

Early School Exposure, Test Scores, and Noncognitive Outcomes[†]

By THOMAS CORNELISSEN AND CHRISTIAN DUSTMANN*

We estimate the effects of receiving additional schooling before age five on cognitive and noncognitive outcomes, exploiting unique school entry rules in England that cause variation in the age at school entry and the effective length of the first school year, and combining survey data with administrative school records up to six years after exposure. We find significant effects on both cognitive and noncognitive outcomes at ages five and seven, particularly so for boys with a disadvantaged parental background. At age 11, effects on cognitive outcomes have disappeared, while there is still evidence for effects on noncognitive outcomes. (JEL I21, J13)

How early in life should formal schooling start? Some argue that interventions such as early preschool attendance are extremely effective for skill development (see Cunha and Heckman 2007, Cunha et al. 2006, Heckman 2008, and the comprehensive review in Currie and Almond 2011). Based on such evidence, President Barack Obama proposed in his 2013 State of the Union Address “to make high-quality preschool available to every single child in America.”¹ Yet generalizable and conclusive evidence based on a clean design to identify the causal effect of universal early schooling remains scarce. While studies analyzing well-designed randomized programs such as the Perry Preschool Project, the Carolina Abecedarian Project, and the more wide-scale Head Start program provide evidence of positive effects of early schooling in the United States,² they do so only in relatively specific contexts. That is, not only are these programs all targeted at disadvantaged children, but they include both schooling and a mix of interventions (e.g., home visits in the

*Cornelissen: Department of Economics, University of York, Heslington, York YO10 5DD, United Kingdom (email: thomas.cornelissen@york.ac.uk); Dustmann: Department of Economics, University College London and CREAM, 30 Gordon Street, London WC1H 0AX, United Kingdom (email: c.dustmann@ucl.ac.uk). Kate Ho was coeditor for this article. Dustmann acknowledges funding through the ERC Advanced Grant 323992-DMEA and by the German Research Foundation DFG (DU1024/1-1). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <https://doi.org/10.1257/pol.20170641> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

¹See <http://www.whitehouse.gov/the-press-office/2013/02/12/remarks-president-state-union-address> for the 2013 State of the Union Address.

²The Perry Preschool Project and the Carolina Abecedarian Project have been extensively evaluated on various outcomes, ranging from short-term child development to labor market and other long-term outcomes (Blau and Currie 2006; Currie 2001; Anderson 2008; Heckman et al. 2010a,b; Masse and Barnett 2002; and Schweinhart et al. 2005). Head Start has been analyzed using quasi-experimental (Anderson, Foster, and Frisvold 2010; Currie and Thomas 1995, 1999; Garces, Thomas, and Currie 2002; Carneiro and Ginja 2014; and Ludwig and Miller 2007) and experimental (US Department of Health and

Perry Preschool Project and interventions to improve health, nutrition, and parent involvement in the Head Start program). Evidence of the effects of exposure to early schooling in more universal school environments is much rarer, with few studies employing randomized designs. In the extant studies of general pre-primary education in several countries, participation is voluntary and enrollment rates in early schooling are in the range of 60–70 percent (Berlinski, Galiani, and Manacorda 2008; Berlinski, Galiani, and Gertler 2009; Cascio 2009; Cascio and Schanzenbach 2013; Gormley and Gayer 2005; and Magnuson, Ruhm, and Waldfogel 2007).³

In this paper, we estimate the causal effects of increased exposure to *universal* early schooling, exploiting a unique variation in the rules of entry into the first year of elementary school in England, where the enrollment rate is almost complete. Our treatment affects children four to five years of age, which is the age range that would be affected by the introduction of universal preschool programs in the United States, where children currently enroll in the first year of universal schooling (kindergarten) only at ages five to six. Given this universality and the age range of affected children, our quasi-experiment closely simulates the case of extending preschool programs in the United States.

Our identification is based on school entry regulations that stipulate up to three different entry dates into the *same* academic year (i.e., school entry in the first, second, or third term of the academic year), to which children are assigned by birth month cutoff dates that vary regionally. This variation allows us to compare children *in the same year cohort and grade* (holding advancement in the school curriculum constant) who are the *same (absolute and relative) age at testing*, but who have spent different times in the first grade because of having entered at different dates. This comparison therefore identifies the effect of increasing exposure to elementary education through an earlier school entry age. The variation further allows conditioning on birth month fixed effects (to adjust for birth month effects and age at the test) and local authority fixed effects (to control for region characteristics that might be correlated with the school entry rules and also affect child outcomes).

The school entry rules in most other countries, by contrast, induce no such variation in the length of exposure to early schooling among children in the same grade who sit the test at the same time because there is only one possible school starting date per academic year.⁴ Most papers exploiting school-entry cutoffs identify the effect

Human Services 2010) research designs. See also Elango et al. (2016) for a synthesis of the literature on small- and large-scale, targeted, and universal early childcare programs.

³These studies look at effects of early education at school entry (e.g., kindergarten class in the United States) or immediately before school entry (e.g., prekindergarten in the United States). There is a much larger literature on early childhood programs that often cover children as young as two to three years of age (see, e.g., Baker, Gruber, and Milligan 2008; Bernal and Keane 2010, 2011; Blanden et al. 2014; Cornelissen et al. 2018; Datta Gupta and Simonsen 2010; Fort, Ichino, and Zanella 2016; Havnes and Mogstad 2011, 2015; and Loeb et al. 2007). Baker (2011), Elango et al. (2016), and Ruhm and Waldfogel (2012) provide overview articles of the literature on the effects of early childhood education. Although some studies focus on center-based childcare just before school entry (e.g., Drange, Havnes, and Sandsør 2016), the curriculum is far more play-based than in our setting, in which we investigate the onset of formal schooling.

⁴With only one uniform school entry date per academic year, variation in exposure can only be generated by measuring the outcome at different values of school exposure (Gormley and Gayer 2005; Dee and Sievertsen 2018; Cascio and Lewis 2006; Black, Devereux, and Salvanes 2011; and Carlsson et al. 2015), such as comparing children who at the same age differ by one year in their exposure to schooling because they were assigned to different academic year cohorts. While this in principle identifies an exposure effect, such children differ strongly in their

of changing both age-at-entry and age-at-test, while keeping exposure constant (see, e.g., Bedard and Dhuey 2006; Datar 2006; Elder and Lubotsky 2009; Fertig and Kluge 2005; Fredriksson and Öckert 2014; Landersø, Skyt Nielsen, and Simonsen 2017; McEwan and Shapiro 2008; Mühlenweg and Puhani 2010; and Puhani and Weber 2007). To see why this is, suppose that exposure to schooling (EXP), the age-at-test ($AGET$), and age-at-school-entry ($AGEE$) have separate effects on an outcome y according to $y = \beta_0 + \beta_1 EXP + \beta_2 AGET + \beta_3 AGEE + e$. Given the identity $AGET = AGEE + EXP$, any linear regression can at most include two of the three terms, and the coefficients on those terms pick up composite effects. Moreover, EXP is usually constant because the analysis is typically based on school test scores of children in the same grade, implying the same exposure to schooling. Including either $AGEE$ or $AGET$, while EXP is constant, identifies the composite effect $\beta_2 + \beta_3$ of being older at entry and being older at the time of the test, with no change in exposure, as illustrated in Figure 1, panel B.⁵ From this body of research we know that children who enter school at an older age and who are older at the test do better than their younger classmates. However, this insight has limited policy relevance because it is hard to see how any practical school entry policy could even out such age-related differences.

We, in contrast, exploit school-entry rules that allow entry in different terms of the academic year and thus cause variation in EXP , the length of the school year. By including EXP and holding $AGET$ constant, we are thus in a position to identify the composite effect $\beta_1 - \beta_3$, as illustrated in Figure 1, panel C. Hereafter, we refer to this effect as the exposure effect; that is, the effect of prolonged exposure to schooling (β_1) obtained by starting school earlier ($-\beta_3$), which implies less time spent in the childcare environment preceding school entry. It is precisely this effect that matters for the debate over whether early exposure to formal schooling should be increased because such increase can only be achieved by lowering the age at which children start school.⁶

We add to the existing literature in several important ways. First, we offer an unusually tight identification strategy for the effect of additional exposure to schooling obtained through an earlier school entry at the expense of time spent in the childcare environment that precedes school entry. In contrast to other studies, the variation we exploit allows us to identify this effect net of birth month effects and

relative age compared to their classmates, and in their progression through the school curriculum. The estimated parameter is thus confounded by these two factors.

⁵Some studies of school-entry age effects look at long-term outcomes measured after schooling is completed (Fredriksson and Öckert 2014; Black, Devereux, and Salvanes 2011; and Landersø, Skyt Nielsen, and Simonsen 2017). This breaks the collinearity between the age when the outcome is measured and the age at school entry, which allows controlling for the age at the observation of the outcome. This leads to the conceptually different effect of being younger (in absolute and relative terms) at school entry (and at all points throughout the school career), but having one more year of experience between the school leaving date and the date at which the outcome is measured.

⁶Two papers use a similar type of exogenous variation as we do but, unlike us, only identify reduced-form effects due to data limitations: Crawford, Dearden, and Meghir (2007) exploits the same school-entry rules as we do, and Leuven et al. (2010) uses unique features of Dutch school entry rules allowing children to enter school immediately after their fourth birthday (causing variation in the length of the first school year), combined with the timing of the summer holidays. Other related studies that explicitly focus on late schooling, such as schooling around the school-leaving age, include Del Bono and Galindo-Rueda (2006), Oreopoulos (2006), and Carlsson et al. (2015).

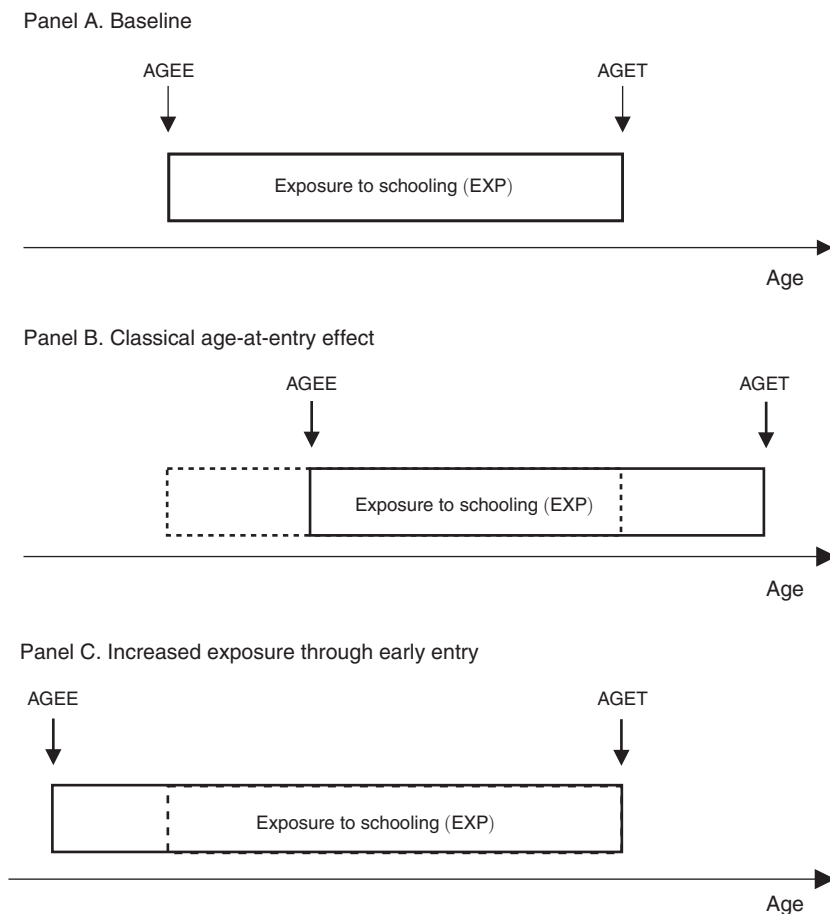


FIGURE 1. AGE-AT-ENTRY EFFECT AND EXPOSURE EFFECT

Notes: Moving from panel A to panel B illustrates the effect of varying the age-at-school-entry (AGEE) while holding exposure to schooling constant. This varies the age at the test (AGET) in the same way as the age at entry. This is the variation usually studied in the classical age-at-entry literature. Moving from panel A to panel C illustrates the variation in exposure to schooling generated by changing the age-at-entry but keeping the age-at-test constant. This variation is the one exploited in this study.

net of effects of absolute and relative age at test. This parameter is policy relevant and informs such debates as that on the expansion of public preschool programs. Our design also includes a one-sided noncompliance (illustrated below) that allows us to identify a treatment effect on the untreated (ATU); that is, the effect of expanding schooling on those who are not yet in school. This is a rare special case in which the local average treatment effect (LATE) identified by linear IV estimation has a clear interpretation and external validity.

Second, we identify this effect in a context of *universal* early schooling, not for a targeted intervention or for voluntary preschool attendance. We do so by exploiting variation in the length of schooling among enrolled children in a context in which the enrollment rate is 96 percent. Our results thus help inform the debate at what age formal

schooling should start. The starting age of formal schooling differs widely across countries, with the United Kingdom among the countries in which formal schooling starts the earliest (at age four to five). Yet, to date there exists little evidence on what the optimal starting age for formal schooling is. Third, we trace out the evolution of the effect over subsequent grades based on school test scores, as well as parental, teacher, and self-assessments taken at ages 5, 7, and 11. Fourth, after having established the overall effect, we conduct a subgroup analysis by gender interacted with socioeconomic background to reassess the hypothesis of Elder and Lubotsky (2009) that children from disadvantaged backgrounds benefit more from early schooling.

Finally, by combining administrative data with unique survey data, we are able to examine an unusually rich set of cognitive, noncognitive, and behavioral outcomes assessed by both parents and teachers, drawing at the same time on a very large number of observations, which turns out to be important for obtaining precise estimates at later ages. Analyzing the effect of early education on noncognitive skills is particularly interesting because these skills may have important long-term impacts (Chetty et al. 2011; Heckman, Pinto, and Savelyev 2013). However, the evidence on how early education and childcare affect noncognitive skills is not altogether clear: whereas some studies find positive effects of early education on noncognitive skills (Berlinski, Galiani, and Gertler 2009; Heckman, Pinto, and Savelyev 2013), others find negative effects (Baker, Gruber, and Milligan 2008; Loeb et al. 2007; Magnuson, Ruhm, and Waldfogel 2007) or report mixed results (Datta Gupta and Simonsen 2010).

For cognitive outcomes, we find that an additional month of exposure to early schooling before the age of 5 (holding age-at-test constant) increases test scores at the end of the first school year by approximately 6–9 percent of a standard deviation. This effect is smaller 2 years later at age 7, albeit still present and significant, but it largely disappears at age 11. We also show that the early test score effects are larger for low socioeconomic status (SES) than high SES boys (but not girls), closing the early achievement gap at age 7 between low and high SES boys by 60–80 percent of its initial magnitude. Even if the overall cognitive effects are temporary, closing early SES achievement gaps may have important implications, in particular in school systems in which early decisions about future school attendance are based on early test scores.

For noncognitive and behavioral outcomes, we find more persistent effects, at least up to age 11—the end of our observation window—and again evidence for stronger effects for low SES boys. Much in line with findings by Chetty et al. (2011) and Heckman, Pinto, and Savelyev (2013) for the STAR experiment and the Perry Preschool project, respectively, our analysis suggests noncognitive effects to be more persistent. We further explore reasons for why SES affects the effect of early schooling for boys but not for girls and conclude that, rather than low SES parents behaving differently according to the gender of their child, boys and girls seem to *respond* differently to low SES, a view that finds support in some strands of the developmental and child psychology literature.

The remainder of the paper is structured as follows. Section I provides information on the institutional background and data used in our analysis. Section II describes our empirical strategy and estimation procedure and clarifies the interpretation of the estimated parameter. Section III presents the results, and Section IV concludes the paper.

I. Background and Data

A. Early Schooling in Britain

Children in England usually enter the first year of elementary school, the so-called reception class, at the age of four in the academic year in which they turn five. Attendance at reception class is close to universal⁷ and this class, although followed by more formal education in year one and year two of elementary school, is seen as the start of an elementary education that is clearly more learning oriented than the play-based nursery (preschool) education. Elementary education ends with year 6 at age 11, when the child moves into secondary education.⁸ Hereafter, we refer to reception class, year one, year two, and year six as the first, second, third, and seventh school year or grade.

The types of skills taught during the first year include rudimentary writing skills, the use of capital letters, and rudimentary counting. By the end of first grade, children should be able to “read a range of familiar and common words and simple sentences independently, ... write their own names and other things such as labels and captions, and begin to form simple sentences, ... count reliably up to 10 everyday objects, ... recognize numerals 1 to 9, begin to relate addition to combining two groups of objects and subtraction to ‘taking away,’ ... [and] use a mouse and keyboard to interact with age appropriate computer software” (UK Department of Education 2008).

Throughout this paper, we focus on the cohort of children born between September 1, 2000, and August 31, 2001, that, because there is no redshirting for a later year in the United Kingdom, enters elementary school in the 2005–2006 academic year. At that time, despite a recent convergence toward a single entry month (September) policy, there was substantial geographical variation across local authorities in school entry policies, which we exploit in this paper.⁹ Specifically, around 60 percent of the children in our sample were subject to the single-point entry policy in September (policy area A). The two second most frequent policies involved multiple entry points: about 20 percent were covered by a policy that anticipated school entry in September or January (policy area B), and 15 percent by a policy that anticipated entry in September, January, or April, depending on birth month (policy area C).¹⁰ Entry in an earlier or later than prescribed term is usually

⁷In administrative student records covering the full population of pupils in state-maintained schools in England, we find that 96 percent of pupils enrolled in year 2 attended reception class. Delaying school entry by an entire year or grade retention is not at all common in England. Our administrative data show that over 99 percent of children attending first grade in 2005–2006 and third grade in 2007–2008 were born between September 1, 2000, and August 31, 2001. Children in private schools are not included in the database being used. In England, only 5 percent of pupils aged 7 are enrolled in private schools, and this proportion is smaller for age 5 (Blundell, Dearden, and Sibiet 2010).

⁸For a complete overview of the English education system, see Gillard’s (2018) *Education in England: A Brief History*, available at <http://www.educationengland.org.uk/history/> or visit the website of the Department for Education at <http://www.gov.uk/df>.

⁹In England, there are approximately 150 local authorities (local government entities) with approximately 160,000 inhabitants on average (see <http://media.education.gov.uk/assets/files/xls/lfs/lfs2001aleaxls.xls>).

¹⁰Admission policies are explained in more detail in online Appendix E.

allowed if the parents wish it.¹¹ School funding, however, unlike the locally varying admission rules, comes from central government. The main criterion for fund allocation to schools is the number of pupils.

B. Data

Both the Millennium Cohort Study (MCS), from which we draw our survey data, and the National Pupil Database (NPD), from which we take our administrative data, provide longitudinal coverage of the September 1, 2000, to August 31, 2001, birth cohort. The MCS has administered surveys during the early months of these children's lives and then again at ages 3, 5, 7, and 11. These surveys cover a broad range of household characteristics, ranging from socioeconomic indicators, health of household members, and neighborhood characteristics to parenting practices and parent-child interactions. They also include assessments of child behavior and the child's cognitive, physical, and noncognitive abilities, reported by both parents and teachers and, from age seven onward, also in self-assessment questionnaires.¹² The MCS also provides the exact month of entry into the first school year, which enables computation of the actual exposure and estimation of the first stage without which we could not identify the causal effect of interest. School identifiers allow the MCS children to be matched to their results from the in-school end-of-year assessment from first and third grade. The first-grade assessment covers both cognitive areas such as language, literacy, problem solving, and numeracy, and noncognitive abilities such as social behavior and attitudes, creative development, and physical development (e.g., motor skills).¹³ The assessment scales for each of these scores are detailed in online Appendix G. The third-grade assessment covers test scores in reading, writing, math, and science. For the different outcomes at ages 5 and 7, the MCS sample used in the analysis comprises close to 8,000 children. Children registered as having special educational needs were removed from the analysis,¹⁴ and the sample is restricted to England to ensure comparability with our second dataset, the NPD, and because school entry rules and the early years curriculum differ in other parts of the United Kingdom.

The NPD contains administrative records of the total population of students at state schools in England and records student test scores on the nationwide assessments administered at different stages of the school curriculum. For reasons of comparability between the two datasets, we extract from the NPD the same academic year cohort covered by the MCS; that is, children born between September 1, 2000, and

¹¹ A few local authorities demand special justification (such as a doctor or social worker's recommendation) for an earlier than prescribed entry. A limit to late entry is set by national law stating that schooling becomes compulsory in the term following a child's fifth birthday. On average, we find an 82 percent compliance rate with school entry rules.

¹² For a detailed description of the survey design, recruitment processes, and fieldwork, see Dex and Joshi (2005).

¹³ This assessment, called the Foundation Stage Profile (FSP), is based on observation of the child throughout reception class, aided by a booklet in which teachers must regularly record the children's achievements. For details on the FSP, see the *Foundation Stage Profile Handbook* issued by the Department for Education and Skills in 2003 and available online at http://doc.ukdataservice.ac.uk/doc/6847/mrdoc/pdf/foundation_stage_profile_handbook.pdf.

¹⁴ The term "special educational needs" refers to conditions that include severe learning disabilities. We exclude children with these conditions in order to have a more homogeneous sample for our test score regressions, but we verified that including them does not change the results.

August 31, 2001. For this cohort, the NPD includes the assessments at the end of the first, third, and seventh grades. For first grade test scores, the NPD is a 10 percent sample of children in state-maintained schools in England from the targeted birth cohort (roughly 40,000 children); for test scores from age 7 onward, it covers the full population of these students (approximately 400,000). Besides test scores, the NPD also provides certain student background characteristics gathered from school records, including information on age, gender, ethnicity, whether English is spoken at home, eligibility for free school meals, and whether the child has special educational needs.

We base our analysis on a set of outcomes at ages 5, 7, and 11 taken from these two datasets. The National Pupil Database provides us with first-grade (age 5), third-grade (age 7), and seventh-grade (age 11) cognitive test scores from in-school assessments. With respect to noncognitive outcomes, we draw on the MCS data containing noncognitive assessments from parental, teacher, and self-reports on child behavior and noncognitive outcomes. These latter include, among others, the child's personal, social, and emotional development at age 5, as well as information on the teacher-child relationship, academic interest, self-perception, and disruptive behavior at ages 7 and 11.¹⁵

The cognitive, noncognitive, and behavioral outcomes used in our analysis are described in online Appendices F and G. Unless otherwise noted in the tables, we normalize all scores to a mean of zero and a standard deviation of one across the whole sample.

II. Estimation

A. Estimated Parameters and Empirical Strategy

In our empirical specification, the composite effect of receiving additional early schooling by entering school at an earlier age, which we refer to as the exposure effect ($\beta_1 - \beta_3$ in our discussion in the introduction), is γ_1 in the equation

$$(1) \quad y_{imr} = \gamma_0 + \gamma_1 EXP_{imr} + \sum_j \alpha_j DAGET_{imr}^j + \mathbf{Z}'_{imr} \delta + \mu_m + \rho_r + v_{imr},$$

where y_{imr} is an outcome for individual i born in birth month m and attending school in local authority r , and EXP_{imr} is the length of exposure to schooling up to the test; $DAGET_{imr}^j$ are an exhaustive set of dummies for age at the test indexed by j and measured in months. The vector \mathbf{Z}_{imr} includes background variables that are included to increase precision, but are not necessary for the validity of the IV identification strategy that we describe below. The birth month fixed effects μ_m control for seasonal variation in the outcome across birth months. For those outcomes that are assessed around the same time for all children (such as test scores from school

¹⁵We did not use teacher reported noncognitive outcomes at ages 7 and 11 from the MCS because these were gathered through class teacher questionnaires that have a large number of nonrandom missing values, because both parents and children must give consent for the class teacher to be interviewed. As on this reduced sample we could not reproduce the results from our main test score regressions, we concluded that it is selective and did not use it. There are also some age five outcomes from the MCS that we did not use, including the psychometric test scores from the "British Ability Scales," because these measurements were taken *during* the first school year when some children would have barely been exposed to schooling.

exams), they effectively control for age-at-test; hence, for these outcomes, we drop the age-at-test dummies. We include local authority fixed effects, ρ_r , to control for region-specific unobserved factors, such as teaching quality.

The error term v_{imr} includes unobserved child characteristics, such as intellectual ability or maturity. Given that parents have discretion over the choice of the school entry term, observed exposure EXP_{imr} is likely to be correlated with v_{imr} . For example, if parents of high-ability children tend to bring school entry forward while parents of low-ability children tend to delay school entry, then the exposure effect estimated by applying OLS to (1) will be upward biased.

We address this possible endogeneity by instrumenting actual exposure (EXP) with expected exposure (EEXP) prescribed by the school entry rules. Recall that we sample a cohort of children that all enter elementary school in the 2005/2006 academic year, but depending on the local school-entry policy, there may be up to three possible entry months into the first school year—September 2005, January 2006, and April 2006—corresponding to the three terms of the academic year. The three most frequent school-entry policies are the following.¹⁶ In school entry policy A, all children irrespective of their birth month are scheduled to enter school in the first term of the academic year (September 2005); thus, in policy area A, there is no birth month cutoff, and school-entry rules do not cause any variation in expected exposure. In policy area B, children born before March 2001 are scheduled to enter in the first term, and children born from March 2001 onward are scheduled to enter in the second term of the academic year. Finally, in policy area C, there are two birth-month cutoffs and three possible school-entry dates. Children born before January 2006 are supposed to enter school in the first term, children born from January to March 2006 are supposed to enter in the second term, and children born from April 2006 are supposed to enter in the third term.

Overall, there are thus three possible rule-prescribed entry months into the first school year—September 2005, January 2006, and April 2006—and the school year runs until the end of July 2006. Consequently, expected (rule-prescribed) exposure only takes on three different values: 4 months (if the expected school entry is April), 7 months (if the expected school entry is January), and 11 months (if the expected school entry is September).¹⁷ Because we do not want to impose the assumption of a linear relation between expected exposure and actual exposure, we split expected exposure up into dummies.

We estimate equation (1) using the two-stage least squares (TSLS) method based on the first-stage regression:

$$(2) \quad EXP_{imr} = \pi_0 + \pi_1 EEXP7 + \pi_2 EEXP11 + \sum_j \phi_j DAGET^i_{imr} + \mathbf{Z}'_{imr} \theta \\ + \mu_m + \rho_r + \varepsilon_{imr}$$

¹⁶ Admission policies are explained in more detail in online Appendix E.

¹⁷ Seventy-eight percent of the children in our sample are expected to have 11 months exposure, 16 percent are expected to have 7 months exposure, and about 6 percent are expected to have 4 months exposure.

in which $EEXP7$ is a dummy indicating 7 months expected exposure (January entry), $EEXP11$ is a dummy indicating 11 months expected exposure (September entry), and 4 months expected exposure is the reference group.

By its definition, the instrument of expected exposure determined by the school entry rules depends on birth month and policy area, both of which have their own effects on the outcome. For example, there could be potential differences in the teaching quality of local authorities that could be systematically related to each authority's school entry policy, leading to a correlation between outcomes (via teaching quality) and expected exposure (via the school entry rules). Conditioning on the birth month fixed effects μ_m and the local authority effects ρ_r eliminates these differences. Hence, this exploits a difference-in-difference type of variation, in which we compare the difference in outcomes between children born in a local authority where a difference in birth months causes a difference in expected exposure, with the corresponding difference in outcomes between children born in a local authority where a difference in birth month does not cause variation in exposure.

The usual difference-in-differences common trends assumption applied to our context is that birth month effects have to be uniform across regions, and region fixed effects have to be uniform across birth months. That is, the additively separable specification in birth month and region fixed effects in (1) and (2) must be correct, in the sense that there should be no interaction effects between birth month and region. For example, there must be no birth-month specific differences in teaching quality across the different policy areas, or no regional differences in the seasonality of birth month effects. We provide empirical tests for the validity of the instrument in the next section. In online Appendix B, we show that we get almost identical results when we use an alternative regression discontinuity research design which relaxes the difference-in-differences common trends assumption but instead relies on the (not necessarily weaker) assumption that the running variable (age-at-test/birth month) is correctly specified via a given continuous function, and that being born before or after the cutoff is exogenous (e.g., parents do not manipulate the birth month of their child).

When using the outcomes observed in the NPD dataset, we need to implement a *two-sample* TSLS estimation procedure. That is, because we observe no actual exposure in that dataset, we use the coefficient estimates of the first-stage regression (2) from the MCS dataset to predict actual exposure \widehat{EXP}_{imr} in the NPD data. The TSLS estimate in the NPD data is then obtained by running regression (1) with \widehat{EXP}_{imr} in the place of EXP_{imr} .¹⁸ Throughout the analysis, we cluster standard errors at the level of the local authority, and for estimations involving the MCS dataset, we apply the sample weights provided for that dataset.

¹⁸Following Inoue and Solon (2010), we adjust the standard errors by multiplying the second step covariance matrix by $1 + 1/\hat{\sigma}^2[n_{MCS}/n_{NPD}]\hat{\beta}_{TSLS}'\hat{\Sigma}_\eta\hat{\beta}_{TSLS}$, where $\hat{\sigma}^2$ is the mean squared residual from the second-stage regression, $\hat{\beta}_{TSLS}$ is the estimated $K \times 1$ coefficient vector from the second-stage regression, n_{NPD} is the sample size from the second-stage regression, n_{MCS} is the sample size from the first-stage regression, and $\hat{\Sigma}_\eta$ is the estimated $K \times K$ covariance matrix of the K residual vectors from all K first stages. In our application, this correction factor adjusts the standard errors upward by factors between 1.02 and 1.20.

TABLE 1—BALANCING TESTS

Dependent variable	Dataset	<i>EEXP7</i>	<i>EEXP11</i>	Observations	<i>p</i> -value joint significance
Naming vocabulary score at age 3	MCS	−0.037 (0.051)	0.027 (0.069)	6,752	0.30
English not first language at home	MCS	−0.011 (0.010)	−0.005 (0.011)	7,805	0.53
English not first language at home	NPD	−0.001 (0.006)	−0.002 (0.007)	42,702	0.96
Mother left education before the age of 16	MCS	0.008 (0.027)	0.016 (0.030)	7,778	0.86
Single parent	MCS	0.007 (0.032)	0.016 (0.033)	7,805	0.82
Parents on income support (age 3)	MCS	−0.009 (0.027)	0.002 (0.029)	7,134	0.73
Homeowner	MCS	0.022 (0.043)	0.013 (0.045)	7,805	0.83
Poverty indicator	MCS	−0.024 (0.030)	−0.009 (0.035)	7,796	0.56

Notes: The table shows that the instrumental variables are uncorrelated with a number of family background variables. Each line of the table represents a separate regression, in which the family background variable mentioned in the first column is regressed on dummy IV variables for 7 months of expected exposure (*EEXP7*) and 11 months of expected exposure (*EEXP11*) to the first school year, the reference being 4 months of expected exposure. The only control variables are local authority and birth month fixed effects. Standard errors clustered at the level of the local authority are in parentheses. None of the coefficients are individually significant, nor are they jointly significant, at conventional levels of statistical significance.

Source: Data source indicated in the second column as MCS (Millennium Cohort Study) or NPD (National Pupil Database)

B. Instrument Validity

To check the identifying assumptions of our difference-in-differences IV specification, we regress child and parent characteristics—such as a naming vocabulary test score at age 3, whether English is spoken at home, mother’s education, single parenthood, income support received by the parents, etc.—on the instruments *EEXP7* (expected January entry = expected 7 months exposure) and *EEXP11* (expected September entry = expected 11 months exposure), conditioning on birth month and region fixed effects. The results are reported in Table 1 and show that the association of the instruments with these background characteristics turns out to be small and insignificant in all these regressions, with *p*-values between 0.3 and 0.96.

C. Alternative Treatment to School Entry

Although exposure to the first school year should have an effect on subsequent test scores through its orientation toward learning, this effect—and its interpretation—will depend on the comparison outcome; that is, the alternative childcare arrangements that a four-year-old child is exposed to before entering school. Table 2 reports information on childcare arrangements for preschool children from the 2008 Childcare and Early Years Survey of Parents (Speight et al. 2009). As reported in panel A of the table, only 10 percent of 3- to 4-year-old preschool children in

TABLE 2—INCIDENCE OF CHILDCARE ARRANGEMENTS FOR PRESCHOOL CHILDREN IN ENGLAND, 2008

<i>Panel A. Attendance rates of formal and informal childcare, 3–4-year-olds</i>		
	Attendance rate in percent	
Parental childcare only	10	
Mainly center-based childcare	40	
Center-based and informal childcare	31	
Center-based and other formal (such as child minders)	12	
Informal childcare only	3	
<i>Panel B. Time spent in types of formal childcare, conditional on attending, all age groups</i>		
	Weekly hours	
	Median	Mean
Nursery school	14.8	15.3
Nursery class	12.5	14.7
Day nursery	25.5	22.8
Playgroup or preschool	12.5	8.8
Child minder	9	13

Notes: The table shows that attendance rates to formal childcare among preschool children are high, and that formal childcare attendance is mainly part-time. For comparison, full-time school attendance in the first school year is approximately 31 hours a week. The childcare types in panel A are constructed as mutually exclusive groups. The remaining 4 percent of children receive other types or combinations of childcare.

Source: Panel A was compiled from table 3.1 in Speight et al. (2009), and panel B from table 2.8 in Speight et al. (2009). The underlying data source is the Childcare and Early Years Survey of Parents from the year 2008.

England receive parental childcare only, and a mere 3 percent of children receive other informal childcare only, while the majority (83 percent) receives some form of center-based childcare, partly combined with informal care or other types of formal care, such as child minders. As panel B of Table 2 reports, much of the attendance at formal childcare is part-time, with mean and median attendance below 15 hours per week for most types of formal care. For comparison, attendance at school during the first year, our treatment, corresponds to roughly 31 hours per week. This finding implies that for a high proportion of children our treatment consists of increasing the exposure to full-time learning-oriented early schooling at the expense of part-time and more play-oriented center-based care and some parental or informal care. For a smaller proportion of children, the counterfactual is parental or other informal childcare only.

D. Interpretation of the IV Estimator and One-Sided Noncompliance

If treatment effects are heterogeneous, IV estimates can only be meaningfully interpreted if the monotonicity (or uniformity, or “no-defier”) assumption holds. This assumption requires that all individuals who change treatment status in response to a change in the instrument, do so in the same direction (i.e., they either get all switched into the treatment, or all switched out of the treatment, as the instrument is switched from 0 to 1). In this case, IV estimation identifies a local average treatment effect (LATE) representative for the subgroup of compliers (Imbens and Angrist 1994). Because we have a multivalued treatment (4, 7, or 11 months of exposure)

and two dummy variable instruments, our effect is a weighted average of several LATEs, as we clarify in online Appendix C.

In our application, we observe an interesting pattern according to which the only form of noncompliance with the school entry rules is toward *earlier* entry than recommended. We illustrate the compliance pattern in Figure 2, which shows how the interaction of birth month and school entry policy area affects the term of entry into first grade. The figure shows the shares of children entering in the first, second, and third term of the academic year (on the y-axis), by birth month (on the x-axis). Each of the three panels (A, B, and C) in the figure is for a different policy area (a group of local authorities operating the same school entry policy). For each birth month and policy area, there is a “correct” (rule-prescribed) entry term. The share corresponding to the correct entry term is marked with a circle. The “correct” entry term changes at the birth-month cutoff dates, which are marked by vertical lines. In policy area A (panel A), there are no cutoff dates within the academic year cohort because there is a uniform rule of entry in the first term irrespective of birth month. The figure shows an almost perfect compliance with this rule, with the “correct” share being close to 1 over all birth months. In policy area B, children born up to February are subject to the same rule as in policy area A, but children born from March onward are supposed to enter in the second term of the year. The corresponding figure in panel B shows that, while almost all children born before the cutoff comply with the first-term entry rule, only about 40–50 percent of children born from March onward comply with the second-term entry rule. In policy area C, compliance with the first-term entry rule is close to 1 (first section of the graph), compliance with the second-term entry rule (middle section) is around 40 percent, and compliance with the third-term entry rule (last section) is around 30 percent. Overall, noncompliance almost exclusively consists of earlier than rule-prescribed entry, i.e., nonzero shares for incorrect entry terms almost always refer to earlier entry terms than the recommended entry term. A likely motive for this is that early school entry provides a free form of childcare to parents. As we discuss in online Appendix D and show in online Appendix Table A1, the estimated share of individuals who choose to enter school late, even though the rule indicates early entry, is close to 0. That is, there are almost no never-takers (Angrist, Imbens, and Rubin 1996). This constitutes a special case of one-sided noncompliance with two important implications. First, one-sided noncompliance rules out the existence of defiers, and monotonicity is automatically satisfied (Imbens 2014). Second, LATE is equal to the average treatment effect on the untreated (ATU), and has thus strong external validity.¹⁹ The interpretation of our IV estimates, therefore, is that they capture the effect of extending exposure to early schooling for those individuals who currently have low levels of exposure.

¹⁹The reason for this is that if there are no never-takers, all untreated individuals are compliers (with the instrument switched off). Moreover, the IV assumption that the instrument is as good as randomly assigned ensures that treated compliers (with the instrument switched on) and untreated compliers (with the instrument switched off) are similar. Therefore, if there are no never-takers, compliers are representative for the untreated.

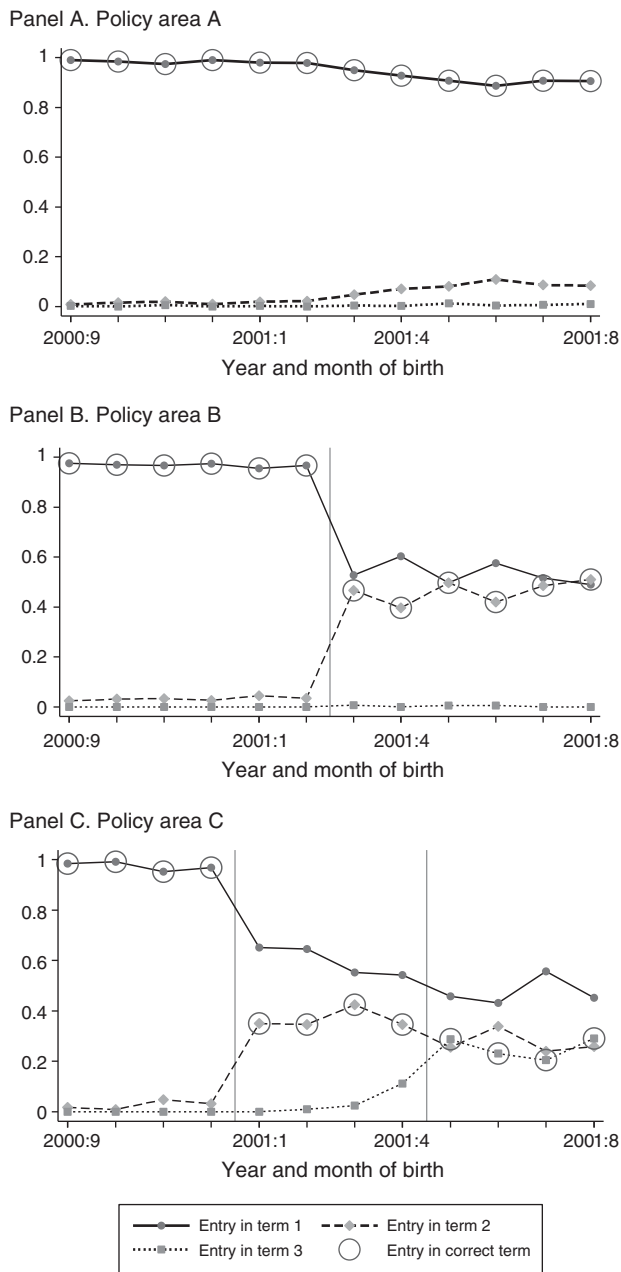


FIGURE 2. FIRST STAGE EXPRESSED BY SHARES OF CHILDREN PER ENTRY TERM

Notes: The figure shows the shares of children entering first grade in the first, second, and third term of the academic year (on the y-axis) by birth month (on the x-axis) and policy area (in the different panels of the figure). Vertical lines mark birth month cutoff dates, and the share marked by a circle refers to the “correct” entry term according to the relevant rule. See Section IID in the main text for a more detailed description.

Source: Millennium Cohort Study

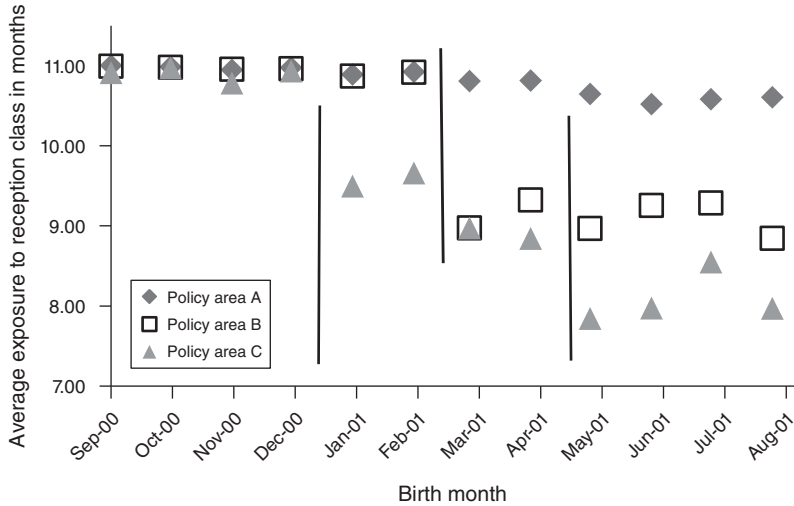


FIGURE 3. FIRST STAGE EXPRESSED BY AVERAGE EXPOSURE

Notes: The figure reports average exposure to the first school year by birth month for three policy areas with different school-entry rules. The figure shows that the school-entry rules have an impact on the average length of the first school year; that is, the first stage of our IV approach. In policy area A, the school-entry rule is that all children enter in September (get 11 months of exposure) regardless of their birth month. The corresponding diamond-shaped data series shows strong compliance with that rule. In policy area B, children born between September and February are supposed to enter school in September (get 11 months of exposure), while children born from March onward should enter in January (get 7 months of schooling). Consequently, there is a drop in average exposure from the birth month of March onward in the corresponding square-shaped data series. In policy area C, children born between September and December are supposed to enter school in September (get 11 months of exposure), children born between January and April are supposed to enter in January (get 7 months of exposure), and children born from May onward are supposed to enter in April (get 4 months of exposure). In line with these rules, the corresponding triangular-shaped time series drops in the birth months of January and May.

Source: Millennium Cohort Study

III. Results

A. Graphic Representation of the First Stage and Reduced Form

The discontinuities in the shares of children entering in each of the three terms shown in Figure 2 translate into corresponding discontinuities in the average duration of schooling in first grade by birth month and policy area. In Figure 3, we show that in policy area B, average exposure drops from 11 months for children born before the cutoff date to around 9 months for children born after the cutoff date. In policy area C, average exposure drops from 11 to 9 months at the first cutoff, and then to around 8 months at the second cutoff.

To illustrate the reduced form of the relation, Figure 4 plots the average standardized test scores from the NPD dataset against birth months for the different policy areas. To filter out a common birth month (age-at-test) trend, the figure shows test scores for policy areas B and C relative to policy area A (in which exposure does not systematically vary by birth month). For policy area B (panel A), comparing the averages before and after the March cutoff date reveals a drop in first-grade test

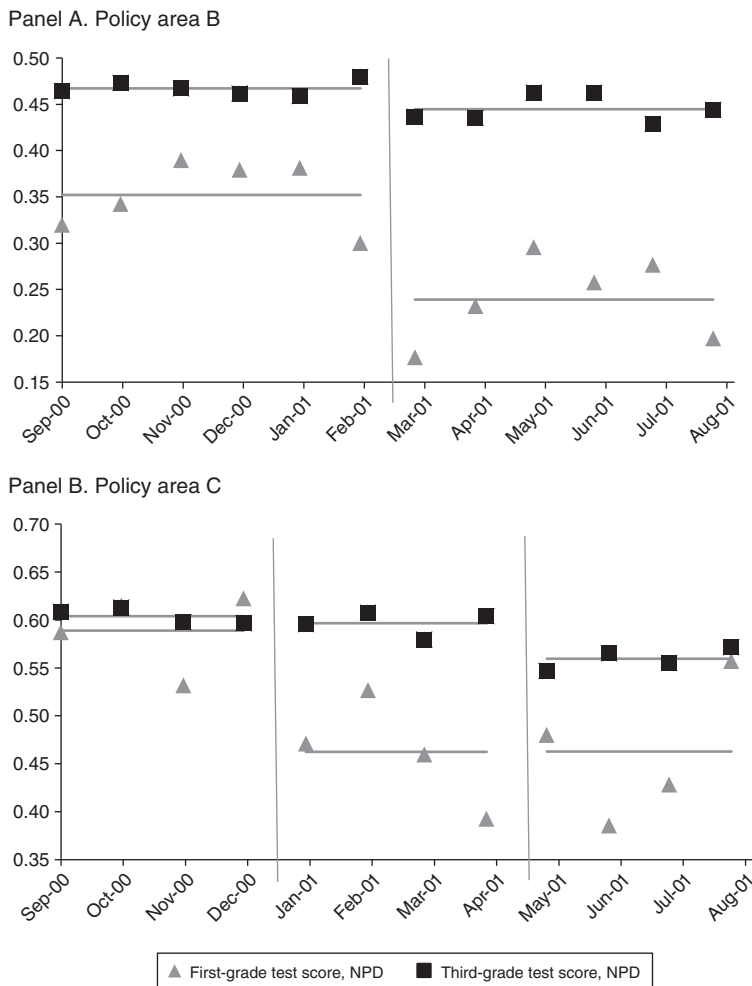


FIGURE 4. REDUCED FORM

Notes: The figure reports birth-month averages for first-grade and third-grade test scores for policy areas B (panel A) and C (panel B). The values shown are relative to policy area A in order to eliminate common birth-month (age-at-test) effects. The vertical lines represent the cutoff dates from the school-entry rules. Horizontal bars represent averages over windows defined by cutoffs. Children to the left of the cutoff get on average more exposure to reception class than pupils to the right of the cutoff (see Figures 2 and 3). Here, we show the associated difference in test scores. In policy area B, first-grade test scores drop by about 10 percent of a standard deviation and third-grade test scores by about 3 percent of a standard deviation around the March cutoff date. In policy area C, first-grade test scores drop by about 15 percent of a test score standard deviation around the January cutoff date, and do not change noticeably around the May cutoff date. The third-grade test scores do not drop around the January cutoff date, and drop by around 4 percent of a standard deviation around the May cutoff date.

Source: National Pupil Database

scores of about 10 percent of a standard deviation. For the third-grade test score, the decrease is only about one-third as large. In policy area C (panel B), the first-grade test score data reflect a drop equivalent to 15 percent of a test score standard deviation around the January cutoff date but no discernible drop around the May cutoff date.

TABLE 3—EARLY EXPOSURE EFFECTS ON THE FSP TOTAL SCORE AT THE END OF FIRST GRADE (AGE FIVE)

Model	OLS		Reduced form		IV	IV-2S	IV	IV-2S
	MCS	MCS	NPD	MCS	Both	MCS	Both	
Dataset	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
<i>Panel A. Outcome equation (dependent variable: FSP total score)</i>								
Exposure	0.063 (0.011)	0.031 (0.011)	0.031 (0.006)	0.087 (0.025)	0.089 (0.019)	0.082 (0.025)	0.084 (0.017)	
Observations	7,805	7,805	42,091	7,805	42,091	7,806	42,091	
<i>Panel B. First-stage equation (dependent variable: Exposure)</i>								
Expected exposure				0.350 (0.060)	0.350 (0.060)			
Expected exposure = 7 months						0.827 (0.379)	0.827 (0.379)	
Expected exposure = 11 months						2.339 (0.459)	2.339 (0.459)	
Observations				7,805	7,805	7,805	7,805	
F-statistic				34.9	34.9	18.1	18.1	

Notes: Panel A shows estimated effects of a one-month increase in exposure to the first school year on the total test score from the in-school assessment at the end of the first grade at age five. Column 1 shows a simple OLS regression on the endogenous exposure variable. Columns 2 and 3 show reduced-form estimates from regressions on expected exposure in the MCS and the NPD datasets. Columns 4 and 5 show IV estimates using linearly coded expected exposure as the instrument for the MCS dataset and the NPD dataset. Columns 6 and 7 show IV estimates using expected exposure coded as two dummy variables for the two datasets. Columns labeled IV-2S show two-sample TSLS estimates with the first stage estimated with MCS data and the second stage with NPD data. Panel B reports the first-stage results for the IV specifications. The reported *F*-statistic is for a test of excluded instruments. Control variables are dummies for gender, free school meal eligibility, English first language at home, ethnicity, birth month, and local authority. Standard errors clustered at the level of the local authority are in parentheses.

Source: Data source indicated in the table header as MCS (Millennium Cohort Study) or NPD (National Pupil Database)

The third-grade test score data show no drop around the January cutoff date and only a small drop of around 4 percent of a standard deviation around the May cutoff date.²⁰

B. Cognitive Test Scores

We now conduct a regression analysis of the cognitive test scores from the school exams, estimating equation (1) above. Table 3 first shows the different elements of our empirical strategy for one outcome, the total score from the teacher assessment at the end of the first school year, and for the two datasets we are using. Exposure is measured in months and, unless otherwise stated, outcomes are normalized to a mean of zero and a standard deviation of one. The exposure effect thus picks up the effect of a one-month increase in exposure measured in terms of the standard deviation of the outcome. The first column of Table 3 shows estimates of a simple OLS regression disregarding the problem that exposure is endogenous, where we find an exposure “effect” of about 6 percent of a standard deviation. In the next two columns, we report reduced-form or intention-to-treat (ITT) effects obtained by regressing the

²⁰The reason the data points for the triangular first-grade test scores fluctuate more strongly around their mean than the third-grade test scores is that they are provided in the NPD data only as a 10 percent sample, whereas the third-grade scores cover the full population.

TABLE 4—EARLY EXPOSURE EFFECTS ON COGNITIVE TEST SCORES AT DIFFERENT AGES, BY GENDER

	Language skills			Numeracy skills		
	1st grade Age 5	3rd grade Age 7	7th grade Age 11	1st grade Age 5	3rd grade Age 7	7th grade Age 11
<i>Panel A. Overall effect</i>						
Actual exposure	0.100 (0.020)	0.026 (0.007)	0.004 (0.004)	0.073 (0.016)	0.014 (0.005)	−0.002 (0.004)
Observations	42,091	410,359	390,696	42,090	410,217	393,833
<i>Panel B. Effect by gender</i>						
Actual Exposure × male	0.088 (0.022)	0.020 (0.008)	−0.001 (0.006)	0.066 (0.018)	0.009 (0.006)	−0.009 (0.006)
Actual Exposure × female	0.112 (0.023)	0.032 (0.007)	0.009 (0.005)	0.082 (0.020)	0.019 (0.006)	0.005 (0.004)
Observations	42,091	410,359	390,696	42,090	410,217	393,833

Notes: The table shows IV (two-sample TLS) estimates for the effect of a one-month increase in exposure to the first school year on cognitive outcomes at different ages. Control variables are dummies for gender, free school meal eligibility, English first language at home, ethnicity, birth month, and local authority. Separate results for males and females are obtained by interacting all regressors with male/female dummies. Standard errors clustered at the level of the local authority are in parentheses.

Source: National Pupil Database (with first stage estimated in the Millennium Cohort Study)

outcome on expected exposure, with very similar reduced-form estimates of around 0.03 across the MCS and the NPD datasets. Columns 4 and 5 report IV estimates of about 0.09 obtained from the MCS and NPD datasets when expected exposure is used as a linear regressor.²¹ The magnitude of the IV effect remains the same when specifying the instrument as two dummies instead of a linear regressor as shown in columns 6 and 7. Here, test statistics for the *F*-test of excluded instruments from the first stage are around 18, implying that the instruments are strong.²² We verified that excluding policy area C, which has the lowest compliance rates, results in very similar estimates (results available upon request). Overall, the IV results indicate that an additional month of exposure to early schooling increases the first-grade total test score by about 9 percent of a standard deviation, a result that is remarkably similar across the MCS and NPD datasets. As we show in online Appendix Table A2 and discuss in online Appendix B, we also find a very similar magnitude if we implement a regression discontinuity design instead of the difference-in-differences design.

In Table 4, we report the exposure effects on indices of language and numeracy skills, aggregated from a range of cognitive subject-specific scores taken at ages 5, 7, and 11 (see online Appendices F and G for a description of the different test

²¹ The reduced-form (ITT) models reported in the second and third columns are similar to the models estimated by Crawford, Dearden, and Meghir (2007) and Leuven et al. (2010). Both of these papers do not observe the actual month of school entry, and in estimating the effects of the expected age-at-entry or school exposure on test scores, they identify reduced-form effects. Our results suggest that ITT effects can be considerably smaller than IV effects: our first-stage estimate of 0.35 in the second panel of the table implies that the reduced-form estimate of about 0.03 has to be scaled up by a factor of $1/0.35 = 2.9$ to arrive at the IV estimate of 0.087.

²² Stock and Yogo (2005) defines a strong instrument in terms of several different criteria, one of them being that the bias induced in the hypothesis testing be small enough that a nominal 5 percent hypothesis test actually rejects it no more than 15 percent of the time. The critical value for this criterion (with one endogenous variable and two instruments) is 11.59, and by this measure, our instruments are strong.

scores), differentiating by gender, and based on the two-sample TSLS estimates, corresponding to column 7 of Table 3.²³ At age five, the effect on language skills is of roughly the same magnitude as the IV effect on the total score reported in the previous table, while the effect on numeracy skills is slightly smaller. Estimating separate effects by gender, obtained by interacting *all* regressors with gender dummies, we find that girls have slightly higher effects than boys in the language and numeracy scores. Further, the results in Table 4 also show a fading out of the effects at higher grades. The effects on test scores assessed at the end of third grade of elementary school at around age 7 are about 20 to 30 percent of the magnitude found for the first-grade test scores. At age 7, an additional month of early schooling increases language skills by about 2 to 3 percent of a test score standard deviation, while effects on numeracy skills are in the range of 1 to 2 percent. The effects have largely disappeared, however, 4 years later at age 11, although a weakly significant but small effect remains for girls in language skills.²⁴ Overall, therefore, we find substantial effects of earlier school attendance on cognitive outcomes at age 5 and smaller effects at age 7, which are more pronounced for girls than for boys. Our estimates for age 11 (6 years after school enrollment) hint at some effects for girls, but these are small in magnitude.

The finding that our effects on cognitive test scores diminish at higher grades could imply that the effects are due to an initial disadvantage of the children who enter reception class later than their peers, but that these children ultimately catch up. A policy implication of this would be that important decisions based on test scores, such as future school type, should not be taken at a too early stage. Yet, despite this fading out of effects on cognitive test scores, early interventions could still have lasting effects by boosting noncognitive skills.²⁵ To investigate this possibility further, we analyze the effects of early schooling on noncognitive skills and behavioral outcomes in the next section.

C. Early Exposure to Schooling and Noncognitive and Behavioral Outcomes

Table 5 reports the results of applying our IV strategy to aggregated noncognitive scores based on teacher, parent, and self-assessments at ages 5, 7, and 11 from the NPD and MCS datasets. These scores are described in detail in online Appendices F and G. For age 5 (panel A), we have information on three teacher-assessed noncognitive outcomes from both the NPD and MCS data. We find positive effects

²³For outcomes that are available both in the NPD and the MCS data, the two sample TSLS estimates are our preferred specification because of their much higher precision given the much larger NPD sample size. In online Appendix Table A3, we show the same outcomes using the MCS data. Although the results for age 5 are very similar, the effects at ages 7 and 11 are much less precisely estimated in the MCS. Because the NPD data does not provide the detailed measures on socioeconomic status (SES) included in the MCS data, we first differentiate our results by gender, and turn to interactions with SES in Section IIID using MCS data only.

²⁴Although the average gender difference found is small, the finding that girls have higher effects on average is in line with the gender differences identified by Anderson (2008), Cascio (2009), and Havnes and Mogstad (2011).

²⁵Cascio and Staiger (2012) argues that, because of knowledge accumulation, the standard deviation of knowledge is likely to rise at higher grades. Standardizing the outcome with the grade-specific standard deviation would therefore mechanically lead to fade-out at higher grades even if the underlying effect on absolute knowledge is constant. While Cascio and Staiger (2012) shows that this mechanism can explain fade-out to some extent, it is unlikely to be the sole explanation for the strong fade-out to virtually zero that we observe in our data.

TABLE 5—EARLY EXPOSURE EFFECTS ON NONCOGNITIVE AND BEHAVIORAL SCORES

Dataset	Creative development		Physical development		Personal, social, and emotional development	
	NPD	MCS	NPD	MCS	NPD	MCS
<i>Panel A. Age 5, NPD and MCS</i>						
Exposure	0.065 (0.015)	0.057 (0.026)	0.058 (0.012)	0.055 (0.024)	0.063 (0.014)	0.056 (0.025)
Observations	42,090	7,805	42,090	7,805	42,091	7,805
Exposure × male	0.057 (0.018)	0.038 (0.037)	0.069 (0.018)	0.053 (0.035)	0.058 (0.019)	0.051 (0.035)
Exposure × female	0.071 (0.019)	0.078 (0.033)	0.045 (0.016)	0.058 (0.029)	0.068 (0.019)	0.062 (0.029)
Observations	42,090	7,805	42,090	7,805	42,091	7,805
			Teacher relationship I	Academic interest I	Positive self-perception I	Disruptive behavior I
<i>Panel B. Age 7, MCS</i>						
Exposure			0.064 (0.026)	0.058 (0.036)	0.039 (0.045)	-0.035 (0.030)
Observations			6,159	6,264	6,339	6,390
Exposure × male			0.098 (0.042)	0.120 (0.060)	0.034 (0.065)	-0.006 (0.044)
Exposure × female			0.027 (0.039)	-0.0048 (0.031)	0.042 (0.047)	-0.075 (0.045)
Observations			6,159	6,264	6,339	6,390
			Teacher relationship II	Academic interest II	Positive self-perception II	Disruptive behavior II
<i>Panel C. Age 11, MCS</i>						
Exposure			0.051 (0.043)	0.028 (0.029)	0.005 (0.026)	-0.081 (0.029)
Observations			5,317	5,861	6,250	6,246
Exposure × male			0.108 (0.059)	0.128 (0.045)	0.022 (0.037)	-0.105 (0.046)
Exposure × female			0.007 (0.050)	-0.050 (0.041)	-0.015 (0.040)	-0.060 (0.031)
Observations			5,317	5,861	6,250	6,246

Notes: The table shows IV estimates for the effect of a one-month increase in exposure to the first school year on noncognitive and behavioral outcomes. Panel A shows teacher assessments from the end of first grade, available in both the NPD and MCS datasets. Estimates involving the NPD dataset are estimated by two-sample TSLS. In panels B and C, the outcomes are normalized factors obtained from a factor analysis on several outcomes. See online Appendix F for a description of the dependent variables used in this table. Control variables are dummies for gender, free school meal eligibility, English first language at home, ethnicity, birth month, and local authority. Separate results for males and females are obtained by interacting all regressors with male/female dummies. Standard errors clustered at the level of the local authority are in parentheses.

Source: Data source for panel A indicated in the table header as MCS (Millennium Cohort Study) or NPD (National Pupil Database). Data source for panels B and C is the MCS.

for physical development (covering coordination and fine motor control); creative development; and personal, social, and emotional development. The effects are of similar magnitude for boys and girls and suggest a positive impact of earlier exposure to schooling on important noncognitive behavioral outcomes. The effect size is about 6–7 percent of a standard deviation for an additional month of school entry, which is only slightly smaller than the effects on cognitive test scores at the same age shown in the previous table.

Panels B and C of Table 5 report the results for outcomes assessed 2 and 6 years later, at age 7 and 11, by parents and by the child. At these later ages, information on noncognitive skills is only available from the MCS, which is of a smaller sample size. We find that early exposure affects a range of behavioral responses at age 7. For instance, starting elementary school earlier improves academic interest and the relationship with the teacher for boys, and reduces disruptive behavior for girls, while effects on positive self-perception have positive point estimates for both genders but are statistically insignificant. Remarkably, these effects remain statistically significant and of a similar magnitude at age 11 (see panel C, Table 5), where now the beneficial effect on disruptive behavior also gains significance for boys. This is in clear contrast to the fast fading-out of the effects on the cognitive skills, which we documented in the previous section.

Overall, these findings confirm expectations that earlier exposure of 4-year-olds to same-age peers in a professional childcare setting, as well as exposure to early learning, rather than being harmful, actually fosters a range of important social skills throughout ages 5, 7, and 11.²⁶ This is the case in particular for boys, for whom we find sizable beneficial effects (with magnitudes of around 10 percent of a standard deviation) at age 11 on the teacher relationship, academic interest, and disruptive behavior; while for girls, from age 7 onward, only the effect on disruptive behavior remains significant.

D. The Role of Socioeconomic Status

Given that providing early schooling programs is costly, it is important to understand whether such programs are particularly effective for certain groups, to which they could then be targeted. Hence, the literature on the effects of early schooling or childcare programs usually looks at heterogeneous effects by parental background. Elder and Lubotsky (2009), for example, finds stronger age-at-entry effects for children with higher socioeconomic family backgrounds, and Magnuson, Ruhm, and Waldfogel (2007) finds that positive cognitive effects from prekindergarten attendance are more long lasting for disadvantaged children. Likewise, observational studies of the effects of preschool programs in the United States on test scores also tend to find that the benefits are often greater for children from disadvantaged backgrounds (Currie 2001), and a similar pattern has been uncovered for universal childcare programs in Germany and Norway (Cornelissen et al. 2018; Havnes and Mogstad 2011, 2015).

To analyze the role of parental background, we measure socioeconomic status (SES) using the National Statistics Socioeconomic Classification, an instrument devised by the UK Office for National Statistics and provided as part of the MCS

²⁶This is in line with the view of child psychologists, that for children of this age group, exposure to peers and caregivers other than parents provides opportunities for child development that cannot be experienced at home, particularly so if the quality of nonparental care is high (Lamb and Ahnert 2007). Two commonly used measures of quality—student-to-teacher ratios and teacher salaries—suggest that the quality of care in UK elementary schooling is by no means low. In 2006, the ratio of students to teaching staff in elementary education was 19.8, similar to the ratios in France and Germany but higher than the ratio of 14.5 for the United States (OECD 2008, Table D.2.2). The ratio of an experienced elementary school teacher's salary to GDP per capita was 1.3 in England in 2006, compared to the United States of 0.97 and an OECD average of 1.22 (OECD 2008, Table D.3.1).

dataset. This measure classifies parental occupation into 14 categories, with the 3 highest categories being entrepreneurs of large establishments, higher managerial and administrative occupations, and higher professional occupations, and the lowest being semi-routine occupations, routine occupations, and “never worked or long-term unemployed.”²⁷ We define family SES as the highest SES among the parents, and create dummy variables for “low SES,” corresponding to the bottom quartile of family SES, and “high SES,” corresponding to the three top quartiles of family SES. In terms of the underlying occupational categories, “low SES” includes the lowest three of these categories mentioned above. Being based on broadly defined occupational choice, this measure is likely to be largely determined by past educational and occupational choices and hence much less likely to be endogenous to (or an outcome of) the school entry decision for the child than alternative SES measures such as household income. As online Appendix Table A4 shows, SES is indeed strongly correlated with socioeconomic family characteristics. For example, the share of homeowners among low SES children is about 0.44 versus 0.71 among high SES children, while the share of children with a low-educated mother is about 0.62 among low SES children but about 0.41 among high SES children.²⁸

In Table 6, we report the exposure effects for ages 5, 7, and 11 on cognitive and noncognitive outcomes, allowing for interactions of early school exposure with indicator variables for high and low socioeconomic status (SES). We use common factors that aggregate the different outcomes used in Tables 4 and 5 into one overall cognitive score and one overall noncognitive score for each age group. The results in Table 6 therefore also provide an overall synthesis of our results.²⁹ For cognitive and noncognitive outcomes, there emerges a strong pattern in which the early exposure effects for boys are driven primarily by low SES boys, both at ages five and seven. At age 11, effects on cognitive outcomes have faded away, but a uniform positive effect on the noncognitive outcomes persists across SES for boys. For girls, the positive exposure effect appears to be more uniform across socioeconomic groups, and largely fades away at later ages.

Higher returns to early schooling for low-SES boys compared to high-SES boys imply that additional early schooling can contribute to closing the achievement gap between high- and low-SES boys. To investigate to what extent this is the case, we relate the exposure effects to the “initial” SES achievement gap. Because the exposure variable is centered around seven months of exposure, the coefficient on low SES picks up the achievement gap among children with seven months of

²⁷ See <http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html> for a description of the measure. This SES measure is based on a parent’s occupation when the child is five years old. If a parent is not working at that point, then his/her last known occupation from previous survey waves is used. We classify repeatedly unemployed parents, for whom no prior information on occupation is available, as long-term unemployed.

²⁸ In an international comparison, the United Kingdom occupies a medium place similar to the United States when it comes to the strength of the correlation between family background and educational achievement (e.g., Figures 1–4 in Waldinger 2007).

²⁹ Using aggregate scores allows us to present the pattern of results with multiple outcomes and multiple interactions in a compact way. Nevertheless, for comparison, we report in online Appendix Table A5 the exposure effects for ages 5, 7, and 11 on the disaggregated cognitive and noncognitive outcomes as in Tables 4 and 5 allowing for interactions of early school exposure with SES. The pattern of results for the disaggregated outcomes replicates the pattern of results for the aggregate scores in Table 6.

TABLE 6—EARLY EXPOSURE EFFECTS ON AGGREGATED OUTCOMES AT AGES 5, 7, AND 11, BY GENDER AND SES

	Cognitive factor			Noncognitive factor		
	Age 5	Age 7	Age 11	Age 5	Age 7	Age 11
<i>Panel A. Male</i>						
Exposure × high SES	0.030 (0.039)	−0.028 (0.032)	−0.010 (0.055)	0.015 (0.032)	0.038 (0.035)	0.089 (0.043)
Exposure × low SES	0.168 (0.051)	0.098 (0.049)	0.017 (0.057)	0.101 (0.048)	0.093 (0.043)	0.090 (0.062)
Low SES	−0.756 (0.153)	−0.840 (0.158)	−0.515 (0.172)	−0.492 (0.145)	−0.260 (0.134)	−0.093 (0.190)
<i>Panel B. Female</i>						
Exposure × high SES	0.091 (0.029)	−0.012 (0.037)	−0.003 (0.044)	0.076 (0.027)	0.008 (0.023)	0.013 (0.030)
Exposure × low SES	0.094 (0.034)	−0.008 (0.039)	0.049 (0.050)	0.054 (0.037)	−0.039 (0.029)	−0.045 (0.047)
Low SES	−0.311 (0.103)	−0.299 (0.103)	−0.557 (0.123)	−0.136 (0.076)	0.099 (0.080)	0.113 (0.147)
Observations	7,769	5,761	4,647	7,768	6,133	4,732

Notes: The table shows IV estimates of the effect of a one-month increase in exposure to the first school year on cognitive and noncognitive outcomes. The overall cognitive and noncognitive outcomes are constructed as common factors of the more disaggregated cognitive and noncognitive outcomes used in Tables 4 and 5. The exposure variable is centered around seven months of exposure. The coefficient on low SES thus captures the achievement gap between high and low SES children with seven months of exposure to the first school year. Control variables are dummies for gender, birth month, and local authority. Separate results for males and females are obtained by interacting all regressors with male/female dummies. Standard errors clustered at the level of the local authority are in parentheses.

Source: Millennium Cohort Study

exposure (i.e., who have two terms of schooling). For example, the coefficient of -0.492 on “low SES” for boys in the regression of the noncognitive outcome score at age 5 in Table 6 indicates an achievement gap between high- and low-SES boys of 49 percent of a standard deviation. At the same time, the returns to exposure on noncognitive outcomes of low-SES boys at age 5 exceed those of high-SES boys by about 8.5 percent of a standard deviation ($0.101-0.015$). An additional term (additional 4 months) of exposure would therefore reduce the SES gap in noncognitive outcomes at age 5 by about 34 percent of a standard deviation (4×8.5 percent), which amounts to almost three-quarters of the initial gap of 49 percent. At age seven, an additional four months of schooling would even almost close the SES achievement gap in noncognitive outcomes for boys. Analogous calculations for cognitive outcomes at age five and seven suggest a similar pattern, according to which additional four months of early schooling can close the initial SES achievement gap for boys by about two-thirds to three-quarters. At age 11, however, additional early schooling hardly affects the initial SES gap, which itself is generally smaller at that age than at the younger ages.

Our results for ages five and seven lend support to the hypothesis of Elder and Lubotsky (2009) that earlier exposure to a more formal school environment is beneficial for children from lower socioeconomic backgrounds. These authors show that the combined (positive) age-at-entry and age-at-test effect is smaller for children from disadvantaged backgrounds. Our evidence reinforces their findings by showing

that a direct substitution of time spent in the childcare environment that precedes school entry for time spent at school, irrespective of the age-at-test, is highly beneficial for children from the lower end of the family background distribution, not just for cognitive but also for noncognitive and behavioral outcomes. Our results also suggest that this mechanism is driven by effects on boys only because girls seem to benefit uniformly from early schooling. Furthermore, the stronger effects for low-SES boys persist up to age 7, and effects become more uniform across SES for boys at age 11.

IV. Discussion and Conclusions

In this paper, we investigate the effects of the length of exposure to early schooling before age 5 on cognitive and noncognitive outcomes at the ages of 5, 7, and 11. Our results show that, holding the age-at-test constant, receiving an additional month of early school exposure at age 4–5 at the expense of time spent in the counterfactual childcare environment increases test scores at ages 5 and 7 by about 6–11 percent and 1–3 percent of a test score standard deviation, respectively, but effects on test scores have largely faded away by age 11. While this seems to suggest that there is no benefit from additional early schooling for longer term cognitive development, we also show that the early test score effects are larger for low-SES boys, and that an additional term of early schooling reduces the achievement gap between low- and high-SES boys by 60–80 percent of its initial magnitude because of the higher differential returns to low-SES boys. Thus, even if the overall cognitive effects are temporary, closing early SES achievement gaps may have important implications. It is particularly relevant if early decisions about future school attendance are based on cognitive test scores—as, e.g., in the German tracking system, where tracking choices are made as early as grade four, and therefore well within the window where we find effects.

For noncognitive and behavioral outcomes, we find more persistent effects, at least up to age 11, the end of our observation window, and again evidence for stronger effects for low-SES boys. Much in line with findings by Chetty et al. (2011) or Heckman, Pinto, and Savelyev (2013) for the STAR experiment and the Perry Preschool project, respectively, our analysis thus suggests cognitive effects to be rather transitory, while noncognitive effects seem more persistent. Given that the birth cohort that we investigate has not yet left school, this conjecture defines an interesting agenda for future research.

The reason why boys from low-SES backgrounds have a stronger beneficial effect of additional schooling may be that their counterfactual outcome, when not being in school, is worse. Our finding of this pattern for boys but not for girls is in line with recent literature suggesting that girls are less affected by an adverse family background than boys. For example, family income seems to affect boys' educational outcomes more than girls' (see, e.g., Milligan and Stabile 2011), girls tend to perform better at school than boys despite being on average exposed to less favorable family backgrounds (Fortin, Oreopoulos, and Phipps 2015), and the noncognitive development of boys seems to be more harmed by social disadvantage, nontraditional family structures, or a lack of parental input than that of girls

(Bedard and Witman 2015, Bertrand and Pan 2013, Autor and Wasserman 2013, Autor et al. 2017, and Brenøe and Lundberg 2018).³⁰

If parents behave differently toward sons than toward daughters as, e.g., suggested by Baker and Milligan (2016), and if these differences in parental responses to child gender vary by SES, then they may explain why low-SES boys might have a worse counterfactual outcome when not enrolled in school, and thus a higher positive effect of additional early schooling. In online Appendix Table A4, we provide descriptive evidence showing that even though parental characteristics and behaviors differ markedly by SES, these differences across family background are very similar for boys and girls. This suggests that a worse counterfactual outcome for low-SES boys as compared to girls cannot be explained by differential parental behavior. The most plausible alternative explanation is thus that boys and girls *respond* differently to moving from a low-SES background to a more structured school environment, while parental behaviors toward them are similar.

Our findings are relevant for the debate over the optimal school starting age; that is, the concern that expanding universal schooling to ever earlier ages must necessarily have negative effects because school is simply not the right childcare environment for the very young. Our results show that the effect of additional early schooling at age four to five achieved by bringing the school starting age slightly forward is positive and has especially large effects up to age seven for boys from weaker socioeconomic backgrounds. It also has persistent effects on noncognitive skills until age 11 that are more uniform across SES for boys. This finding is particularly relevant from a US perspective where only about two-thirds of four-year-olds are enrolled in any educational pre-primary program (McFarland et al. 2017), and coverage to four-year-olds of the major public preschool programs has largely stalled at around 40 percent since 2010 (Barnett et al. 2016).

REFERENCES

- Anderson, Michael L.** 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.
- Anderson, Kathryn H., James E. Foster, and David E. Frisvold.** 2010. "Investing in Health: The Long-Term Impact of Head Start on Smoking." *Economic Inquiry* 48 (3): 587–602.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin.** 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.
- Autor, David, and Melanie Wasserman.** 2013. *Wayward Sons: The Emerging Gender Gap in Labor Markets and Education*. Third Way, March.
- Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman.** 2017. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." National Bureau of Economic Research (NBER) Working Paper 22267.

³⁰Developmental psychologists have noted that boys and girls may react differently to risk factors such as poverty, family breakup, and parental mental illness (see Werner 2000, Rutter 2000). They have also shown that during the first decade of life, boys are more vulnerable than girls to certain risk factors, including poverty and disharmony at home (Werner and Smith 1989, 1992), and that being raised by a single mother has stronger and more long-lasting adverse effects on boys (Hetherington, Stanley-Hagan, and Anderson 1989). There is also evidence that females benefit more from protective factors that lie within the individual (personality traits, cognitive skills), while males benefit more from protective factors provided by the environment. These latter include the structure, organization, and rule enforcement that can be provided at school, which has been identified as a stronger protective factor for boys than for girls (Werner 2000).

- Baker, Michael.** 2011. "Innis Lecture: Universal Early Childhood Interventions: What Is the Evidence Base?" *Canadian Journal of Economics/Revue canadienne d'économique* 44 (4): 1069–1105.
- Baker, Michael, Jonathan Gruber, and Kevin Milligan.** 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4): 709–45.
- Baker, Michael, and Kevin Milligan.** 2016. "Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries." *Journal of Human Capital* 10 (4): 399–441.
- Barnett, W. Steven, Allison H. Friedman-Krauss, Rebecca E. Gomez, Michelle Horowitz, G.G. Weisfeld, Kirsty Clarke Brown, and James H. Squires.** 2016. *The State of Preschool 2015: State Preschool Yearbook*. National Institute for Early Education Research. New Brunswick: NJ.
- Bedard, Kelly, and Elizabeth Dhuey.** 2006. "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects." *Quarterly Journal of Economics* 121 (4): 1437–72.
- Bedard, Kelly, and Allison Witman.** 2015. "Family Structure and the Gender Gap in ADHD." https://paa.confex.com/paa/2016/mediafile/ExtendedAbstract/Paper4409/ADHD_1-16.pdf.
- Berlinski, Samuel, Sebastian Galiani, and Paul Gertler.** 2009. "The Effect of Pre-primary Education on Primary School Performance." *Journal of Public Economics* 93 (1–2): 219–34.
- Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda.** 2008. "Giving Children a Better Start: Preschool Attendance and School-Age Profiles." *Journal of Public Economics* 92 (5–6): 1416–40.
- Bernal, Raquel, and Michael P. Keane.** 2010. "Quasi-structural Estimation of a Model of Childcare Choices and Child Cognitive Ability Production." *Journal of Econometrics* 156 (1): 164–89.
- Bernal, Raquel, and Michael P. Keane.** 2011. "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers." *Journal of Labor Economics* 29 (3): 459–512.
- Bertrand, Marianne, and Jessica Pan.** 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1): 32–64.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes.** 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." *Review of Economics and Statistics* 93 (2): 455–67.
- Blanden, J., E. Del Bono, K. Hansen, S. McNally, and B. Rabe.** 2014. "Evaluating a Demand-Side Approach to Expanding Free Preschool Education." https://www.nuffieldfoundation.org/sites/default/files/files/Childoutcomes_final.pdf.
- Blau, David, and Janet Currie.** 2006. "Pre-school, Day Care, and After-School Care: Who's Minding the Kids?" In *Handbook of the Economics of Education*, Vol. 2, edited by E. Hanushek and F. Welch, 1163–1278. Amsterdam: Elsevier.
- Blundell, Richard, Lorraine Dearden, and Luke Sibietta.** 2010. "The Demand for Private Schooling in England: The Impact of Price and Quality." Institute for Fiscal Studies (IFS) Working Paper 10/21.
- Brenøe, Anne Ardila, and Shelly Lundberg.** 2018. "Gender Gaps in the Effects of Childhood Family Environment: Do They Persist into Adulthood?" *European Economic Review* 109: 42–62.
- Carlsson, Magnus, Gordon B. Dahl, Björn Öckert, and Dan-Olof Rooth.** 2015. "The Effect of Schooling on Cognitive Skills." *Review of Economics and Statistics* 97 (3): 533–47.
- Carneiro, Pedro, and Rita Ginja.** 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6 (4): 135–73.
- Cascio, Elizabeth U.** 2009. "Do Investments in Universal Early Education Pay Off? Long-Term Effects of Introducing Kindergartens into Public Schools." National Bureau of Economic Research (NBER) Working Paper 14951.
- Cascio, Elizabeth U., and Ethan G. Lewis.** 2006. "Schooling and the Armed Forces Qualifying Test: Evidence from School-Entry Laws." *Journal of Human Resources* 41 (2): 294–318.
- Cascio, Elizabeth U., and Diane Whitmore Schanzenbach.** 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity* 43 (2): 127–78.
- Cascio, Elizabeth U., and Douglas O. Staiger.** 2012. "Knowledge, Tests, and Fadeout in Educational Interventions." National Bureau of Economic Research (NBER) Working Paper 18038.
- Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan.** 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *Quarterly Journal of Economics* 126 (4): 1593–1660.
- Cornelissen, Thomas, and Christian Dustmann.** 2019. "Early School Exposure, Test Scores, and Noncognitive Outcomes: Dataset." *American Economic Journal: Economic Policy*. <https://doi.org/10.1257/pol.20170641>.
- Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg.** 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy* 126 (6): 2356–2409.
- Crawford, Claire, Lorraine Dearden, and Costas Meghir.** 2007. "When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England." Centre for the Economics of Education (CEE), London School of Economics.

- Cunha, Flavio, and James Heckman.** 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.
- Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov.** 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, Vol. 1, edited by E. Hanushek and F. Welch, 697–812. Amsterdam: Elsevier.
- Currie, Janet.** 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15 (2): 213–38.
- Currie, Janet, and Douglas Almond.** 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1315–1486. Amsterdam: Elsevier.
- Currie, Janet, and Duncan Thomas.** 1995. "Does Head Start Make a Difference?" *American Economic Review* 85 (3): 341–64.
- Currie, Janet, and Duncan Thomas.** 1999. "Early Test Scores, Socioeconomic Status and Future Outcomes." National Bureau of Economic Research (NBER) Working Paper 6943.
- Datar, Ashlesha.** 2006. "Does Delaying Kindergarten Entrance Give Children a Head Start?" *Economics of Education Review* 25 (1): 43–62.
- Datta Gupta, Nabanita, and Marianne Simonsen.** 2010. "Non-cognitive Child Outcomes and Universal High Quality Child Care." *Journal of Public Economics* 94 (1–2): 30–43.
- Dee, Thomas S., and Hans Henrik Sievertsen.** 2018. "The Gift of Time? School Starting Age and Mental Health." *Health Economics* 27 (5): 781–802.
- Del Bono, Emilia, and Fernando Galindo-Rueda.** 2006. "The Long Term Impacts of Compulsory Schooling: Evidence from a Natural Experiment in School Leaving Dates." University of Essex Institute for Social and Economic Research (ISER) Working Paper 2006-44.
- Dex, Shirley, and Heather Joshi, eds.** 2005. *Children of the 21st Century: From Birth to Nine Months*. Bristol: Bristol University Press.
- Drange, Nina, Tarjei Havnes, and Astrid M. J. Sandsør.** 2016. "Kindergarten for All: Long Run Effects of a Universal Intervention." *Economics of Education Review* 53: 164–81.
- Elango, Sneha, Jorge Luis García, James J. Heckman, and Andrés Hojman.** 2016. "Early Childhood Education." In *Economics of Means-Tested Transfer Programs in the United States*, Vol. 2, edited by Robert A. Moffitt, 235–97. Chicago: University of Chicago Press.
- Elder, Todd E., and Darren H. Lubotsky.** 2009. "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers." *Journal of Human Resources* 44 (3): 641–83.
- Fertig, Michael, and Jochen Kluge.** 2005. "The Effect of Age at School Entry on Educational Attainment in Germany." IZA Institute of Labor Economics Discussion Paper 1507.
- Fort, Margherita, Andrea Ichino, and Giulio Zanella.** 2016. "Cognitive and Non-cognitive Costs of Daycare 0-2 for Girls." Centre for Economic Policy Research (CEPR) Discussion Paper 11120.
- Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps.** 2015. "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *Journal of Human Resources* 50 (3): 549–79.
- Fredriksson, Peter, and Björn Öckert.** 2014. "Life-Cycle Effects of Age at School Start." *Economic Journal* 124 (579): 977–1004.
- Garces, Eliana, Duncan Thomas, and Janet Currie.** 2002. "Longer-Term Effects of Head Start." *American Economic Review* 92 (4): 999–1012.
- Gillard, Derek.** 2018. *Education in England: A History*. www.educationengland.org.uk/history.
- Gormley, William T., Jr., and Ted Gayer.** 2005. "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program." *Journal of Human Resources* 40 (3): 533–58.
- Havnes, Tarjei, and Magne Mogstad.** 2011. "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy* 3 (2): 97–129.
- Havnes, Tarjei, and Magne Mogstad.** 2015. "Is Universal Child Care Leveling the Playing Field?" *Journal of Public Economics* 127: 100–114.
- Heckman, James J.** 2008. "Schools, Skills, and Synapses." *Economic Inquiry* 46 (3): 289–324.
- Heckman, James, Seong Hyeok Moon, Rodrigo Pinto, Peter Savelyev, and Adam Yavitz.** 2010a. "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the HighScope Perry Preschool Program." *Quantitative Economics* 1 (1): 1–46.
- Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz.** 2010b. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1–2): 114–28.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev.** 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–86.

- Hetherington, E. Mavis, Margaret Stanley-Hagan, and Edward R. Anderson.** 1989. "Marital Transitions: A Child's Perspective." *American Psychologist* 44 (2): 303–12.
- Imbens, Guido W.** 2014. "Instrumental Variables: An Econometrician's Perspective." *Statistical Science* 29 (3): 323–58.
- Imbens, Guido W., and Joshua D. Angrist.** 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.
- Inoue, Atsushi, and Gary Solon.** 2010. "Two-Sample Instrumental Variables Estimators." *Review of Economics and Statistics* 92 (3): 557–61.
- Lamb, Michael E., and Lieselotte Ahnert.** 2007. "Nonparental Child Care: Context, Concepts, Correlates, and Consequences." In *Handbook of Child Psychology*, Vol. 4, edited by K. Ann Renninger and Irving E. Sigel, 950–1016. Hoboken: John Wiley and Sons.
- Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen.** 2017. "School Starting Age and the Crime-Age Profile." *Economic Journal* 127 (602): 1096–1118.
- Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink.** 2010. "Expanding Schooling Opportunities for 4-Year-Olds." *Economics of Education Review* 29 (3): 319–28.
- Loeb, Susanna, Margaret Bridges, Daphna Bassok, Bruce Fuller, and Russell W. Rumberger.** 2007. "How Much Is Too Much? The Influence of Preschool Centers on Children's Social and Cognitive Development." *Economics of Education Review* 26 (1): 52–66.
- Ludwig, Jens, and Douglas L. Miller.** 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122 (1): 159–208.
- Magnuson, Katherine A., Christopher Ruhm, and Jane Waldfogel.** 2007. "Does Prekindergarten Improve School Preparation and Performance?" *Economics of Education Review* 26 (1): 33–51.
- Masse, Leonard N., and W. Steven Barnett.** 2002. "A Benefit Cost Analysis of the Abecedarian Early Childhood Intervention." National Institute for Early Education Research. <https://files.eric.ed.gov/fulltext/ED479989.pdf>.
- McEwan, Patrick J., and Joseph S. Shapiro.** 2008. "The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates." *Journal of Human Resources* 43 (1): 1–29.
- McFarland, Joel, Bill Hussar, Cristobal de Brey, Tom Snyder, Xiaolei Wang, Sidney Wilkinson-Flicker, Semhar Gebrekristos, et al.** 2017. *The Condition of Education 2017*. National Center for Education Statistics (NCES) 2017-144. US Department of Education Institute of Education Sciences. Washington, DC, May.
- Milligan, Kevin, and Mark Stabile.** 2011. "Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions." *American Economic Journal: Economic Policy* 3 (3): 175–205.
- Mühlenweg, Andrea M., and Patrick A. Puhani.** 2010. "The Evolution of the School-Entry Age Effect in a School Tracking System." *Journal of Human Resources* 45 (2): 407–38.
- Oreopoulos, Philip.** 2006. "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter." *American Economic Review* 96 (1): 152–75.
- Organisation for Economic Co-operation and Development (OECD).** 2008. *Education at a Glance 2008: OECD Indicators*. Organisation for Economic Co-operation and Development.
- Puhani, Patrick A., and Andrea M. Weber.** 2007. "Does the Early Bird Catch the Worm?" *Empirical Economics* 32 (2–3): 359–86.
- Ruhm, Christopher, and Jane Waldfogel.** 2012. "Long-Term Effects of Early Childhood Care and Education." *Nordic Economic Policy Review* 1: 23–51.
- Rutter, Michael.** 2000. "Resilience Reconsidered: Conceptual Considerations, Empirical Findings, and Policy Implications." In *Handbook of Early Childhood Intervention*. 2nd ed., edited by Jack P. Shonkoff and Samuel J. Meisels, 651–82. Cambridge: Cambridge University Press.
- Schweinhart, L.J., J. Montie, Z. Xiang, W.S. Barnett, C.R. Belfield, and M. Nores.** 2005. *Lifetime Effects: The HighScope Perry Preschool Study through Age 40*. Ypsilanti: HighScope Press.
- Speight, Svetlana, Ruth Smith, Ivana La Valle, Vera Schneider, Jane Perry, Cathy Coshall, and Sarah Tipping.** 2009. "Childcare and Early Years Survey of Parents 2008." National Centre for Social Research Department for Children, Schools and Families Research Report DCSF-RR136.
- Stock, James H., and Motohiro Yogo.** 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, edited by Donald W.K. Andrews and James H. Stock, 80–108. New York: Cambridge University Press.
- UK Department of Education.** 2008. "Practice Guidance for the Early Years Foundation Stage." Department for Children, Schools and Family. Manchester, May.

- US Department of Health and Human Services.** 2010. *Head Start Impact Study: Final Report: Executive Summary*. Office of Planning, Research and Evaluation, Administration for Children and Families. Washington, DC, January.
- Waldinger, Fabian.** 2007. "Does Ability Tracking Exacerbate the Role of Family Background for Students' Test Scores?" <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.9069&rep=rep1&type=pdf>.
- Werner, Emmy E.** 2000. "Protective Factors and Individual Resilience." In *Handbook of Early Childhood Intervention*. 2nd ed., edited by Jack P. Shonkoff and Samuel J. Meisels, 115–32. Cambridge: Cambridge University Press.
- Werner, Emmy E., and Ruth S. Smith.** 1989. *Vulnerable but Invincible: A Longitudinal Study of Resilient Children and Youth*. New York: Adams Bannister Cox Pubs.
- Werner, Emmy E., and Ruth S. Smith.** 1992. *Overcoming the Odds: High Risk Children from Birth to Adulthood*. Ithaca: Cornell University Press.