described below. Throughout the analysis, we exclude foreign citizens from our sample because they may have migrated to Germany after beginning school and thus may not have been affected by the school entry cut-off date.

4.1 Social Security Records

Our primary data source is three decades of social security records, covering 1975 to 2006, used to estimate our reduced-form equations (see equation (3)), which relate long-term outcomes to birth month (and date). These data, collected for every individual covered by the social security system, include detailed information on such variables as education, wages, unemployment, occupation, and exact date of birth. Not included are civil servants, the self-employed, and military personnel.[25] From this database, we select all men and women born between 1961 and 1976. The 1961 cohort is the first cohort for which the effective school

[25]In 2001, 77.2 percent of all workers in the German economy were covered by social security and are hence recorded in the data (Bundesagentur für Arbeit, 2004).

entry cut-off falls between June and July. The 1976 cohort is 30 years old in the last year of our data and should thus have completed post-secondary education. For these cohorts, we observe the entire work history from labor market entry onwards, which allows precise computation of their potential and actual labor market experiences. The wages variable refers to April 30 of each year and is deflated using the Consumer Price Index, with 1995 as the base year.[26] We distinguish four educational categories: "no post-secondary education" refers to individuals who graduated from an L or M and did not complete an apprenticeship; "apprenticeship" includes individuals who completed an apprenticeship (as part of the formal German vocational education system) but did not complete college or university, regardless of school type completed; and "college" and "university" refer to individuals who have graduated from college or university, respectively. Our unemployment variable refers to registered unemployment and includes only individuals entitled to unemployment benefits.

4.2 Microcensus

Our second dataset is the scientific use files of the German Microcensus for 1976, 1978, 1980, 1982, 1985, and 1987. We use these data for our first-stage analysis (see equation (2)), which relates the individual middle school type to birth date. We restrict the sample to the same birth cohorts as for the social security records, 1961 to 1976. These data, rather than specifying the exact birth month, provide only a binary indicator for individuals' being born either during the January through April period or the May through December period. Because the school entry cut-off falls in June/July, however, comparing these two periods introduces measurement error in the age of school entry. It could also yield biased estimates in the presence of a strong birth season effect. We therefore assess these biases using data from the School Census for Bavaria and Hesse, which does contain information on birth

[26] In our data, up to 5 percent of the observations are top-coded at the highest wage level for which social security contributions must be paid. In imputing the censored part of the wage distribution, we assume that residuals are normally distributed and allow for heterogeneity in the variance by age group; see Dustmann et al. (2009) for details.

month.

4.3 School Census

Our third data source, the School Census for Bavaria and Hesse, covers all students attending general and vocational schools in these two German states. It is available for the academic school years 2004/05 to 2008/09 for Bavaria, and for the academic school years 2002/03 to 2008/09 for Hesse. The dataset does include information on type of school and grade attended, as well as birth month and birth year. We use these data to assess the bias in our first-stage estimates resulting from the lack of exact birth month information in the Microcensus data for the 1988 to 1994 birth cohorts. We also use them to illustrate the impact of birth month on school type completed for the 1986-1987 and 1984-1987 birth cohorts for Bavaria and Hesse, respectively.

4.4 1987 Census

Our fourth data source is the 1987 census which, unlike the social security records, contains information on the exact school type individuals graduated from. Like the German Microcensus, however, the 1987 census (the last census in Germany before the 2011 census) includes no information on exact birth month, asking respondents instead whether they were born before or after May 24. We use these data to estimate the effect of being born later versus earlier in the year on school type completed by the 1961 to 1963 birth cohorts, who were between 24 and 26 years old at the time of the census and should thus have completed their secondary education.

5 Results

We begin our analysis by showing the first stage that date of birth affects the type of middle school attended (Section 5.1). We then report reduced-form estimates of the impact of birth month and birth date on education completed and labor market outcomes like wages and unemployment. Dividing these reduced-form estimates by the first-stage estimates, we obtain two-sample two-stage least squares (TS-2SLS) estimates of the impact of middle school type attended on these outcomes (Sections 5.2 and 5.3). For first-stage as well as education outcomes, we report results jointly for men and women, since we find no statistically significant differences between the two. For wages, we report results for men only, for reasons we explain in Section 5.3.

5.1 Birth Date and Middle School Type

In Table 3, Panel A, we compare for the 1961 to 1976 birth cohorts (for whom we can study long-term labor market and education outcomes) school type choices at age 14 for children born earlier and later during the year (i.e., January through April or May through December), using data from the German Microcensus which do not contain information on the exact birth month. Because of the limited possibilities to switch school type between grade 5 and grade 9 or 10, the school type attended at age 14 is a good proxy for school type attended throughout middle school.[27] We find that children born later in the year are 2.0 percentage points more likely to attend an H school and 1.9 percentage points less likely to attend an L school than children born earlier in the year, for a total effect of 3.9 percentage points. Hence, the weights a (= 0.019) and b (= 0.020) in the weighted average of the two local average treatment effects in equation (4) are roughly equal: the relative weight for the local average treatment effect on long-term outcomes of attending an M as opposed to an L middle school is 0.49 (0.019/0.039), while the relative weight for the local average treatment effect of attending an H as opposed to an M middle school is 0.51 (0.020/0.039).

As discussed in Section 3, comparing children born earlier and later in the year rather than children born in June versus July leads to two potential biases: seasonal effects (i.e., children born in fall and winter may academically underperform children born in spring and

[27] We illustrate the strong "persistence" of school type chosen throughout grades 5 to 9/10 in Table A3, in which we use data from the School Census for Bavaria and Hesse to investigate whether the impact of birth month on school type choice varies across grades. The results indicate that the effect is roughly constant throughout middle school.

summer irrespective of school entry age) and measurement error (the school entry cut-off date is in July not May). The school census, which contains information on birth month (although not for the 1961 to 1976 cohorts for whom we observe long-term effects) allows us to assess the magnitude of both sources of bias. Based on these data, children born in July are 5.2 (versus 2.0) percentage points more likely to attend an H and 4.2 (versus 1.9) percentage points less likely to attend an L school than children born in June, for a total effect of 9.4 (versus 3.9) percentage points (row (i) of Panel B, Table 3). Figure A1 in Appendix C.1, which displays the share of 14-year-olds attending an H (Panel A) or an L (Panel B) school by birth month, confirms a strong relationship between birth month, and hence age of school entry, and middle school type choice; in particular, the figure shows a clear discontinuity in school type selection around the June/July school entry cut-off date.

In row (ii) of Panel B, we compare, as we do in Panel A using the Microcensus, the school type attended of children born between May and December with those born between January and April. The first stage estimate is now smaller, 7.2 percentage points as opposed to the baseline estimate of 9.4 percentage points. It is therefore important to keep in mind that the first-stage estimate of 3.9 percentage points in Table 3, Panel A, is likely to be underestimated, implying that the resulting TS-2SLS estimates are probably smaller in absolute value than the numbers given below.

Even after adjusting for measurement error, the first-stage results for the recent cohorts in Bavaria and Hesse appear stronger than those for the 1961-1976 cohorts in West Germany (7.2 vs. 3.9 percentage points). In Appendix C.2, we provide evidence that this difference can be wholly explained by the older cohorts' being less compliant with the school entry age cut-off rule than the more recent cohorts. Hence, the mechanism generating the date-of-birth effect on school type choice-namely, that a higher relative age increases the likelihood of attending a more academic school-seems to be very similar for both older and more recent cohorts.

Overall, the birth month has a precisely estimated and quite sizable effect on the type of middle school attended: at least 3.9 percent of students are shifted to a more academic middle school because of their birth month. As previously emphasized, these students are then locked into their respective school types for 5 to 6 years because the opportunities to switch school type in middle school are limited (see also footnote 27 and Table A3.)

5.2 Birth Date, Middle School Type, and Completed Education

Having established that birth month affects middle school type, we now investigate the link between month or date of birth, type of middle school, and type of secondary education. We then carry out a more direct investigation of the up- and downgrading between school types at the end of middle school.

Education Completed In Figure 3, we plot the highest degree completed by date of birth. We distinguish three disjoint post-secondary education outcomes: no post-secondary education (Panel A), apprenticeship (Panel B, the typical pathway for L or M graduates), and college or university (Panel C, the typical pathway for H graduates). The results are based on social security records for the years 1975 to 2006 and refer to all individuals born between 1961 and 1976 aged 30 and older. We use a polynomial of order 5 to fit the points. The vertical line indicates the school entry cut-off date (i.e., July 1), around which, remarkably, neither figure shows a clear discontinuity.[28]

In Table 4, we report the corresponding reduced-form (Panel A) and TS-2SLS estimates (Panel B) obtained by dividing the reduced-form estimates by the first-stage estimate of 0.039. In column (1), we simply compare the educational choices of individuals born in July with those of individuals born in June. The estimates in columns (2) to (4) differ by sampling window (January-December in columns (2) and (3); April-September in column

[28]Note however that the figures reveal strong seasonal effects, with children born in April or May generally outperforming children born in January. These patterns are consistent with those documented by for instance Buckles and Hungerman (2008). We would like to emphasize that as discussed in Section 3.1, seasonal effects do not pose a threat to our identification strategy as long as it is random whether a woman gives birth just before or just after the school entry cut-off date.

(4)) and the order of the birth date polynomial included as a control in the discontinuity regression (5th order, 6th order, and 2nd order in columns (2), (3), and (4), respectively).[29]

All specifications yield similar results: Attending a more academic school through grades 5 to 9/10 has no positive impact on education. Our point estimates even suggest that attending a more academic middle school slightly *reduces* the probability to complete college of university. Note however that this negative impact in the TS-2SLS estimates may be exaggerated because our first stage estimate is likely to be biased downward. Below we also show that the negative effect can be explained by July born individuals being less likely to pursue college education, although they completed an H school, than those who are born in June, a phenomenon we refer to as "downgrading".

Overall, although birth month has a strong effect on the type of school attended throughout the middle school years, it has little effect on the education completed. This finding suggests that, among students shifted to a more academic middle school because of birth month, a substantial share switched school types at the end of middle school. According to Figure 1, such switching may be an "upgrading" of M students to a general or specialized H school, a phenomenon we refer to as a "second chance." It may also be a "downgrading" through failure to enroll in college or university after graduating from an H school. Next, we investigate both the extent of such upgrading and downgrading.

Up- and Downgrading To investigate the extent of upgrading, we first draw on the School Census for Bavaria for the 1986-1987 birth cohorts and that for Hesse for the 1984-1987 birth cohorts. These data allow us to measure the school type individuals graduated from when they are 22 years old and also to distinguish between general and specialized H schools. In Table 5, row (i) of Panel A, we report estimates for the effect of birth date on

[29]These specifications compare educational outcomes of individuals born just before the school entry cut-off date with those born just after the school entry cut-off date, in the same year. Hence, these individuals belong to different school cohorts. We have also compared educational outcomes of individuals born in July with those born in June one year later, and who thus belong to the same school cohort. Our findings are very similar.

graduating from a *general* H school by age 22 (reduced-form, column (1)) and on attending an H school at age 14 (first stage, column (2); see also Table 3, Panel B, row (i)). The estimates are based on a comparison of graduation rates for those born in June versus those born in July (which corresponds to specification (1) in Table 4). The numbers show that by age 22, children born in July are only 2 percentage points more likely to *graduate* from a general H school than children born in June (column (1)). This finding contrasts with our estimates in column (2), in which children born in July are 5.2 percentage more likely to *attend* an H school through grades 5 to 9/10.

In row (ii) of Panel A, Table 5, we include in column (1) students who graduated from a general or specialized H school by age 22 (see Section 2.1 for details). The difference between June- and July-born children now disappears entirely, suggesting that for the particular group of students we consider, the type of middle school attended has no impact on the probability that the individual graduates from any type of H school. These results refer to the 1984-1987 birth cohorts which we cannot follow into the labor market. We confirm the finding that birth date has no impact on the probability that an individual graduates from a generalized of specialized H school for 1961-1963 birth cohorts, using data from the 1987 census in row (iii) of Panel A, Table 5. Note that the 1987 census only includes information on whether the individual was born earlier or later during the year. Again, while children born in July are 2 percentage points more likely to attend a general or specialized H school at age 14, there is no difference between the two groups in the probability to have graduated from a a general or specialized H school by age 22.

In Panel B, Table 5, we provide evidence that a substantial amount of switching occurs also between L and M schools after ninth grade. Here, we base our analysis again on the 1987 census. The results show that children born earlier during the year are only 0.3 percentage points more likely to graduate from at least an M school than children born later during the year. This finding contrasts with our estimates in column (2), which shows that children born earlier during the year are 1.9 percentage points (see also Table 3, Panel A, column (2)) more likely to at least attend an M school during middle school.

Another opportunity to switch school types is given following graduation from a (general or specialized) H school in the form of choosing not to pursue a college or university education. We investigate this type of "downgrading" in Figure 4, where, using the social security records for the 1961-1976 birth cohorts, we plot the share of students who do not attend college, although they graduated from an H school, against date of birth. The figure reveals a clear discontinuity around the school entry cut-off date: among individuals who graduated from an H school, individuals born after the cut-off date are about 0.8 percentage point *more* likely not to complete college or university than individuals born after the cut-off date. This downgrading, it should be noted, can explain why attending a more academic middle school decreases the probability of university or college completion (see Table 4).

In sum, these results indicate that, for students who attend a more academic middle school because of birth month, the type of middle school has a surprisingly small effect on school type completed and education completed. This observation not only suggests that the costs of switching school types after middle school are small but also that the *indirect* effect of middle school type (i.e., through secondary education type) on long-term outcomes like wages is close to zero. Nonetheless, if the type of secondary education is held constant, the type of middle school may affect long-term labor outcomes *directly* through exposure to a more academic school environment throughout middle school. To investigate this direct effect, we next use 1975-2006 social security records for all individuals born between 1961 and 1976 aged 30 and over to link the month (and exact date) of birth-and thus middle school type-to labor market outcomes.

5.3 Birth Date, Middle School Type, and Labor Market Outcomes

Note that since individuals who are born in July start school later, and hence enter the labor market later, they have, at any given age, accumulated less labor market experience than children born in June. To identify the causal effect of school type selection in middle school on wages, we need to eliminate this experience effect. This requires knowledge of the returns to potential experience. Estimating these returns is particularly challenging for women, due to the changing selection of women into work over the life cycle. We therefore focus here on men only. We eliminate the experience effect by first estimating the returns to potential experience based on a 4th order polynomial using OLS and then subtracting these estimates from raw log-wages. Our results remain robust to alternative ways of estimating the returns to potential experience (see also Table A4).

WAGES In Panel A, Figure 5, we plot the raw log wages by birth date using a polynomial of order 5. The figure shows a clear wage disadvantage for children born in July, i.e., just after the school entry cut-off date. As explained above, this is because at any given age, individuals born in July have accumulated less labor market experience than individuals born in June. The wage disadvantage disappears, however, once we plot log wages adjusted for differences in experience (see Panel B).

In Table 6, we present the results for the same specifications of equation (5) as in Table 4, reporting the reduced-form estimates in Panel A and the TS-2SLS estimates in Panel B. In row (i) of Panel A, we display the effects of birth date on wages without adjusting for differences in potential work experience. The findings mirror those in Figure 4, Panel A: in all specifications, workers born after the school entry cut-off date earn 0.3 to 0.4 percent lower wages. In line with Figure 4, Panel B, this statistically significant effect reduces to a statistically insignificant point estimate of around zero when we adjust wages for differences in potential work experience. Dividing these experience-adjusted reduced-form estimates by our first-stage estimate of 0.039 (see column (3) of Panel A, Table 3), we obtain statistically insignificant TS-2SLS point estimates ranging between -2.4 and +0.6 percent (Panel B, Table 6). Keeping in mind that the first-stage estimate is likely to be biased downward (as discussed in Section 5.1), the true effect is thus likely to be even closer to zero. Using a one-sided hypothesis test (appropriate because we expect school quality to positively affect wages), we can exclude that attending a more academic middle school increases wages by more than 1.7 percent based on our baseline specification in column (1).

These findings are further supported by additional robustness checks that use different methods to adjust for potential experience (e.g., considering only individuals older than 40 years whose wages do not increase with experience). The findings of these tests are reported and explained in detail in Table A4 in Appendix C.3. Taken together, these result suggest that, once secondary education type is held constant, the *direct* effect of middle school type on wages is small.

OTHER LABOR MARKET OUTCOMES We next turn to the impact of middle school type attended on three other labor market outcomes: occupational choice (distinguishing between blue- and white-collar occupations), the share of days in registered unemployment, and the share of days working full time since labor market entry. Occupational choice is a particularly interesting variable because historically, L schools prepare students for apprenticeships in blue-collar occupations, whereas M schools prepare them for apprenticeships in white-collar occupations. We again base our analysis on 1975-2006 social security records for individuals born between 1961 and 1976 aged 30 and over. Since we find no statistically significant differences between men and women, we report results jointly for the two sexes.

We report reduced-form and TS-2SLS estimates in Table 7, which has the same structure as Tables 4 and 6. Both the reduced-form (Panel A) and TS-2SLS estimates (Panel B) are closely centered around zero and typically not statistically significant from zero, suggesting that once secondary education type is held constant, the *direct* effect of attending a more academic middle school on occupational choice, unemployment, and full-time work is small. Our findings on unemployment and full-time work further imply that our wage results in Table 6 do not suffer from an employment selection bias.

6 Discussion, Interpretation, and Conclusions

Simple correlations (see Table 2), derived from survey data, between middle school type attended and school completion indicate that students attending an H school at age 14 are 50 percentage points more likely to graduate from an H school than students attending an M school at age 14. These latter in turn are 64 percentage points more likely to graduate from at least an M school than students attending an L school at age 14. There is also a strong association between school type attended at age 14 and wages, which implies that individuals who attended an H (M) school at age 14 earn at least 10 percent higher wages than those who attended an M (L) school at that age.

These associations, however, confound the causal effect of middle school type attended with the selection effect of more proficient students being sorted into more academic schools. Our experiment, in contrast, compares the outcomes for individuals with the same abilities who are allocated to different middle school types because of birth date. The resulting causal estimates imply that, in stark contrast to the strongly positive associations indicated by the simple associations, attending a more academic middle school has little effect on school type completed, education completed, wages, unemployment, or occupational choice. Note that we identify a local average treatment effect for individuals shifted to a better middle school because of birth date. Our model in section 3.2 indicates that these are individuals within a certain expected ability range that falls close to indifferent between the two school types (see Proposition 1).

There are two channels through which attending a more academic middle school may affect long-term outcomes: indirectly, through the type of secondary education, and directly, when the secondary education type is held constant. As regards the first, we find that attending a more academic middle school has no effect on secondary education type, as a considerable share of students switch between school types at the end of middle school (age 15/16) and at the end of high school (age 18/19). In our model, this up- or downgrading implies that the academic costs of switching school types are low (see Proposition 2).

If middle school type has no effect on secondary education type, it can affect wages and other labor market outcomes only through the direct channel (when secondary education type is held constant). Our model (section 3.2) implies that, in the absence of switching costs, this direct effect equals the impact of peers and teachers in period 1 (Proposition 4). Our findings on wages and other long-term outcomes provide strong evidence that this direct effect is zero (i.e., $\gamma_1^3 = \delta_1^3 = 0$; see Table A1).

Our overall conclusion, therefore, is that even longer exposure to a less challenging school environment may not have any detrimental long-term effects as long as the education system provides enough opportunities for students to revise the educational choices made earlier in life. This message is positive in that it implies that disadvantaged youth can make up for substantial weaknesses in early education at later stages of their educational career provided they are given the opportunity to reassess and correct initial choices. This provision of "second chances", however, has received little attention in the literature despite being, we believe, fundamental for the assessment of any educational system, regardless of whether school choice is determined primarily by ability (as in Germany) or residential location (as in the U.S. or UK). Educational systems in different countries should thus be judged not only on how and when they segregate students, but also on the extent to which they allow students to remedy initial choices that may have been based on incomplete or faulty information about their learning potential.

References

[1] Abdulkadiroglu, A., Angrist, J.D., Dynarski, S.M., Kane, T.J., and P.A. Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics*, 125: 699-748.

[2] Abdulkadiroglu, A., J. D. Angrist, and P. A. Pathak. 2011. "The Elite Illusion: Achievement Effects of Boston and New York Exam Schools." NBER Working Paper 17264.

[3] Altonji, J., T. Elder and C. Taber. 2005. "Selection on Observed and Unobserved Vari-

ables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113: 151-184.

[4] Ammermüller, A. and J.S. Pischke. 2009. "Peer Effects in European Primary Schools: Evidence from PIRLS." *Journal of Labor Economics*, 27: 315-348.

[5] Angrist, J. D. and A. B. Krueger. 1991. "Does Compulsory School Attendance Affect Education and Earnings? *Quarterly Journal of Economics*, 106: 979-1014.

[6] Angrist, J. D., S. M. Dynarski, T. J. Kane, P. A. Pathak, and C. Walters. 2010. "Inputs and Impacts in Charter Schools: KIPP Lynn." *American Economic Review Papers and Proceedings*, 100: 239-243.

[7] Bedard, K. and E. Dhuey. 2006. "The Persistence of Early Childhood Maturity: International Evidence of Long-term Age Effects." *Quarterly Journal of Economics*, 121: 1437-1472.

[8] Bifulco, R., J. Fletcher and S. Ross. 2011. "The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy*, 3: 25-53.

[9] Black, S. and S. Machin. 2010. "Housing Valuations of School Performance". Chapter 10 of *Handbook of the Economics of Education*, Vol. 3, E. A. Hanushek, S. Machin, and L. Woessmann (ed.), North-Holland.

[10] Black, S., P. Devereux, and K. G. Salvanes. 2010. "Under Pressure: The Effect of Peers on Outcomes of Young Adults." NBER Working Paper 16004.

[11] Black, S., Devereux, P. and K. G. Salvanes. 2011. "Too Young to Leave the Nest: The Effects of School Starting Age." *Review of Economics and Statistics*, 93: 455-467.

[12] Bound, J. and D. Jaeger. 2000. "Do Compulsory Attendance Laws Alone Explain the Association between Earnings and Quarter of Birth?", *Research in Labor Economics*, 19: 83-108.

[13] Bound, J., D. Jaeger and R.M. Baker. 1995. "Problems with Instrumental Variables Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable Is Weak", *Journal of the American Statistical Association*, 90: 443-450.

[14] Brunello, G. and D. Checchi . 2007. "Does School Tracking Affect Equality of Opportunity? New International Evidence." *Economic Policy*, 52: 781-861.

[15] Buckles, K. and D.M. Hungerman. 2008. "Season of Birth and Later Outcomes: Old Questions, New Answers." NBER Working Paper 14573.

[16] Cascio, E. and D. Whitmore Schanzenbach. 2007. "First in Class? Age and the Education Production Function." NBER Working Paper 13663.

[17] Chatty, R., Friedman, J.N., Hilger, N.. Saez, E., Whitmore Schanzenbach, D. and D. Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *Quarterly Journal of Economics*, 126, 1593-1660.

[18] Chatty, R., Friedman, J.N., and J. Rockoff. 2011. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." NBER Working Paper 17699.

[19] Clark, D. 2010. "Selective Schools and Academic Achievement." *B.E. Journal of Economic Analysis & Policy*, 10 (Advances), Article 9.

[20] Crawford, C., L. Dearden, and C. Meghir. 2007. *When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England.* London: Institute for Fiscal Studies.

[21] Cullen, J., B. Jakob, and S. Levitt. 2005. "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." *Journal of Public Economics*, 85: 729-760.

[22] Cullen, J., B. Jakob, and S. Levitt. 2006. "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries." *Econometrica*, 74: 1191-1230.

[23] Duflo, E., P. Dupas and M. Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review*, 101: 1739-1774.

[24] Dustmann, C., J. Ludsteck and U. Schoenberg. 2009. "Revisiting the German Wage Structure." *Quarterly Journal of Economics*, 124: 843-881.

[25] Elder, T. E. and D. H. Lubotsky. 2009. "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers." *Journal of Human Resources*, 44: 641-683.

[26] Evans, W. and R. Schwab. 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics*, 110: 941-974.

[27] Figlio, D. N. and M. E. Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics*, 51: 497-514.

[28] Fredriksson, P. and B. Öckert. 2006. "Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labour Market Performance." IFAU Working Paper 12, Uppsala, Sweden.

[29] Gould, E. D., V. Lavy, and M. D. Paserman. 2004. "Immigrating to Opportunity: Estimating the Effect of School Quality Using a Natural Experiment on Ethiopians in Israel." *Quarterly Journal of Economics*, 119: 489-526.

[30] Gould, E. D., V. Lavy, and M. D. Paserman. 2009. "Long-term Classroom Peer Effects: Evidence from Random Variation in Enrollment of Disadvantaged Immigrants." *Economic Journal*, 119: 1243-1269.

[31] Grenet, J. 2010. "Academic Performance, Educational Trajectories and the Persistence of Date of Birth Effects. Evidence from France." Mimeo, Paris School of Economics.

[32] Hanushek, E.A., Kain, J. and S. Rivkin. 2005. "Teachers, Schools, and Acadamic Achievement." *Econometrica*, 73: 417-458.

[33] Hanushek, E. A. and L. Wössmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries." *Economic Journal*, 116: C363-C376.

[34] Hoxby, C. M. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.

[35] Hoxby, C. M. and S. Murarka. 2009. "Charter Schools in New York City: Who Enrolls and How They Affect Student Achievement." NBER Working Paper 14852.

[36] Inoue, A. and G. Solon. 2010. "Two-Sample Instrumental Variables Estimators." *Review of Economics and Statistics*, 92: 557-561.

[37] Jackson, C. K. 2010. "Do Students Benefit from Attending Better Schools? Evidence from Rule-Based Student Assignments in Trinidad and Tobago." *Economic Journal*, 120: 1399-1429.

[38] Kinsler, J. 2011. "Beyond Levels and Growth: Estimating Teacher Value-Added and it Persistence." Mimeo, University of Rochester.

[39] Lavy, V. and A. Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics*, 3: 1-33.

[40] Lee, D.S. and D. Card. 2007. "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics*, 144: 655-674.

[41] Manning, A. and J.S. Pisckhe. 2006. "Comprehensive versus Selective Schooling in England in Wales, What Do We Know?" IZA Discussion Paper No. 2072.

[42] Malamud, O. and C. Pop-Eleches. 2011. "School Tracking and Access to Higher Education among Disadvantaged Groups." *Journal of Public Economics*, 95(11-12): 1538-1549.

[43] Maurin, E. and S. McNally. 2007. "Educational Effects of Widening Access to the Academic Track: A Natural Experiment." Institute for the Study of Labor (IZA) Discussion Paper 2596, Bonn, Germany.

[44] Muehlenweg, A. and P.A. Puhani. 2010. "Persistence of the School Entry Age Effect in a System of Flexible Tracking." *Journal of Human Resources*, 45: 407-435.

[45] Neal, D. 1997. "The Effects of Catholic Secondary Schooling on Educational Achievement." *Journal of Labor Economics*, 15: 98-123.

[46] Pop-Eleches, C. and M. Urquiola. 2010. "Going to a Better School: Effects and Behavioral Responses." Mimeo, Columbia University.

[47] Puhani, P. A. and A. M. Weber. 2007. "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Early Educational Effects of Age of School Entry in Germany." *Empirical Economics*, 32: 359-386.

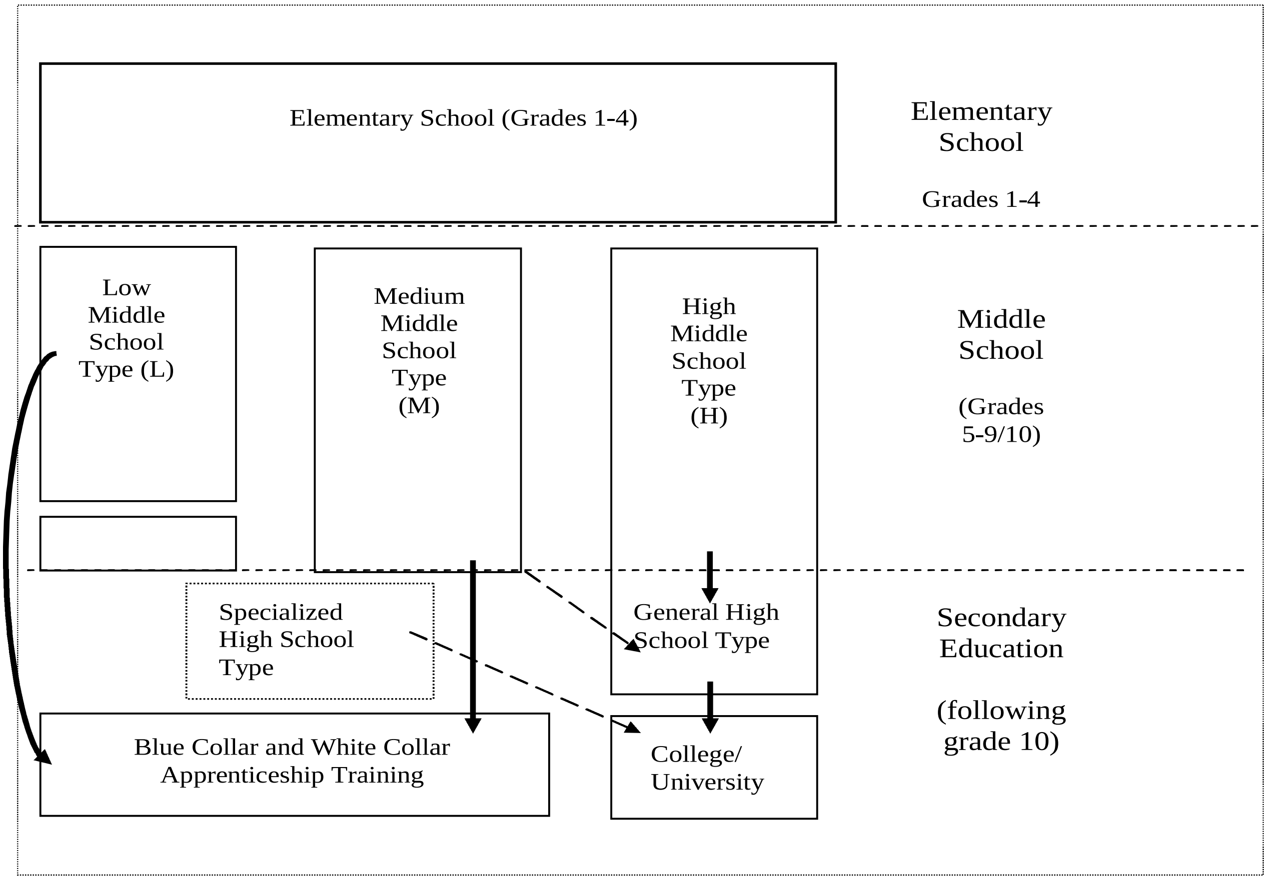
[48] Rothstein, J. 2010. “Teacher Quality and Educational Production: Tracking, Decay, and Student Achievement.” *Quarterly Journal of Economics,* 125: 175-214.

[49] Solli, I. 2011. “Left Behind by Birth Month.” Mimeo, University of Stavanger.

[50] Schanzenbach-Whitmore, D. 2009. “Experimental Estimates of Peer Effects.” Mimeo, Northwestern University.

[51] Waldinger. 2007. “Does Ability Tracking Exacerbate the Role of Family Background for Students’ Test Scores?” Mimeo, University of Warwick.

Figure 1: The German Education System



Note : The figure provides an overview of the German education system. Students who have attended a middle school (grades 5-9/10) of low (L) type would typically start an apprenticeship in a blue collar occupation after 9th grade. Students who have attended a middle school of medium (M) type would typically start an apprenticeship in a white collar occupation after 10th grade. Students who have attended a middle school of high (H) type would typically continue at this school until grade 13 and then enter college or university. Graduation from a specialized high school type grants access to college or university, but possibly restricts the field of study.

Figure 2: Research Design

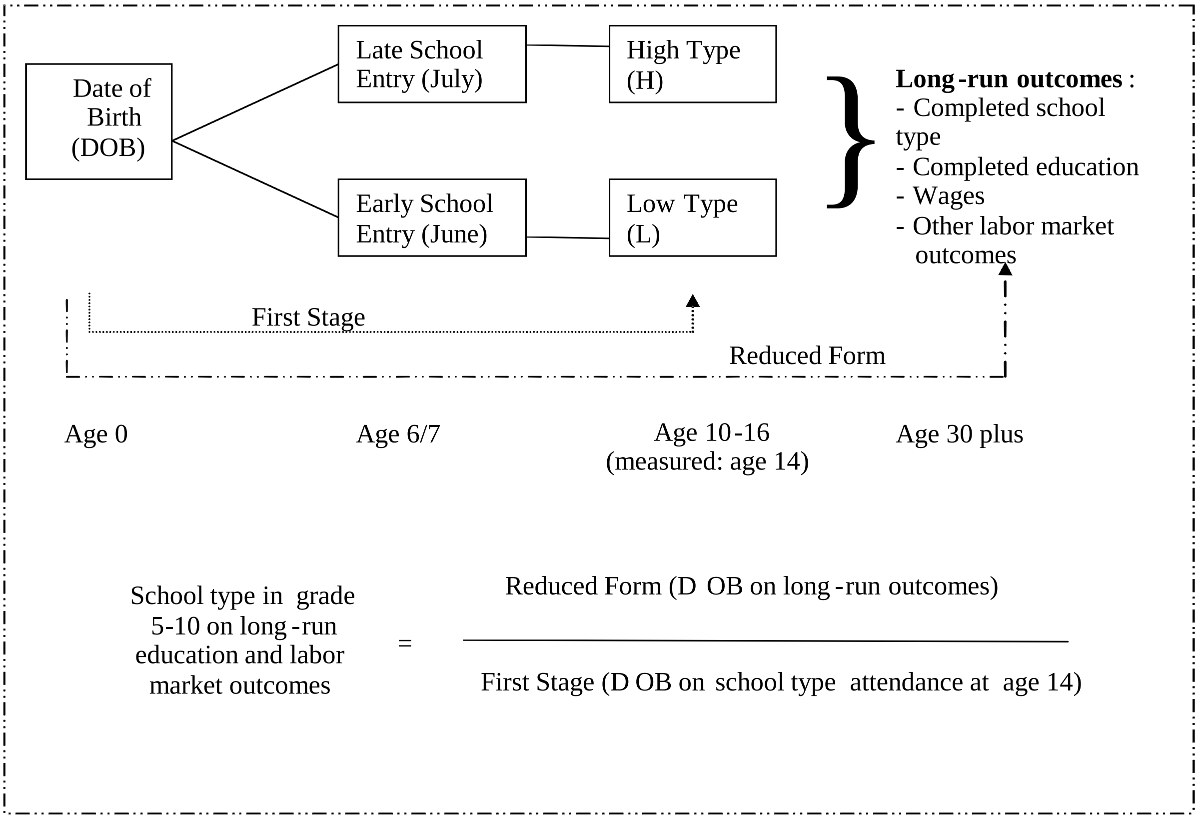


Table 1: Differences between School Types: Peer Exposure, Teacher Quality, and Teaching Technology

	Low (L)	Medium (M)	High (H)
Peer Exposure (9th grade)			
Test Scores			
Reading score	-0.900	-0.183	0.768
Mathematics score	-0.908	-0.217	0.840
Parental Background			
Mother's education (years)	9.88	11.71	13.91
Father's education (years)	9.91	11.53	14.3
% of households with less than 25 books	39.81	23.49	5.48
% of households below median income	76.02	65.32	38.72
Share Girls	40.68	49.19	54.37
Teacher Quality			
Minimum Length of Study	4 years	4 years	5 years
Salary	46,872	46,872	49,530
Teaching Intesnity (9th Grade)			
total hours per week	32	32	36
2nd and 3rd foreign language	0	0	6
Natural and computer science	3	6	8
Social science	3	6	4
Physical education	4	4	2
Vocational Subjects	6	0	0
Learning Goals (9th grade)			
German:	writing CVs, filling out forms, discussions and rules to resolve comflict.	discussions and rules to resolve conflict	detailed explanations, analysis and interpretation of various types of texts; creative writing
Mathematics:	equations with rational numbers;Phytagoras; prisma, pyramids, cones, cylinder; descriptive statistics.	real numbers and powers; linear equations with two unknowns; functions excluding exponential and broken polynomial functions; trigonometry, prisma, pyramids, cones, cylinder	functions, including exponential, trigonometric and broken polynomial functions; Phytagoras, trigonometry, prisma, pyramids, cones, cylinder; stoachstics

Note : The table reports differences in peer exposure at age 15, teacher quality, teaching intensity and learning goals in grade 9 between school types. Reading and mathematics test scores are normalized to have mean zero and standard deviation 1.

Data Source s: Peer Exposure: Programme for International Student Assessment (PISA), 2003 and 2006. Teacher Quality: various Ministries of Education of various West German states. Teaching Intensity: Ministry of Education, State of Bavaria. Learning Goals: Ministries of Education of various West German states (Baden-Württemberg, Bavaria, Hesse, Lower Saxony, North Rhine-Westphalia, Rhineland-Palatine).

Table 2: School Type Completion and Log-Run Outcomes

Panel A: School Type Attendance at Age 14 and School Type Completion	completed a H type (including specialized)	completed at least M type
attended low (L) type at age 14/15	0.066	0.226
attended medium (M) type at age 14/15	0.308	0.866
attended high (H) type at age 14/15	0.808	0.902
p-value	0.000	0.000

Panel B: School Type Completion and Long-Run Outcomes	Wages	Unemployment	No Post-Secondary Education	College/ University
completed low (L) type	2.376	0.076	0.230	0.016
completed medium (M) type	2.485	0.042	0.080	0.032
completed high (H) type	2.672	0.037	0.080	0.541
p-value	0.000	0.000	0.000	0.000

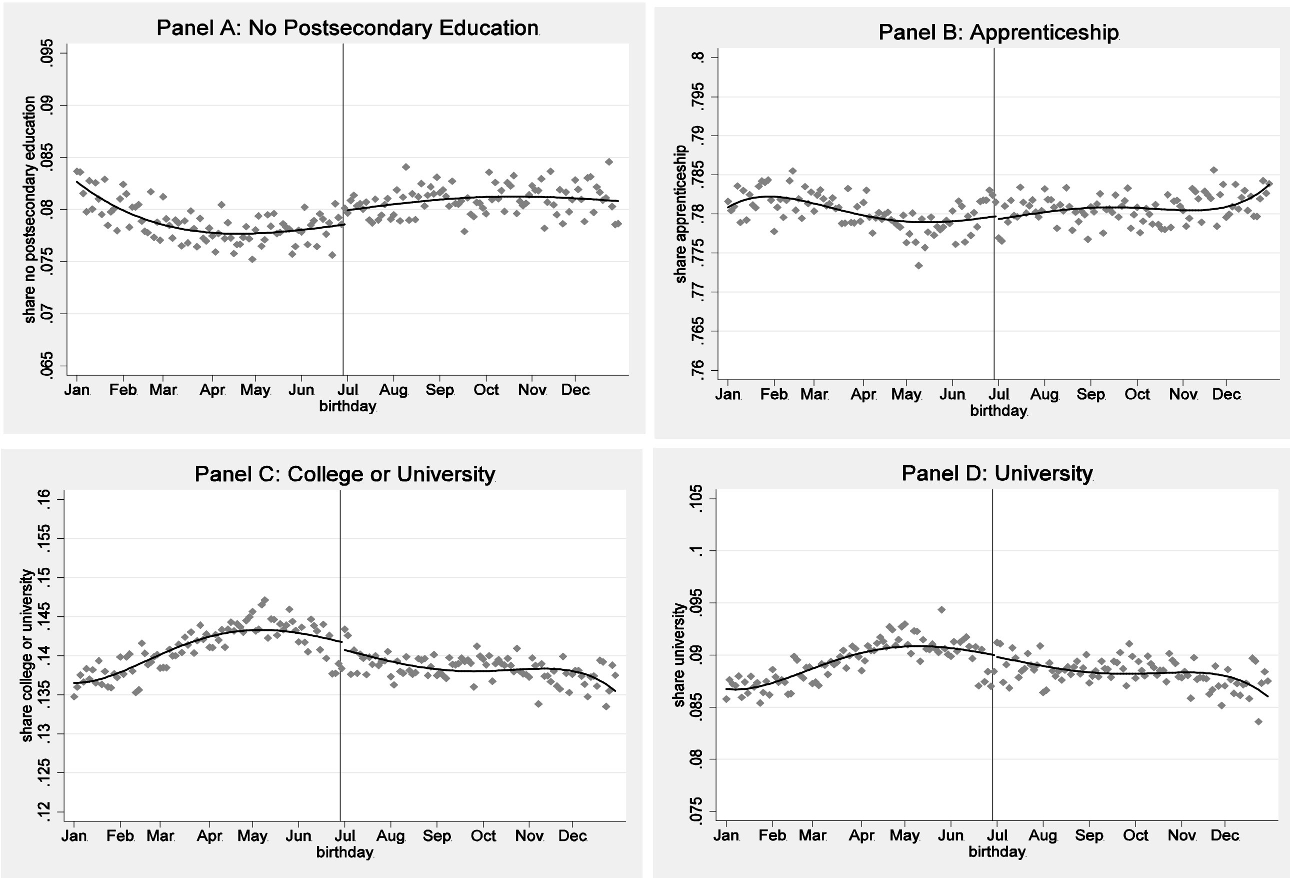
Note : In Panel A, we report the probability that an individual who attended an L, M, or H middle school type at age 14 or 15 completes a general or specialized H school type, or at least an M school type at age 21 (or older). The sample consists of West German citizens who are observed at age 14 or 15 in the 1984 to 2000 waves of the German Socio-Economic Panel as well as in the year they turn 22 in the 1990 to 2007 waves of the German Socio-Economic Panel (1,678 individuals). In Panel B, we display the differences in wages, unemployment rates, apprenticeship completion, and college or university graduation between individuals who completed an L, M and (specialized or general) M school type, respectively. The sample consists of West German citizens born between 1961 and 1976 aged 30 through 46 in the 1991 to 2007 waves of the German Socio-Economic Panel. The last rows in each panel report the p-value for the hypothesis that there are no differences between middle school type attendance at age 14 or 15 and school type completion (Panel A) or school type completion and wages, unemployment, and completed education (Panel B). Sampling weights provided by the GSOEP are used.
Source : German Socio-Economic Panel (GSOEP), 1984-2007.

Table 3: Birth Date and School Type Attendance at Age 14
(First Stage)

	(1) High (H) versus Medium (M) or Low (L)	(2) High (H) or Medium (M) versus Low (L)	(3) Sum
Panel A: Birth Cohorts 1961-1976 (Microcensus)			
(i) May-Dec versus Jan-April	0.020	0.019	0.039
	(0.005)**	(0.005)**	(0.009)**
N=37,808			
Panel B: Measurement Error and Seasonal Effects (School Census, Birth Cohorts 1988-1994)			
(i) July versus June	0.052	0.042	0.094
(s.e.)	(0.002)**	(0.002)**	(0.004)**
N=170,832			
(ii) May-Dec versus Jan-April	0.033	0.039	0.072
(s.e.)	(0.001)**	(0.001)**	(0.002)**
N=990,854			

Note : In Panel A, we report, for selected cohorts born between 1961 and 1976, the difference in the share of students who attend an H (versus M or L) school type (column (1)) or an L (versus a M or H) school type (column (2)) at age 14 between students born earlier (January to April) and later (May to December) during the year. In column (3), the dependent variable is coded 2 if the students is in a high school type, 1 if she is in an M school type and 0 if she is in an L school type, corresponding to the sum of the coefficients in the first two columns. In Panel B, we report the impact on birth month on school type attendance at age 14 for the birth cohorts 1988 to 1994, using two different cut-offs. Robust standard errors in parentheses.
Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.
Source : Panel A: Microcensus, selected years 1976 to 1987. Panel B: School Census for Bavaria and Hesse, 2002-2009.

Figure 3: Birth Date and Completed Education



Note: The figures plot the relationship between birthday and the share of individuals without post-secondary education (Panel A), the share of individuals who completed an apprenticeship (Panel B), the share of individuals who graduated from college or university (Panel C), and the share of individuals who graduated from university (Panel D). Each dot refers to the average share for two birthdays (e.g. 3rd and 4th of January). We also plot predicted shares which we obtain from a regression that controls, in addition to an indicator variable equal to 1 for individuals born after July 1st, for a polynomial of order 5 in the date of birth. The vertical lines indicate the school entry cut-off date.
Source : Social Security Data, 1975-2006.

Table 4: Type of Middle School and Completed Education

Panel A: Reduced Form Estimates		(1) Jun-Jul, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol. 2
(i) No Post-secondary Education					
	Coeff.	0.001	0.001	0.001	0.002
	(s.e.)	(0.000)**	(0.001)*	(0.001)	(0.001)**
(ii) Apprenticeship					
	Coeff.	0.000	0.000	0.000	0.001
	(s.e.)	(0.001)	(0.001)	(0.001)	(0.001)
(iii) College or University					
	Coeff.	-0.002	-0.001	-0.001	-0.002
	(s.e.)	(0.000)**	(0.001)	(0.001)	(0.001)**
(iv) University					
	Coeff.	-0.001	0.000	0.000	-0.001
	(s.e.)	(0.000)	(0.001)	(0.001)	(0.001)
	N	1,961,320	11,609,855	11,609,855	5,905,126

Panel B: Two Sample Two Stage Least Squares Estimates		(1) Jul-Jun, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol. 2
(i) No Post-secondary Education					
	Coeff.	0.037	0.031	0.031	0.042
	(s.e.)	(0.013)**	(0.017)*	(0.017)	(0.017)**
(ii) Apprenticeship					
	Coeff.	0.009	0.001	-0.001	0.019
	(s.e.)	(0.015)	(0.023)	(0.022)	(0.021)
(iii) College or University					
	Coeff.	-0.046	-0.032	-0.030	-0.061
	(s.e.)	(0.016)**	(0.024)	(0.022)	(0.024)**
(iv) University					
	Coeff.	-0.018	-0.010	-0.008	-0.028
	(s.e.)	(0.011)	(0.018)	(0.017)	(0.016)

Note : Panel A reports various reduced-form estimates for the impact of the exact date of birth on the share of individuals without post-secondary education, the share of individuals who completed an apprenticeship or graduated from a general or specialized high school type, the share of individuals who graduated from college or university, and the share of individuals who graduated from university. In column (1), we report the difference in the respective shares between individuals born in July and June; robust standard errors in parentheses. In columns (2) and (3), we display the coefficient on being born on or after July 1st from a regression that controls for a polynomial of order 5 and 6, respectively, in the day of birth, and includes all individuals, born between January and December. In column (4), we restrict the sample to individuals born betweeen April and September, and include a polynomial of order 2 in the day of birth. Standard errors are clustered at the day of birth.

In Panel B, we report the corresponding Two-Sample Two-Stage-Least-Squares estimates for the impact of attending a moe academic middle school on completed education, by dividing the reduced-form estimates in Panel A by the first stage of 0.039 (Table 4, Panel A, column (3)). Standard errors are computed using the Delta Method, see footnote 20 for details.

Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.

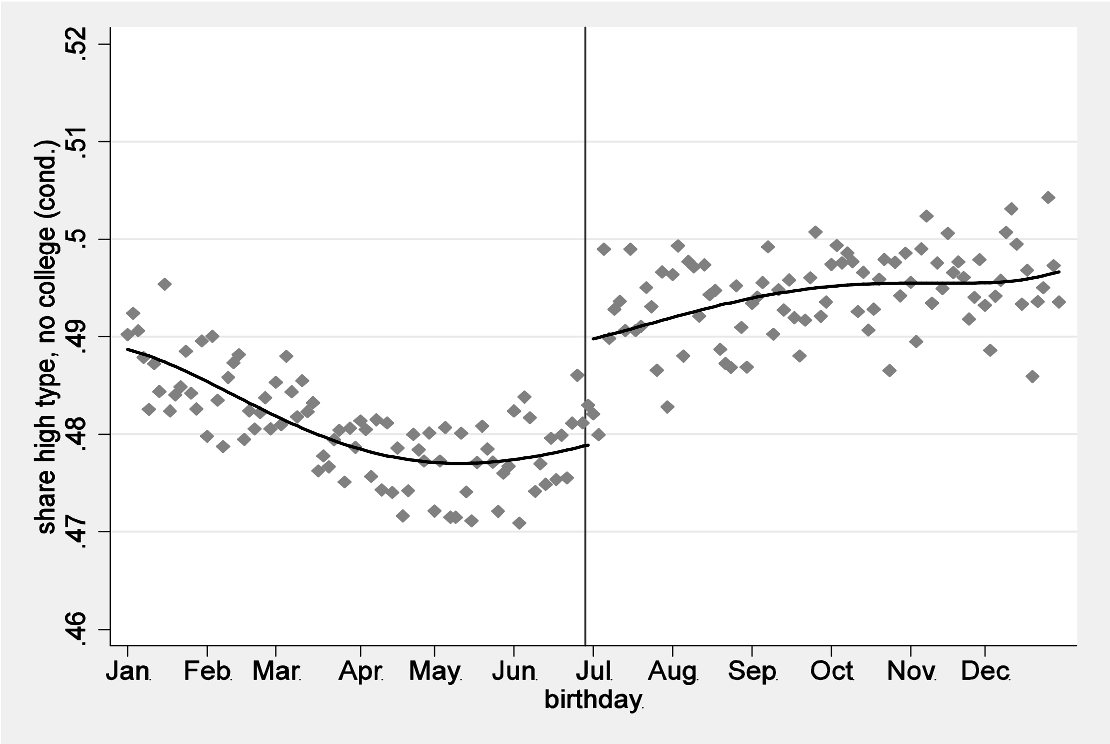
Source : Reduced Form: Social Security Data, 1975-2006. First Stage: Microcensus, selected years 1976 to 1987.

Table 5: Type of Middle School and School Type Completion

	(1) Age 22 (Reduced Form)	(2) Age 14 (First Stage)
Panel A: Graduation from an H school		
(i) Graduation from general H school (School Census)		
July versus June	0.020	0.052
(s.e.)	(0.004)**	(0.002)**
(ii) Graduation from general or specialized H school (School Census)		
July versus June	-0.001	0.052
(s.e.)	(0.005)	(0.002)**
(iii) Graduation from general or specialized H school (1987 Census)		
June-Dec verus Jan-May	0.000	0.020
(s.e.)	(0.001)	(0.005)**
Panel B: Graduation from at least an M school		
Graduation from a M or L school		
June-Dec verus Jan-May	0.003	0.019
(s.e.)	(0.001)**	(0.005)**

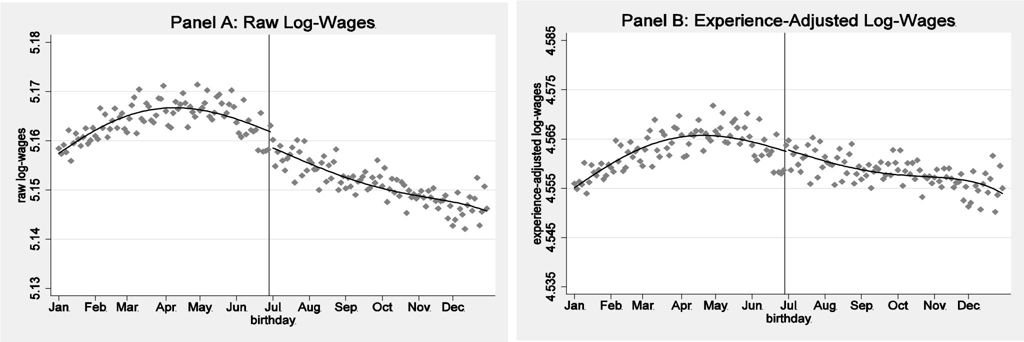
Note: In Panel A, we report in the first column ("Age 22") the difference in the share of students who graduated from a general high school type (row (i)), or from a general or specialized high school type (row (ii)), between students born in July and June for the birth cohorts 1984 to 1987. In row (iii), we compare for the birth cohorts 1961 to 1963 the share of students who graduated from a general or specialized high school between students who were born earlier or later during the year. In the second column ("Age 14"), we show the difference in the share of students attending a high school type at age 14 between students born in July and June (rows (i) and (ii); see also Table 3, Panel B, column (1)) and between students born earlier or later during the year (row (iii); see also Table 3, Panel A, column (1)). See Appendix B for calculation of standard errors of the reduced-form estimates.
In Panel B, we first report, for birth cohorts 1961 to 1963, the difference in the share of students who completed at least a medium school type between students born earlier (January versus May) and later (June versus December) during the year (column (1)). We then display in column (2) the difference the share of students who attend at least a medium school at age 14 between students born earlier or later during the year (see also Table 3, Panel A, column (2)). Robust standard errors in parentheses.
Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.
Source: Panel A, rows (i) and (ii): School Census for Bavaria and Hesse, 2002-2008. Panel A, row (iii) and Panel B: Reduced Form: Census 1987. First Stage: Microcensus, selected years 1976 to 1987.

Figure 4: Birth Date and Downgrading from a High School Type



Note: The figure plots, for individuals who have graduated from a general or specialized high school, the relationship between birthday and the share of individuals who "downgrade" by not graduating from college or university. Each dot refers to the average share for two birthdays (e.g. 3rd and 4th of January). We also plot predicted shares which we obtain from a regression that controls, in addition to an indicator variable equal to 1 for individuals born after July 1st, for a polynomial of order 5 in the day of birth. The vertical line indicates the school entry cut-off date.
Source: Social Secuirty Data, 1975-2006.

Figure 5: Birth Date and Wages (Men only)



Note : The figures plot the relationship between birthday and raw (Panel A) and experience-adjusted (Panel B) log-wages. In Panel B, the experience effect is eliminated by estimating returns to potential experience using OLS, imposing a functional form of a 4th order polynomial in potential experience, and substracting these from raw log-wages. The figures are based on individuals aged 30 and over with a valid wage. Each dot refers to the average wage for two birthdays (e.g. 3rd and 4th of January). We also plot predicted raw and experience-adjusted log-wages obtained from a regression that controls, in addition to an indicator variable equal to 1 for individuals born after July 1st, for a polynomial of order 5 in the date of birth. The vertical lines indicate the school entry cut-off date.
Source : Social Secuirty Data, Men, 1975 to 2006.

Table 6: Type of Middle School and Wages

Panel A: Reduced Form Estimates

	(1) Jun-Jul, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol. 2
(i) Raw wages (age 30 and higher)				
Coeff.	-0.005	-0.003	-0.003	-0.004
(s.e.)	(0.001)**	(0.001)**	(0.001)**	(0.001)**
(ii) Wages net of experience (age 30 and higher)				
Coeff.	-0.001	0.000	0.000	-0.001
(s.e.)	(0.001)	(0.001)	(0.001)	(0.001)
N	810,679	4,807,959	4,807,959	2,444,420

Panel B: Two-Sample Two-Stage Least Squares Estimates

	(1) Jul-Jun, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol.2
(i) Wages net of experience (age 30 and higher)				
Coeff.	-0.024	0.000	0.006	-0.023
(s.e.)	(0.025)	(0.033)	(0.032)	(0.029)

Note: In Panel A, we report various reduced-form estimates for the impact of day of birth on log-wages, for men aged 30 and older with a valid wage. Our first outcome variable is the raw log-wage. We then eliminate the experience effect, by estimating returns to potential experience using OLS and imposing a functional form of a 4th order polynomial in potential experience. In column (1), we report the difference in raw and experience-adjusted log-wages between individuals born in July and June; standard errors clustered at the individual level in parentheses. In columns (2) and (3), we display the coefficient on being born on or after July 1st from a regression that controls for a polynomial of order 5 and 6, respectively, in the day of birth, and includes all individuals, born between January and December. In column (4), we restrict the sample to individuals born betweeen April and September, and include a polynomial of order 2 in the day of birth. Standard errors are clustered at the individual level as well as at the day of birth.

In Panel B, we report the corresponding Two-Sample Two-Stage Least Squares estimates for the impact of attending a better middle school on (experience-adjusted) log-wages, by dividing the reduced form estimates in Panel A by the first stage of 0.039 (Table 3, Panel A, column (3)). Standard errors are computed using the Delta Method, see footnote 20 for details.

Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.

Source: Reduced Form: Social Security Data, Men, 1975-2006. First Stage: Microcensus, selected years 1976 to 1987.

Table 7: Type of Middle School and Occupational Choice, Experience, and Unemployment

Panel A: Reduced Form Estimates

		(1) Jun-Jul, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol. 2
(i) White collar occupation					
	Coeff.	0.000	0.001	0.002	0.000
	(s.e.)	(0.001)	(0.001)	(0.001)	(0.001)
(ii) Share days worked					
	Coeff.	0.001	0.000	0.000	0.001
	(s.e.)	(0.000)	(0.001)	(0.001)	(0.002)
(iii) Share days in unemployment					
	Coeff.	0.000	0.000	0.000	0.000
	(s.e.)	(0.000)	(0.000)	(0.000)	(0.002)
	N	1,573,136	9,317,405	9,317,405	4,739,278

Panel B: Two-Sample Two-Stage Least Squares Estimates

		(1) Jul-Jun, none	(2) Jan-Dec, pol. 5	(3) Jan-Dec, pol. 6	(4) Apr-Sept, pol. 2
(i) White collar occupation					
	Coeff.	0.013	0.037	0.040	-0.001
	(s.e.)	(0.021)	(0.029)	(0.028)	(0.025)
(ii) Share days worked					
	Coeff.	0.015	0.005	0.005	0.013
	(s.e.)	(0.010)	(0.013)	(0.014)	(0.061)
(iii) Share days in unemployment					
	Coeff.	-0.005	-0.004	-0.004	0.004
	(s.e.)	(0.004)	(0.005)	(0.005)	(0.061)

Note : In Panel A, we report various reduced-form estimates for the impact of day of birth on whether the current occupation is a white collar occupation, and on the share days spent working and in unemployment since labor market entry, for individuals aged 30 and over. In column (1), we report the difference in these outcomes between individuals born in July and June; standard errors clustered at the individual level in parentheses. In columns (2) and (3), we display the coefficient on being born on or after July 1st from a regression that controls for a polynomial of order 5 and 6, respectively, in the day of birth, and includes all individuals, born between January and December. In column (4), we restrict the sample to individuals born betweeen April and September, and include a polynomial of order 2 in the day of birth. Standard errors are clustered at the individual level as well as at the day of birth.
In Panel B, we report the corresponding Two-Sample Two-Stage Least Squares estimates for the impact of attending a more academic middle school on these labor market outcomes, by dividing the reduced form estimates in Panel A by the first stage of 0.039 (Table 3, Panel A, column (3)). Standard errors are computed using the Delta Method, see footnote 20 for details.
Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level. Number of observations refer to the number of workers.
Source : Reduced Form: Social Security Data, 1975-2006. First Stage: Microcensus, selected years 1976 to 1987.

A Model Appendix

A.1 School Type Attendance and Productivity

In period 1, expected productivity depends on month of birth and students born in June are Δ less productive on average than students born in July. In periods 2 and 3, the initial advantage resulting from delayed school entry has fully disappeared, meaning that productivity no longer depends on birth month. The teaching technology in the two types of schools is linear in the student's ability a, and given by $\alpha^j + \beta^j a$, where the subscript j denotes the school type ($j = L, H$), and $\alpha^L > \alpha^H$ and $\beta^L < \beta^H$. q_t^j denotes the quality of teachers in period t ($t = 1, 2$) and school type j. $\hat{\Pi}_t^j$ denotes parental expectations about the average ability of peers in each school type in period t. To contrast parental expectations from the average ability which is realized in equilibrium, we denote the latter by Π_t^j. The effects of teachers and peers whom students are exposed to in period t on productivity in period k are denoted by δ_t^k and γ_t^k, respectively. The academic cost of switching from the low to the high school type at the end of period 1 is denoted by c.

-Table A1 here-

At the beginning of period 1, parents receive a noisy signal about their child's ability, denoted by θ, expressed as $\theta_i = a_i + \varepsilon_i$ for children born in July and $\theta_i = a_i - \Delta + \varepsilon_i$ for children born in June. Parents use this signal to update their own beliefs about their child's ability; updated beliefs are denoted here by $\hat{a} = E[a|\theta]$. Assuming that a_i and ε_i are normally distributed with mean μ and variance σ_a^2 and with mean zero and variance σ_ε^2, $E[a|\theta] = \hat{a} = \frac{\sigma_a^2\theta + \sigma_\varepsilon^2\mu}{\sigma_a^2 + \sigma_\varepsilon^2}$ for children born in July, and $E[a|\theta] = \hat{a} = \frac{\sigma_a^2(\theta+\Delta) + \sigma_\varepsilon^2\mu}{\sigma_a^2 + \sigma_\varepsilon^2}$ for children born in June. We use $G(\hat{a})$ and $F(a|\hat{a})$, respectively, to denote the cumulative distribution functions of the updated ability $\hat{a}$ and the true ability a, conditional on $\hat{a}$. $G(\hat{a})$ is normally distributed with mean μ and variance $\frac{\sigma_a^4}{\sigma_a^2+\sigma_\varepsilon^2}$. $F(a|\hat{a})$ is normally distributed with mean $\hat{a}$ and variance $\frac{\sigma_a^2\sigma_\varepsilon^2}{\sigma_a^2+\sigma_\varepsilon^2}$.

A.2 Equilibrium and Proof of Proposition 1

This section describes our model's equilibrium and presents the proof of Proposition 1. We begin with school type selection in period 2 when ability is fully known. A student who attended an L school in period 1 prefers an L over an H school in period 2 if

$$\underbrace{\alpha^L + \beta^L a + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_1^L + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L}_{\text{flow utility in period 2, low type in period 2}} + \underbrace{\alpha^L + \beta^L a + \gamma_1^3 \hat{\Pi}_1^L + \delta_1^3 q_1^L + \gamma_2^3 \hat{\Pi}_2^L + \delta_2^3 q_2^L}_{\text{flow utility in period 3, low type in period 2}} >$$

$$\underbrace{\alpha^H + \beta^H a + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_2^L + \gamma_2^2 \hat{\Pi}_2^H + \delta_2^2 q_2^H}_{\text{flow utility in period 2, high type in period 2}} + \underbrace{\alpha^H + \beta^H a + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_2^L + \gamma_2^2 \hat{\Pi}_2^H + \delta_2^2 q_2^H - c}_{\text{flow utility in period 3, high type in period 2}}$$

Since the teaching technology in an H school is more sensitive to ability than in an L school (i.e. $\beta^H > \beta^L$), student utility in an H school is increasing in ability at a faster rate than in an L school. Hence, there is an ability threshold a_L^* such that all students whose ability is identified as below a_L^* sort into L schools, while all students whose ability is identified above a_L^* sort into H schools.

Next, we consider a student who attended an H school in period 1. Her decision problem is the same as that above, except for there being no cost of switching from an L to an H school. It then follows that the ability threshold at which this student is indifferent between the two school types, a_H^*, is below the threshold at which the student who attended an L school in period 1 is indifferent between the two school types. It is then easy to show that $a_L^* = a_H^* + \frac{c}{2(\beta^H - \beta^L)}$.

As regards school type selection in period 1 when student ability is uncertain, in this period, parents whose child is born in July and has the expected ability $\widehat{a}$ send their child to an H school if

$$\underbrace{\alpha^H + \beta^H \hat{a} + \gamma_1^1 \hat{\Pi}_1^H + \delta_1^1 q_1^H}_{\text{first period, high type}} + \underbrace{F(a_H^*|\hat{a})(\alpha^L + \beta^L \frac{\int_{-\infty}^{a_H^*} adF(a|\hat{a})}{F(a_H^*|\hat{a})} + \gamma_1^2 \hat{\Pi}_1^H + \delta_1^2 q_1^H + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L)}_{\text{2nd period, downgrade to low type}} +$$

$$\underbrace{(1 - F(a_H^*|\hat{a}))(\alpha^H + \beta^H \frac{\int_{a_H^*}^{\infty} adF(a|\hat{a})}{1 - F(a_H^*|\hat{a})} + \gamma_1^2 \hat{\Pi}_1^H + \delta_1^2 q_1^H + \gamma_2^2 \hat{\Pi}_2^H + \delta_2^2 q_2^H)}_{\text{2nd period, stay in high type}} +$$

$$\underbrace{F(a_H^*|\hat{a})(\alpha^L + \beta^L \frac{\int_{-\infty}^{a_H^*} adF(a|\hat{a})}{F(a_H^*|\hat{a})} + \gamma_1^3 \hat{\Pi}_1^H + \delta_1^3 q_1^H + \gamma_2^3 \hat{\Pi}_2^L + \delta_2^3 q_2^L)}_{\text{3rd period, downgrade to low type}} +$$

$$\underbrace{(1 - F(a_H^*|\hat{a}))(\alpha^H + \beta^H \frac{\int_{a_H^*}^{\infty} adF(a|\hat{a})}{1 - F(a_H^*|\hat{a})} + \gamma_1^3 \hat{\Pi}_1^H + \delta_1^3 q_1^H + \gamma_2^3 \hat{\Pi}_2^H + \delta_2^3 q_2^H)}_{\text{3rd period, stay in high type}} \qquad >$$

$$\underbrace{\alpha^L + \beta^L \hat{a} + \gamma_1^1 \hat{\Pi}_1^L + \delta_1^1 q_1^L}_{\text{first period, low type}} + \underbrace{F(a_L^*|\hat{a})(\alpha^L + \beta^L \frac{\int_{-\infty}^{a_L^*} adF(a|\hat{a})}{F(a_L^*|\hat{a})} + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_1^L + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L)}_{\text{2nd period, stay in low type}} +$$

$$\underbrace{(1 - F(a_L^*|\hat{a}))(\alpha^H + \beta^H \frac{\int_{a_L^*}^{\infty} adF(a|\hat{a})}{1 - F(a_L^*|\hat{a})} + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_1^L + \gamma_2^2 \hat{\Pi}_2^H + \delta_2^2 q_2^H - c)}_{\text{2nd period, upgrade to high type}} +$$

$$\underbrace{F(a_L^*|\hat{a})(\alpha^L + \beta^L \frac{\int_{-\infty}^{a_L^*} adF(a|\hat{a})}{F(a_L^*|\hat{a})} + \gamma_1^3 \hat{\Pi}_1^L + \delta_1^3 q_1^L + \gamma_2^3 \hat{\Pi}_2^L + \delta_2^3 q_2^L)}_{\text{3rd period, stay in low type}} +$$

$$\underbrace{(1 - F(a_L^*|\hat{a}))(\alpha^H + \beta^H \frac{\int_{a_L^*}^{\infty} adF(a|\hat{a})}{1 - F(a_L^*|\hat{a})} + \gamma_1^3 \hat{\Pi}_1^L + \delta_1^3 q_1^L + \gamma_2^3 \hat{\Pi}_2^H + \delta_2^3 q_2^H - c)}_{\text{3rd period, stay in high type}}.$$

Since the child's utility in an H school is increasing in her expected ability at a faster rate than her utility in an L school, there exists an ability threshold $\hat{a}^*_{\text{July}}$ such that, in period 1, all children whose expected ability is below this threshold attend an L school while all children whose expected ability is above the threshold attend an H school.

For a child who is born in June, the parents' decision problem is similar except that the child's utility in the first period in an H and L school is replaced by $\alpha^H + \beta^H(\hat{a} - \Delta) + \gamma_1^0 \hat{\Pi}_1^H + \delta_1^0 q_1^H$ and $\alpha^L + \beta^L(\hat{a} - \Delta) + \gamma_1^0 \hat{\Pi}_1^L + \delta_1^0 q_1^L$, respectively. It then follows that the expected ability threshold at which children born in June are indifferent between attending an L or an H school in the first period, $\hat{a}^*_{\text{June}}$, exceeds the threshold at which children born in July are indifferent; that is, $\hat{a}^*_{\text{June}} > \hat{a}^*_{\text{July}}$. Consequently, children who are shifted from an L to an H school because of a July (rather than a June) birth date are children in the expected ability range $[\hat{a}^*_{\text{July}}, \hat{a}^*_{\text{June}}]$, as given in Proposition 1.

In equilibrium, parental expectations of peer ability in each school type and period must correspond to the realized average ability of children in each school type and period; that is, $\hat{\Pi}_t^j = \Pi_t^j$. These realized expectations are determined by the ability thresholds $\hat{a}^*_{\text{July,}}$ $\hat{a}^*_{\text{June}}$, a_L^*, and $a_{H.}^*$. We illustrate the computation of Π_t^j using the average ability of children who attend an L school in period 2 (Π_2^L), who fall into four groups: children born in July or June who attended an L school in period 1 (who have a measure of $\int_{-\infty}^{\hat{a}^*_{\text{June}}} F(a_L^*|\widehat{a})dG(\widehat{a})$ and $\int_{-\infty}^{\hat{a}^*_{\text{July}}} F(a_L^*|\widehat{a})dG(\widehat{a})$, respectively), and children born in July or June who attended the high school type in period 1 (who have a measure of $\int_{\hat{a}^*_{\text{June}}}^{\infty} F(a_H^*|\widehat{a})dG(\widehat{a})$ and $\int_{\hat{a}^*_{\text{July}}}^{\infty} F(a_H^*|\widehat{a})dG(\widehat{a})$, respectively). Thus, for example, the average ability of children born in June who attended an H school in period 1 equals $\int\limits_{\hat{a}^*_{\text{June}}}^{\infty} \int\limits_{-\infty}^{a^*_H} \frac{a dF(a|\widehat{a})}{F(a^*_H|\widehat{a})} dG(\widehat{a})/(1-G(\widehat{a}^*_{\text{June}}))$. Hence, Π_2^L equals:

$$\Pi_2^L = \frac{\int\limits_{-\infty}^{\hat{a}^*_{\text{June}}} F(a^*_L|\widehat{a})dG(\widehat{a}) \frac{\int\limits_{-\infty}^{\hat{a}^*_{\text{June}}} \int\limits_{-\infty}^{a^*_L} \frac{a dF(a|\widehat{a})}{F(a^*_L|\widehat{a})} dG(\widehat{a})}{G(\widehat{a}^*_{\text{June}})} + \int\limits_{\hat{a}^*_{\text{June}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a}) \frac{\int\limits_{\hat{a}^*_{\text{June}}}^{\infty} \int\limits_{-\infty}^{a^*_H} \frac{a dF(a|\widehat{a})}{F(a^*_H|\widehat{a})} dG(\widehat{a})}{1-G(\widehat{a}^*_{\text{June}})}}{\int\limits_{-\infty}^{\hat{a}^*_{\text{June}}} F(a^*_L|\widehat{a})dG(\widehat{a}) + \int\limits_{\hat{a}^*_{\text{June}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a}) + \int\limits_{-\infty}^{\hat{a}^*_{\text{July}}} F(a^*_L|\widehat{a})dG(\widehat{a}) + \int\limits_{\hat{a}^*_{\text{July}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a})} +$$

$$\frac{\int\limits_{-\infty}^{\hat{a}^*_{\text{July}}} F(a^*_L|\widehat{a})dG(\widehat{a}) \frac{\int\limits_{-\infty}^{\hat{a}^*_{\text{July}}} \int\limits_{-\infty}^{a^*_L} \frac{a dF(a|\widehat{a})}{F(a^*_L|\widehat{a})} dG(\widehat{a})}{G(\widehat{a}^*_{\text{June}})} + \int\limits_{\hat{a}^*_{\text{July}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a}) \frac{\int\limits_{\hat{a}^*_{\text{July}}}^{\infty} \int\limits_{-\infty}^{a^*_H} \frac{a dF(a|\widehat{a})}{F(a^*_H|\widehat{a})} dG(\widehat{a})}{1-G(\widehat{a}^*_{\text{July}})}}{\int\limits_{-\infty}^{\hat{a}^*_{\text{June}}} F(a^*_L|\widehat{a})dG(\widehat{a}) + \int\limits_{\hat{a}^*_{\text{June}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a}) + \int\limits_{-\infty}^{\hat{a}^*_{\text{July}}} F(a^*_L|\widehat{a})dG(\widehat{a}) + \int\limits_{\hat{a}^*_{\text{July}}}^{\infty} F(a^*_H|\widehat{a})dG(\widehat{a})}$$

A.3 Proposition 2

Next, we derive the local average treatment effect of attending an H school in period 1 (denoted by H_1) on the probability of attending an H school in period 2 (denoted by H_2). Dividing the reduced-form effect (i.e. $\Pr(H_2|Z_i = \text{July}) - \Pr(H_2|Z_i = \text{June})$) by the first-stage effect (i.e. $\Pr(H_1|Z_i = \text{July}) - \Pr(H_1|Z_i = \text{June})$), which corresponds to equation (4) in Section 3.1, yields

$$\frac{\Pr(H_2|Z_i = \text{July}) - \Pr(H_2|Z_i = \text{June})}{\Pr(H_1|\text{July}) - \Pr(H_1|\text{June})} = \frac{\int_{\hat{a}^*_{\text{July}}}^{\hat{a}^*_{\text{June}}} (F(a^*_L|\hat{a}) - F(a^*_H|\hat{a}))dG(\hat{a})}{G(\hat{a}^*_{\text{June}}) - G(\hat{a}^*_{\text{July}})}$$

This effect is highly dependent on the difference between the thresholds a^*_L and a^*_H, which in turn depend on the switching costs c. Ceteris paribus, the larger the switching costs, the larger the impact of school type attended in period 1 on school type attended in period 2. If there are no switching costs, then $a^*_L = a^*_H$ and school type attended in period 1 has no impact on the school type completed. The intuitive argument, however, considers only the direct effect of switching costs on school type completed through its effect on a_L^* and a_H^* and ignores that switching costs also affect the ability thresholds $\hat{a}^*_{\text{July}}$ and $\hat{a}^*_{\text{June}}$. We therefore confirmed the positive relationship between switching costs and the impact of school type attended in period 1 on school type attended in period 2 by running extensive simulations that take these indirect effects into account.

A.4 Propositions 3 and 4

To assess the local average treatment effect of attending an H (rather than an L) school in period 1 on wages (assumed to be equal to productivity) in period 3, we divide the reduced-form effect (i.e. $\Pr(w_3|Z_i = \text{July}) - \Pr(w_3|Z_i = \text{June})$) by the first-stage effect to yield

$$\frac{E(w_3|Z_i = \text{July}) - E(w_3|Z_i = \text{June})}{\Pr(H_1|Z_i = \text{July}) - \Pr(H_1|Z_i = \text{June})} = \underbrace{\gamma_1^3(\hat{\Pi}_1^H - \hat{\Pi}_1^L)}_{\text{direct effect, peers}} + \underbrace{\delta_1^3(q_1^H - q_1^L)}_{\text{direct effect, teachers}}$$
$$+ \underbrace{\frac{\int_{\hat{a}^*_{\text{July}}}^{\hat{a}^*_{\text{June}}}(1 - F(a_L^*|\hat{a}))c\,dG(\hat{a})}{G(\hat{a}^*_{\text{June}}) - G(\hat{a}^*_{\text{July}})}}_{\text{direct effect, switching costs}} + \tag{6}$$
$$\underbrace{\frac{\int_{\hat{a}^*_{\text{July}}}^{\hat{a}^*_{\text{June}}}(F(a_L^*|\hat{a}) - F(a_H^*|\hat{a}))(\alpha^H - \alpha^L + \delta_2^3(q_2^H - q_2^L) + \gamma_2^3(\hat{\Pi}_2^H - \hat{\Pi}_2^L)) + (\beta^H - \beta^L)\int_{\hat{a}_H^*}^{\hat{a}_L^*} a\,dF(a|\hat{a}))dG(\hat{a})}{G(\hat{a}^*_{\text{June}}) - G(\hat{a}^*_{\text{July}})}}_{\text{indirect effect through school type in period 2}}$$

When school type attended in period 2 is held constant, school type attended in period 1 affects wages directly-as captured by the first three terms in equation (6)-reflecting both the exposure to better peers and better teachers in period 1 and the costs of switching from an L to an H school. School type attended in period 1 also affects wages but indirectly, through the type of school attended in period 2. This indirect effect, captured by the fourth term in equation (6), becomes larger, the larger the impact of school type attended in period

1 on school type attended in period 2 (i.e., a_L^* versus a_H^*), and the larger the effect of peers and teachers in period 2 on wages in period 3 (i.e., δ_2^3 and γ_2^3).[30] If switching costs are zero, then $a_L^* = a_H^*$ and the indirect effect disappears. In this case, the local average treatment effect of attending an H school in period 1 on wages in period 3 reduces to the direct effect, again reflecting the exposure to better peers and teachers in period 1; that is, equation (6) becomes $\gamma_1^3(\hat{\Pi}_1^H - \hat{\Pi}_1^L) + \delta_1^3(q_1^H - q_1^L)$. This outcome leads to Propositions 3 and 4.

B Data Appendix: School Census for Bavaria and Hesse

Because this school census contains no direct information on degrees obtained, we proxy graduation from a general H school by first counting the number of students who, by age 17, 18, and onward, had ever attended a general H school in grade 13, making sure not to double count students who repeated a grade. Because students drop out of the census on leaving school, we divide this number by the total number of children born in the month, year, and state. We proxy graduation from a specialized H school as the ratio of the number of students who, by age 22, had ever attended a specialized H school in grade 12 to the total number of children born in the month, year, and state:

$$\widehat{p} = \frac{\#\text{ ever reached grade 13 (12)}}{\#\text{ births}}$$

The standard error of this share is estimated as the square root of

$$Var\left(\widehat{p}\right) = \frac{\widehat{p}\left(1-\widehat{p}\right)}{\#\text{ births}}$$

The variance of the difference in the shares between July and June born children is given by

$$Var\left(\widehat{p}_{July} - \widehat{p}_{July}\right) = \frac{\widehat{p}_{July}\left(1-\widehat{p}_{July}\right)}{\#\text{ births}_{\text{July}}} + \frac{\widehat{p}_{June}\left(1-\widehat{p}_{June}\right)}{\#\text{ births}_{\text{June}}}$$

[30] Again, this argument considers direct effects only, but again we confirmed that this relation holds in equilibrium using extensive simulations.

The share of students graduating from a specialized *or* general H school is computed as the ratio of the students who ever reached grade 12 in a specialized or general H school, over the number of births in question. Standard errors are calculated in the same way as shown above.

C Additional Results

-Table A2 here-

-Table A3 here-

C.1 Month of Birth and School Type Attendance in Middle School (School Census)

Figure A1 plots the share of students who at age 14 were attending an H (Panel A) or an L (Panel B) school against birth month based on the school census data. The figure shows a clear discontinuity in school type selection around the school entry cut-off date: children born in July are 5.2 percentage points more likely to attend an H school and 4.2 percentage points less likely to attend an L school than children born in June, for a total effect of 9.4 percentage points.

-Figure A1 here-

C.2 Comparison of First Stage for Recent Cohorts (School Census) and Older Cohorts (Microcensus)

A comparison of the estimates in Panel A and Panel B of Table 3 clearly indicates a stronger impact of birth date on school type selection for the recent birth cohorts in Hesse and Bavaria than for the older birth cohorts in West Germany, even taking into account the measurement error in the school entry cut-off (9.4% versus 5.0%)[31]. We next investigate whether this

[31] This number is computed as follows. By dividing the first-stage estimate based on children born in June or July of 9.4% (Table 3, Panel B, row (i), column (3)) with the first-stage estimate based on children born earlier or later during the year of 7.2% (Table 3, Panel B, row (ii), column (3)), we can derive an adjustment factor of 1.29 (0.072/0.094). Multiplying the estimates in Panel A by this adjustment factor indicates that, for the 1961 to 1976 birth cohorts, children born in July are 5 percentage points more likely to attend a more academic middle school than students born in June.

difference can be explainable by the older cohorts' being less compliant with the school entry cut-off rule than the more recent birth cohorts. Because Bavaria does not record the year of school entry, meaning we cannot compute its compliance rate, we restrict the analysis to Hesse, whose first-stage estimate is 7.3 percent (compared to a 9.4 percent first-stage estimate for Hesse and Bavaria combined). The compliance rate for the recent birth cohorts in Hesse is 33.5 percent, whereas that for the older birth cohorts (here proxied by the 1963 birth cohort in the 1970 Census for West Germany) is only 21.9 percent.[32] To estimate the impact of relative age (i.e., being one year older at school entry) on school type selection, we divide the effect of birth month on school type selection by the effect of birth month on age of school entry and obtain an effect of 23.0 percentage points (0.050/0.219) for the older birth cohorts. For the younger birth cohorts, we estimate a very similar number: 21.8 percentage points (0.078/0.335). Hence, the lower first-stage results for the 1961-1976 cohorts in West Germany than for the recent cohorts in Bavaria and Hesse are wholly explained by the older cohorts' being less compliant with the school entry age cut-off rule than the more recent cohorts.

C.3 Month of Birth and Wages: Robustness Checks

In Table A4, we report a number of robustness checks on the effect of birth month on wages. All estimates are based on a comparison of individuals born in June or July. The results in columns (1) to (2) are for the baseline sample of men born between 1961 and 1976 aged 30 and above. For comparative purposes, we report our baseline estimate in column (1), where returns to potential experience are estimated using OLS and a functional form of a 4th order polynomial in potential experience is imposed. In column (2), we allow for a more flexible functional form and use a fully flexible set of dummy variables to model the returns

[32] Children in the 1963 cohort who were born July through December should not yet have been in school when the census was carried out in spring 1970, whereas children born January through June should have already been attending school. However, regressing an indicator for school attendance on an indicator for being born in July or June results in a coefficient of only 0.219; that is, a compliance with the school entry cut-off rule of 21.9 percent.

to potential experience. Results are very similar.

In columns (3) and (4), we restrict the sample to individuals aged 35 years and older, and report findings both for raw and experience-adjusted log-wages. Our conclusions are unchanged. In columns (5) and (6) we further restrict the sample to individuals aged 40 years and older whose wages are flat with respect to potential experience. Indeed, for these individuals, both the raw and experience-adjusted wage differentials are very small and not statistically significant from zero.

We therefore conclude that the reduced-form estimates of the impact of birth month on experience-adjusted wages are very small and robust to different specifications.

-Table A4 here-

Table A1: School Type Attendance and Productivity

Period 1, July (Schooling Period)	
low in 1	$\alpha^L + \beta^L \hat{a} + \gamma_1^1 \hat{\Pi}_1^L + \delta_1^1 q_1^L$
high in 1	$\alpha^H + \beta^H \hat{a} + \gamma_1^1 \hat{\Pi}_1^H + \delta_1^1 q_1^H$
Period 1, June (Schooling Period)	
low in 1	$\alpha^L + \beta^L (\hat{a} - \Delta) + \gamma_1^1 \hat{\Pi}_1^L + \delta_1^1 q_1^L$
high in 1	$\alpha^H + \beta^H (\hat{a} - \Delta) + \gamma_1^1 \hat{\Pi}_1^H + \delta_1^1 q_1^H$
Period 2, July and June (Schooling Period)	
low in 2, low in 1	$\alpha^L + \beta^L a + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_1^L$
low in 2, high in 1	$\alpha^H + \beta^H a + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L + \gamma_1^2 \hat{\Pi}_1^H + \delta_1^2 q_1^H$
high in 2, low in 1	$\alpha^H + \beta^H a + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L + \gamma_1^2 \hat{\Pi}_1^L + \delta_1^2 q_1^L - c$
high in 2, high in 1	$\alpha^H + \beta^H a + \gamma_2^2 \hat{\Pi}_2^L + \delta_2^2 q_2^L + \gamma_1^2 \hat{\Pi}_1^H + \delta_1^2 q_1^H$
Period 3, July and June (Working period)	
low in 2, low in 1	$\alpha^L + \beta^L a + \gamma_2^3 \hat{\Pi}_2^L + \delta_2^3 q_2^L + \gamma_1^3 \hat{\Pi}_1^L + \delta_1^3 q_1^L$
low in 2, high in 1	$\alpha^H + \beta^H a + \gamma_2^3 \hat{\Pi}_2^H + \delta_2^3 q_2^H + \gamma_1^3 \hat{\Pi}_1^H + \delta_1^3 q_1^H$
high in 2, low in 1	$\alpha^H + \beta^H a + \gamma_2^3 \hat{\Pi}_2^L + \delta_2^3 q_2^L + \gamma_1^3 \hat{\Pi}_1^L + \delta_1^3 q_1^L - c$
high in 2, high in 1	$\alpha^H + \beta^H a + \gamma_2^3 \hat{\Pi}_2^H + \delta_2^3 q_2^H + \gamma_1^3 \hat{\Pi}_1^H + \delta_1^3 q_1^H$

Note: The table reports the student's productivity in each period, depending on the type of school the student attended in period 1 and 2.

Table A2: Month of Birth and Family Background Characteristics

	Father	Mother
Age	-0.14	-0.098
(s.e.)	(0.13)	(0.11)
College/University	-0.011	-0.004
(s.e.)	(0.01)	(0.01)
At least Apprenticeship	-0.007	-0.004
(s.e.)	(0.01)	(0.01)
N	8,616	8,616

Note: The table reports the difference in age at birth (in years), college and university education, and apprenticeship education of fathers and mothers of children who were born in June or July, respectively. The analysis is based on the German Microcensus for the year 2005, and the sample consists of all children still living with their parents who were born between 1991 and 2004 and thus were less than 14 years old at the time of the survey.
Source: Microcensus 2005.

Table A3: The Stability of the First Stage Across Grades

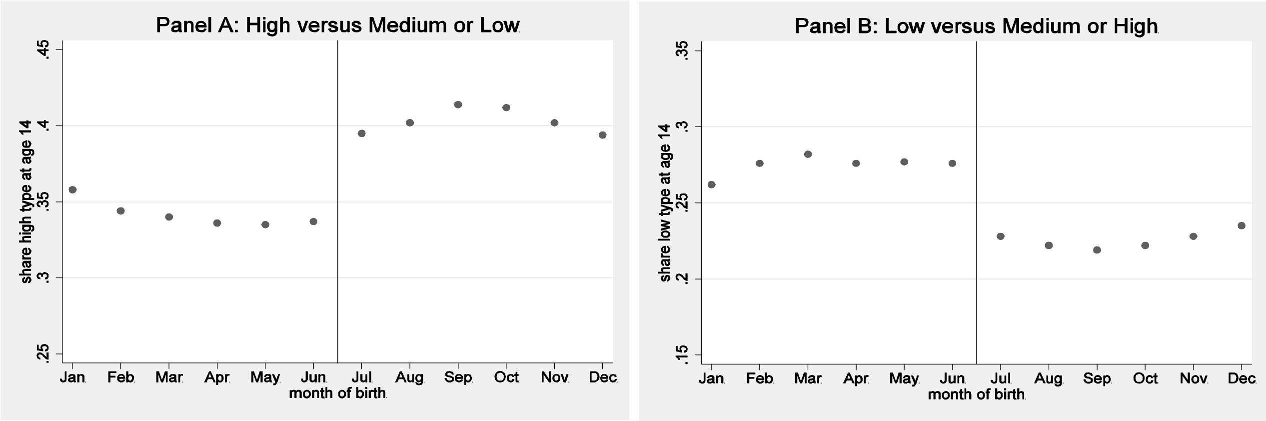
	(1) High versus Medium or Low	(2) High or Medium versus Low	(3) Sum
Grade 5	0.063	0.052	0.115
N=163,138	(0.002)**	(0.002)**	(0.004)**
Grade 6	0.063	0.052	0.115
N=158,781	(0.002)**	(0.002)**	(0.004)**
Grade 7	0.061	0.051	0.112
N=160,354	(0.002)**	(0.002)**	(0.004)**
Grade 8	0.057	0.048	0.105
N=162,491	(0.002)**	(0.002)**	(0.004)**
Grade 9	0.056	0.049	0.105
N=164,200	(0.002)**	(0.002)**	(0.004)**

Note: The table reports the difference in the share of students who attend a H versus a M or L school (column (1)), and in the share of students who attend a L versus a H or M school (column (2)) between students born in July and June, through grade 5 to grade 9. In column (3), the dependent variable is coded 2 if the student is in a H school, 1 if she is in a M school and 0 if she is in a L school, corresponding to the sum of the coefficients in the first two columns. Robust standard errors in parentheses.

Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.

Source: School Census for Bavaria and Hesse, 2002-2009.

Figure A1: School Type Choice at Age 14 and Month of Birth



Note : The figures plot the share of students attending a high (H) (Panel A) and low (L) (Panel B) middle school type at age 14 against the month of birth. Results refer to birth cohorts 1988 to 1994.

Source : School Census for Bavaria and Hesse, 2002-2009.

Table A4: Date of Birth and Wages: Robustness Checks (Men)

	Birth Cohorts 1961-1976, age 30 and up		Birth Cohorts, 1961-1971, age 35 and up		Birth Cohorts 1961-1964, age 40 and up	
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline OLS, pol. 4	Alt. functional form OLS, dummies	Raw	Exerience-adjusted OLS, pol. 4	Raw	Exp.-adjusted OLS, pol. 4
July vs June						
Coeff.	-0.0009	-0.0009	-0.0032	-0.0007	0.0000	-0.0002
(s.e.)	(0.0008)	(0.0008)	(0.0011)**	(0.0011)	(0.0018)	(0.0019)

Note : The table reports various robustness checks for the impact of birth month on log-wages. Columns (1) to (2) refer to birth cohorts 1961 to 1976 and individuals 30 and older. For comparison, we report our baseline estimate in column (1) where returns to potential experience are estimated using OLS and a functional form of a 4th order polynomial in potential experience is imposed. In column (2), we relax the functional form assumption and include a full set of dummy variables for potential experience instead. Columns (3) and (4) refer to birth cohorts 1961 to 1971 and individuals 35 and older, while columns (4) and (6) refer to birth cohorts 1961 to 1964 and individuals 40 and older. We first report the raw and then the experience-adjusted wage differential, where returns to potential experience are, as in our baseline specification, estimated using OLS and a functional form of a 4th order polynomial in potential experience is assumed. Standard errors clustered at the individual level in parentheses.
Coefficients with * are statistically significant at the 5 percent level, those with ** at the 1 percent level.
Source : Social Security Data, Men, 1975-2006.