# Birth Order, Early Childhood Development and Adolescent Outcomes: Evidence from Administrative Data *

Juan-Manuel Del Pozo Iftikhar Hussain

September 2026

**Abstract**

How does the family environment shape early childhood cognitive and noncognitive skills? We leverage unique data for the universe of students in England's public education sector, containing detailed teacher-assessed measures of age-five cognitive, socioemotional, and physical development skills, to provide new evidence on early childhood birth order effects. We first demonstrate that these cognitive and noncognitive age-five skills are strongly associated with subsequent high stakes exam performance as well as behavioral outcomes. For the majority white British population, firstborn siblings experience a modest advantage on age-five cognitive skills. There is no birth order impact on socio-emotional skills; this null effect is precisely estimated. However, a substantial firstborn advantage emerges through adolescence. For closely spaced white British siblings, on the other hand, firstborn children experience a *disadvantage* on age-five skills. Strikingly, this age-five penalty reverses by age-11. We discuss the implications of these results for leading theoretical mechanisms put forward to explain birth order effects. For minority students, firstborn children experience large penalties in age-five skills, which also reverse through adolescence. These birth order trajectories offer novel insights into the evolution of the skills gap between minority and majority groups of students. Our findings establish that the impact of birth order on age-five outcomes varies by skill domain, family structure, and ethnicity, and its effects can reverse between early childhood and adolescence.

* Del Pozo: Departamento de Economía, Pontificia Universidad Católica del Perú, jmdelpozo@pucp.pe ; Hussain: Department of Economics, University of Sussex, iftikhar.hussain@sussex.ac.uk . This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners. Department for Education (DfE) data is used in this publication. The DfE does not accept responsibility for any inferences or conclusions derived from the DfE Data Extracts by third parties. For helpful comments and discussions we thank Amalavoyal Chari, Ellen Greaves and seminar participants at the universities of Stellenbosch, Sussex and York.

# 1 Introduction

Families are a primary driver of children's early childhood and later outcomes (Becker and Tomes, 1976; Cunha and Heckman, 2007; Currie and Almond, 2011). Within the family, birth order (an individual's ordinal position among siblings within the family) has long been recognised as an important feature of family structure that can generate inequality between siblings (Galton, 1874; Adler, 1928; Behrman and Taubman, 1986; Sulloway, 1996). Recent empirical literature on birth order effects from developed countries has leveraged large administrative datasets, documenting large, systematic birth order gradients in outcomes. On average, firstborn and earlier-born siblings outperform later-born ones on education attainment and labor market outcomes, as well as exhibiting improved noncognitive and behavioral outcomes, such as lower levels of adult crime (Black et al., 2005; Black et al., 2018; Breining et al., 2020). Yet, little is known about the role of birth order on early childhood skills, especially noncognitive domains. This study seeks to fill this gap.

Mechanisms put forward to account for birth order effects include the resource dilution hypothesis which proposes that as family size grows, parental time and other inputs for each child fall, generating systematic within-family differences by birth order (Blake, 1981; Black et al., 2005; Price, 2008). Alternative hypotheses include variants of different types of family spillover effects, including the worsening of the intellectual environment with the arrival of later-born siblings, as well as the benefits for early-born siblings from teaching their younger siblings (Zajonc and Markus, 1975). Recent research by Daysal et al. (2026) provides novel evidence of disease spillovers from the firstborn to the secondborn sibling in early childhood, which accounts for some of the firstborn advantage in adult outcomes.

Notably, the absence of any firstborn advantage in neonatal health indicators suggests that the observed birth order effect is not a product of initial biological endowments, but rather 'nurture', i.e. the postnatal environment and parental investment, is the primary driver. [1] This leads to a critical empirical question: at what point in the early childhood

[1] Firstborn infants exhibit systematically *lower* birth weight and Apgar scores compared to their later-born siblings, and a higher probability of being born prematurely (Brenøe and Molitor, 2018; Pruckner et al., 2021). Furthermore, studies find no systematic genetic differences by birth order (Isungset et al., 2022; Muslimova et al., 2026).

developmental trajectory does the firstborn advantage in cognitive, noncognitive and behavioral skills begin to emerge?

Progress in understanding the role that birth order plays in determining early childhood outcomes is hampered by the limited availability of large datasets with early childhood measures. This study leverages administrative data which track the universe of students in the English public school system from school enrollment at age four and five in the 'Reception' grade through the end of secondary school. The data include unique early childhood measures across multiple domains, including cognitive, noncognitive and physical development skills, collected by the Reception teacher, based on classroom observations. This allows for an analysis of whether birth order effects differ across language, numeracy and socio-emotional skills, as well as motor skills and behavioral outcomes. The data facilitate the estimation of dynamic birth order effects, revealing whether early childhood birth order effects are present, and whether they persist, attenuate, or widen as children progress through key developmental stages. Uncovering these effects is important, especially if skill formation is inherently dynamic. If early capabilities foster subsequent human capital accumulation via self-productivity and dynamic complementarity, then a given early childhood gap in skills between siblings can plausibly shrink or grow depending on subsequent inputs and the nature of the skill production function (Cunha et al., 2010). Moreover, many early interventions exhibit fade-out along some dimensions, while other domains show delayed impact (Duncan et al., 2023; Gray-Lobe et al., 2023; Heckman et al., 2013) which underscores the importance of tracing birth order trajectories across domains and over time.

The Reception grade is a unique phase of a child's formal schooling in England, marked by a particular emphasis on play-based learning. The Reception year in primary school is followed by Grade 1 (referred to as Year 1 in England), at which point students embark on a more traditional, academically focused curriculum. The theoretical basis for the Reception curriculum reflects a mix of developmental, social and educational theories (Evangelou et al., 2009). It emphasizes collaborative play, guided learning and 'scaffolding', highlighting a 'key person' approach, ensuring children have a primary carer (teacher) to build trust and emotional security. [2]

[2] Evangelou et al. (2009), and Siraj-Blatchford et al. (2002), provide references to the relevant developmental literature and its relation to the Foundation Stage, which incorporates the Reception grade. Siraj-Blatchford et

These administrative data also permit rich demographic analysis of early childhood outcomes given the diversity of the student population, including variation by ethnic background and immigration histories across different ethnic cohorts, socioeconomic status of the families, as well as gender composition of siblings.[3] This allows us to investigate different potential mechanisms driving birth order effects, including the role of culture, gender interactions and socioeconomic circumstance. Additionally, we exploit school-level information to examine other potential important mediating channels, including key parental inputs, as captured by school quality, peer composition, as well as deferred school enrollment.[4] Specifically, we examine how birth order effects change after accounting for schooling inputs, which may evolve over time between age five and adolescence.

Our first finding is that the mechanisms driving the early childhood results differ markedly between white British majority students and minority students. Hence results are discussed separately for these two groups. The core age-five results for the white British population demonstrate that relative to their secondborn siblings, firstborn siblings experience an overall mean advantage of 0.05 standard deviation for language skills and a more modest advantage for numeracy skills. There is no evidence of any difference between firstborn and secondborn siblings on socioemotional skills. This latter result is surprising because the prior literature demonstrates that for adolescents and adults, firstborns enjoy substantial advantages on noncognitive outcomes, (Black et al., 2018; Houmark, 2023). We also demonstrate that age-five socioemotional skills strongly predict later education attainment as well as behavioral outcomes, so that measurement error in socioemotional skills alone cannot account for this null finding.

These overall effects for the white British population mask substantial heterogeneity by birth spacing, with the firstborn advantage in cognitive skills generally rising substantially as spacing between siblings increases. For example, for siblings who are three years apart, the firstborn advantage relative to secondborn siblings is 0.1 and 0.05 standard deviation in

al. (2002) also highlight important lessons learned from influential experiments such as the Perry Preschool program (Heckman et al., 2013).

[3] The Early Years Foundation Stage Profile administrative data are also described in Cornelissen and Dustmann (2019).

[4] Parents may undertake investments which compensate or reinforce differences among their children (Almond and Mazumder, 2013). The evidence is mixed, with some studies finding compensatory behavior (Bharadwaj et al., 2018; Figlio et al., 2014), while other studies find reinforcing responses (Aizer and Cunha, 2012; Behrman et al., 1994; Datar et al., 2010), and still others finding a mix of patterns depending on the specific domain of human capital under consideration (Grätz and Torche, 2016; Yi et al., 2015).

language and numeracy skills, respectively. For noncognitive skills, the effects are relatively small (0.01 and 0.03 standard deviation for socioemotional and creative development skills, respectively, for siblings with birth spacing of three years). On the other hand, for close spaced siblings (born approximately one year apart, but not in the same grade at school), firstborn children experience a modest *disadvantage* on all age-five skill domains. [5]

In contrast to the headline results for the white British population, for minority children, who constitute over a fifth of the student population, there are very large *negative* penalties for firstborn siblings relative to secondborns across all five early childhood skill domains, ranging from 0.1 standard deviation to 0.2 standard deviation. This pattern holds consistently across diverse minority heritage groups, including students from South Asian, African, and Eastern European backgrounds, indicating that the underlying mechanism driving these birth order developmental gaps operates across cultural boundaries.[6] Further investigation suggests that these effects may be driven by positive English language spillover effects from the older sibling to the younger one.[7]

A key goal of our analysis is to characterize the dynamic evolution of birth order effects, specifically, to determine whether early developmental gaps persist, amplify, or fade as children transition into adolescence. In order to do so, we first corroborate findings from the prior literature demonstrating firstborn advantages on adolescent outcomes: firstborn siblings (from white British majority as well as minority families) have higher age-11 test scores, perform better on age-16 high stakes exams, and have lower school suspension rates. Overall, our results demonstrate that the birth order advantage for firstborn siblings increases between early childhood and adolescence.

Strikingly, we uncover two examples of a *reversal* in birth order effects between early childhood and adolescence. The first case is for relatively closely spaced white British siblings (siblings born approximately one year apart): the firstborn penalty on age-five skills for this group fully reverses by age 11, with first born siblings outperforming their younger siblings on education outcomes, which is sustained through to age-16. The second example of such a reversal is with respect to minority students: the large negative effects for age-five skills for

[5] Note that because birth spacing is endogenous, such variation provides descriptive evidence of heterogeneity in birth order effects.
[6] On the interaction of birth order and cultural norms, see Jayachandran and Pande (2017).
[7] This contrasts with the negative diseases spillover effects for earlier-born siblings identified by Daysal et al. (2026).

firstborn sibling are not sustained, so that firstborn minority siblings perform better than their secondborn counterparts on high-stakes age-16 exams and are less likely to be suspended from school.

Our results have important implications for theoretical explanations of birth order effects. In their simplest form, resource dilution explanations imply that firstborn children receive greater parental inputs in early childhood and should therefore exhibit advantages, relative to secondborns, across both cognitive and noncognitive skills at age five.[8] The fact that we find null effects as well as deficits, which subsequently turn into advantages during later childhood, suggests that resource dilution alone is unlikely to provide a complete explanation. Similar arguments apply to alternative explanations, such as those based on sibling spillovers, or evolving parenting style.[9] Our analysis allows us to rule out important classes of explanations based on differential inputs, including differences in deferred school enrollment, school quality, and peer quality. The estimated birth order effects are largely unchanged after accounting for these factors.

Our primary contribution is to two distinct literatures. First, we contribute to the literature on the impact of birth order on individual outcomes (Black et al., 2005; de Haan, 2010; Pavan, 2016; Lehmann et al., 2018; Breining et al., 2020; Bagger et al., 2021; Daysal et al., 2026; Muslimova et al., 2026). Second, our findings speak to the literature on the role of the family environment in early childhood skill formation. While a large literature documents the importance of parental time and resource investments for cognitive and noncognitive child development (Almond et al., 2018; Duncan et al., 2023; Heckman and Mosso, 2014), considerably less is known about how family structure, including a child's position within the family, shapes these early skills.[10] Estimating the reduced-form impact of birth order on early childhood outcomes is crucial for a better understanding of the early origins of skill gaps and the process of skill acquisition. For example, the prior evidence demonstrates that birth order

[8] On the impact of early childhood parental investment, see, for example, Doyle (2020), demonstrating that an intensive parenting program from pregnancy until age five has substantial effects on both cognitive as well as socioemotional and behavioral outcomes observed through age five.

[9] Prior evidence suggests that parental inputs of resources and time favor earlier-born children (Price, 2008; Averett et al., 2011; Pavan, 2016; Lehmann et al., 2018; Black et al., 2018). Breining et al. (2020) document that maternity leave also favors firstborns relative to secondborn children. Alternative explanations for birth order gradients emphasize sibling spillovers, including siblings learning from and teaching each other (Zajonc and Markus, 1975), competition and rivalry between siblings (Sulloway, 1996), as well as evolving parenting style by birth order driven by a reputation model for strategic parenting (Hotz and Pantano, 2015).

[10] On the impact of divorce on children's adult outcomes, see Johnston et al. (2025).

affects adolescent and adult noncognitive abilities and delinquency behavior (Black et al., 2018; Breining et al., 2020), while a parallel literature points to early childhood noncognitive skills as potentially important mediators of later behavioral outcomes (Heckman et al., 2013). It is therefore important to determine whether the relationship between birth order and later behavioral outcomes arises via early childhood skills.[11] This paper uncovers an important set of findings, demonstrating that birth order is an important driver of early childhood skills. Significantly, these effects are not uniform, varying by skill domain, demographic and ethnic groups, and within-family factors, including birth spacing. We find that the firstborn advantage grows substantially between age five and adolescence. We also uncover striking heterogeneous effects where there is either no firstborn advantage or a firstborn deficit at age five, but where a strong firstborn advantage subsequently appears during adolescence. We are not aware of any prior evidence of such a reversal in the birth order gradient for cognitive or noncognitive skills between early childhood and adolescence.

Recent research has also focused on how policy interventions and exogenous shocks have heterogenous effects by birth order. For example, Doyle (2000) finds that a parenting intervention has a stronger impact for firstborns than non-firstborns on some early childhood cognitive and noncognitive outcomes. Ginja et al. (2020) find that an expansion of parental leave benefits in Sweden exacerbates birth order gradients. Black et al. (2021) demonstrate that the birth order gradient between firstborn and secondborn children strengthens when the thirdborn child is disabled. Our findings complement this literature by showing that the relationship between birth order and child outcomes varies substantially over childhood, suggesting that such differential effects of polices and family shocks may depend not only on birth order but also on the developmental stage at which children are exposed.

In addition, our data provide unusually rich measures of children's skills at age five, which allows us to shed new light on important debates in the birth order and early childhood

[11] Prior studies investigating the links between birth order and early childhood development have lacked the statistical power to perform the detailed investigations needed to map developmental trajectories from early childhood to adolescence across skills and demographic groups. For example, both Lehmann et al. (2018) and Pavan (2016) use the Children of the National Longitudinal Survey of Youth 1979 to analyze birth order effects on early outcomes. Pavan focuses on cognitive skills and finds negative birth order effects for language skills, but no statistically significant effect for math skills. Lehmann et al. (2018) construct two broad index measures of cognitive and noncognitive skills. They find a negative birth order effect for the cognitive skill index at ages 0-3 and 4-6, but find no evidence of a birth order impact on noncognitive skills, though the estimates are imprecise and therefore cannot rule out economically meaningful effects. Similar findings are reported in Rohrer et al. (2015) who also undertake within-family analysis using the German Socio-Economic Panel data as well as the NLSY 1997.

development literature. For example, with respect to age-five physical development, we demonstrate that firstborns *underperform* relative to their secondborn siblings, both for the white British as well as minority groups. This finding aligns with the established literature documenting inferior neonatal outcomes (such as birthweight) for firstborns (Brenøe and Molitor, 2018; Pruckner et al., 2021), which in turn suggests that the initial biological and physical disadvantage associated with being earlier-born may not be fully attenuated in the early years, and persists at least through age five. In relation to age-five creative development, our results demonstrate that for white British children, firstborn siblings outperform secondborn siblings on this measure. This finding speaks to the debate concerning the role of birth order and creativity and individuals' openness to experience. Sulloway (1996) argues that firstborn siblings identify with parental authority and power, whilst later-born children are more open to new ideas, experience and creativity. Black et al. (2018) on the other hand, find no overall association between birth order and the likelihood of sorting into creative occupations. In contrast, our results point to a creativity advantage for firstborns in the early stages of childhood.

Finally, our study also contributes to the literature on the education performance of first- and later-generation children from immigrant families (equivalent to minority students in our setting). The prior literature in this area has emphasized intergenerational education transmission (Borjas 1993; Card 2005; Casey and Dustmann, 2008) as well as critical periods for host-language acquisition (Bleakley and Chin, 2004, 2010). To our knowledge, we are the first to document the distinctive role of birth order for minority children in early childhood, showing that birth order effects differ markedly between the minority and the majority group of children. We document a dramatic reversal of the birth order effect between age five and adolescence, suggesting that within the family, language skill advantages enjoyed by the younger sibling at age five may be subject to dynamic substitutability for this group of students, as firstborn siblings catch up with, and subsequently overtake secondborn siblings on age-specific skills. Furthermore, the birth order patterns we identify provide important clues regarding the dramatic gains made by minority children in closing the large skill gaps relative to white British children between age five and age 11.

# 2 Context and Data

Children in the English public (state) school system typically start formal schooling in the Reception grade at age four, during the academic year in which they turn five. The Reception grade follows a distinct pedagogical framework called the Early Years Foundation Stage, which differs significantly from the more formal, subject-based curriculum starting from Grade 1 ('Year 1'). The Reception curriculum is designed to support children's early development across learning, language, behavior, and social adjustment, while also easing their transition into more formal teaching in grade 1. The focus is on activities such as guided play, storytelling, songs and shared reading. As described in Evangelou et al. (2009) and DCSF (2008, the Reception curriculum is closely aligned with a developmental perspective which emphasizes that play and active engagement with the social environment are key to young children's socioemotional, cognitive and linguistic growth. By the end of the Reception year, teachers complete the Early Years Foundation Stage Profile for each child, a teacher-assessed summary of children's development. These are based on teachers' observations of children's day-to-day activities, gathered throughout the year from a wide range of learning and teaching contexts. The areas for assessment are as follows: Communication, Language and Literacy (spoken language, comprehension, reading, and writing); Problem Solving, Reasoning and Numeracy (counting, calculation and mathematical reasoning); Personal, Social, and Emotional Development (including self-regulation, building relationships, and managing self-care); Creative Development (the cultivation of imagination and self-expression across artistic, musical and physical activities); Physical Development (movement, coordination, gross and fine motor skills); and Knowledge and Understanding of the World (early understanding of nature, objects, and everyday phenomena). Children are scored on individual items within each domain. Online Appendix D lists all the individual items for each domain. For the period under investigation, the local authority (municipality) takes the lead in the moderation process in an effort to ensure that assessments are undertaken on a consistent basis. This includes training teachers undertaking the assessments as well as carrying out regular visits in schools to ensure that teachers follow best practice (QCA, 2003).

## *2.1 Data and Sample Construction*

The main administrative data source for our analysis is England's National Pupil Database. These data include information on all four- and five-year old children in the public education system, enrolled in the Reception grade, as well as higher grades in the primary and secondary education system. Teacher assessments of children's cognitive and noncognitive skills in the Reception grade are available from 2007 (i.e. academic year 2006/2007) and were measured on a consistent basis through to 2012. The family identifier variable in the National Pupil Database is available for the years 2008, 2013 and 2016. This variable is constructed by the administrators by matching students with the same last name and the same full home address. [12]

The starting point for constructing the estimation sample is the set of students enrolled in the Reception grade in the years 2007 to 2012. Since the early childhood measures are available on a consistent basis in the period 2007 to 2012, children with an older sibling in the Reception grade in 2006 or earlier are excluded from the estimation sample. For example, consider students who enter Reception in 2007 and subsequently transition to grade 1 in 2008. The sample is restricted to those students in grade 1 in 2008 who do not have an older sibling in a higher grade (up to grade 11, the final year of compulsory formal academic study) and who *do* have at least one younger sibling enrolled in the Reception grade in years 2008 to 2012 (information on month and year of birth of each child is available for all years). [13] For this case, the cohort or grade spacing gap between siblings in the estimation sample ranges from one to five years. [14] Online Appendix B provides further details of the construction of the sample. Appendix B also reports analysis form a robustness exercise assessing the sensitivity of the core results to the possibility that the eldest sibling is incorrectly coded.

The focus of this study is on siblings of birth order one and two, hence the estimation sample consists of sibling pairs (as in, for example, Breining et al., 2020). Prior large-scale

[12] For their analysis of individuals' interaction with the juvenile justice system in Florida, Breining et al. (2020) adopt a similar approach, determining whether children are siblings using information on co-residency and the same last name. In a robustness exercise, we also utilize a sibling identifier using just the full home address. Results are unaffected by this sample selection rule.

[13] For this 2007 Reception cohort of students, they will be incorrectly coded as the oldest sibling if the gap between the younger sibling and their next older sibling is more than ten years.

[14] Twins (siblings born in the same year and month) and siblings in the same grade are dropped from the sample.

studies suggest that the gradient is either strongest at birth orders one and two or, alternatively, is linear in birth order (Black et al., 2005; Black et al., 2018). In the robustness analysis below, the sensitivity of the findings is assessed by exploiting information on higher birth orders. With regards to lack of information on maternal completed fertility, for the family fixed effect strategy this is not a concern because all time-invariant family characteristics are accounted for.[15]

With respect to age-five skills, we focus on two cognitive skill measures, language and mathematics development (Communication, Language and Literacy; and Problem Solving, Reasoning and Numeracy, respectively), two noncognitive skills, socioemotional and creative development (Personal, Social, and Emotional Development; and Creative Development, respectively), as well as physical development. In the robustness analysis below we also report results for the total Early Years Foundation score. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

Age-five outcomes for the 2007 – 2012 Reception grade cohorts are linked to their end-of-primary school (grade 6) age-11 'Key Stage 2' assessments, measured in the years 2013 – 2018.[16] These age-11 standardized national tests are externally marked. For the age-11 assessments, we focus on reading comprehension and mathematics test scores, which are assessed on a consistent basis over the period of our analysis. Age-16 exam outcomes are for the high-stakes 'GCSE' exams. For these age-16 outcomes, the six cohorts enrolled in grade 11 from 2014 to 2019 are used to construct the siblings sample. See Appendix A for further details.

Table 1 presents descriptive statistics for the cohorts of students enrolled in the Reception grade between 2007 and 2012, for the full sample, the estimation sample, as well as the estimation sample disaggregated by white British versus minority status. Table 1 shows that, compared to the full sample, students in the estimation sample are somewhat less

[15] Although we do not observe whether siblings in our data are full biological siblings or not, nor do we observe family structure information such as partner separation or divorce, existing evidence suggests that negative birth order gradients in educational outcomes are not confined to full biological siblings (Barclay, 2015). In our analysis below, we assess how our results change with respect to one measure of family stability we do observe in our data: whether or not the older and younger sibling are enrolled in the same school at the time of their Reception grade.

[16] Note that 2019 is the last year that assessments and test score data are available before the COVID-19 pandemic disrupted the education system and formal testing arrangements. Students typically progress from one grade to the next each academic year in the English system and grade retention is extremely rare.

disadvantaged, as measured by free lunch status, less likely to be from a minority background, and have higher age-five and age-11 test scores.

Table 1 demonstrates that there is substantial within family variation in the teacher-assessed age-five skill measures, indicating that teachers meaningfully differentiate between siblings' abilities and development rather than simply assigning them similar scores. It is also noteworthy that the within family standard deviations for age-five cognitive skills (language and numeracy) are of a similar magnitude as those for noncognitive skills (socioemotional and creative development). Furthermore, the within family variation is of a similar magnitude for the teacher assessed age-five skills as it is for the age-11 externally assessed reading and mathematics test scores.

The early childhood socioemotional skills measure is the aggregate of three sub-scales: dispositions and attitudes, social development, and emotional development (see online Appendix D). The Cronbach's alpha coefficient for socioemotional skills is 0.90, suggesting that the three socioemotional sub-scales are internally consistent. For early childhood language (four sub-scales) and numeracy skills (three sub-scales), Cronbach's alpha is 0.93 and 0.91, respectively. Each sub-scale is scored out of 9, and together with the three remaining components (creative development, physical development, and knowledge and understanding of the world), each scored out of 9, the maximum attainable score is 117. The total target score for all children is 78 points, with at least 6 points in each of the seven socioemotional and language scales. Figure 1 plots the histograms for language, numeracy, socioemotional, creative and physical development, as well as the aggregate score. There is no evidence of score bunching at the threshold scores of 24 for language skills or 18 for socioemotional skills, nor at 78 for the overall Early Years Foundation Stage score. There is evidence of a modest ceiling effect for some of the domains, but not for the aggregate score.

## *2.2 Relationship between early childhood skills and later outcome*

We assess how skills measured by teachers at age five, as well as test scores measured at age 11, are related to later outcomes. This analysis helps anchor our main results below which addresses how the impact of birth order evolves through different stages of a child's development. We focus on two outcomes: high stakes age-16 GCSE exams and school

suspensions. School suspension and misbehavior partly reflect underlying noncognitive skills, including self-regulation, socioemotional adjustment and engagement, with consequences for later education and labor market outcomes (Bacher-Hicks et al., 2024; Jackson, 2018; Papageorge et al., 2019). If age-five skills forecast adolescent academic and behavioral outcomes then this indicates that these early childhood assessments capture meaningful underlying skills and behaviors and not simply transitory or noisy features of early development.

Table 2 presents evidence using the 2007 and 2008 Reception cohorts from the estimation sample. Outcomes for these two cohorts can be traced through to 2018 and 2019, when students are 16 years old. Columns 1 to 3 present results for the nationally standardized GCSE age-16 exam outcome, and columns 4 to 7 focus on school suspensions. All regressions include indicators for free lunch status, female, month and year of birth, and ethnic group. We also include fixed effects for the school attended at age five, which helps in accounting for potential reference bias in teacher assessments (Elder and Zhou, 2021; Feng et al. 2022).

Column 1 in Table 2 demonstrates that there is a strong positive association between age-five language and numeracy skills and age-16 exam outcomes, while the association between early socioemotional skills and age-16 exams is also positive and significant, but somewhat smaller. Conditional on these three age-five skills, creative and physical development at age five are negatively associated with age-16 exams. Column 2 assesses the association between age-11 test scores in reading and mathematics and age-16 exams, and this shows that the relative impact of the early cognitive skills in column 1 is comparable to that of age-11 test scores.

Columns 4 and 5 in Table 2 indicate that age-five socioemotional skills are strongly negatively associated with the likelihood of a child ever being suspended in grades 1 to 6 (primary school) and grades 7 to 11 (secondary school). This strong negative relationship between school suspensions and socioemotional skills continues to hold even after the inclusion of age-11 test scores (column 7). Age-five creative and physical development, on the other hand, are positively associated with school suspensions.

As seen in column 3, when both age-five skills and age-11 test scores are included in the regression, the direct association between early skills and age-16 exam outcomes is markedly attenuated, consistent with early skills exerting much of their influence via age-11 skills (Heckman et al., 2006). When we undertake this analysis separately for the white British

and minority samples, the relationship between early childhood skills and later outcomes is similar for the two sub-samples.

Finally, online Appendix Table 1 investigates whether within-family differences in age-five skills predict age-11 outcomes.[17] This analysis, which uses the full estimation sample, helps address concerns that teachers may compare siblings, even when they are enrolled in different grades, leading to, for example, reference bias in their skill assessments. For example, teachers may assign similar scores to siblings even when their underlying skills differ substantially. Appendix Table 1 shows that family fixed effect estimates of the association between early childhood cognitive skills and age-11 outcomes, although somewhat smaller in magnitude compared with OLS estimates, are nevertheless powerfully predictive of later outcomes. Meanwhile, the association in the family fixed effect specification between age-five socioemotional skills and age-11 reading and mathematics outcomes is relatively small, but is nevertheless statistically significant.

Overall, these results demonstrate that the Early Years Foundation Stage age-five skills are strongly predictive of later outcomes and, in addition, they are equally or as predictive of secondary school high stakes exam performance and behavioral outcomes as end-of-primary school age-11 test scores. This supports the view that the age-five skill measures contain substantial signal about meaningful underlying or latent developmental constructs.

# 3 Empirical Strategy

The empirical approach adopted in this study leverages information on siblings to compare outcomes for the older versus younger sibling. The core model exploits data for pairs of siblings, while results for three and more siblings are also produced in the robustness analysis. The family fixed effects model ensures that constant family-level characteristics common to siblings are differenced out (Black et al., 2005; Lehmann et al., 2018; Breining et al., 2019). The basic estimation model is as follows:

[17] Our data terminate in 2019, the last year before the COVID-19 pandemic disrupted schooling and assessments. We are therefore unable to trace students enrolled in the Reception grade between 2009 and 2012 to their age-16 GCSE outcomes.

$$y_{if} = \beta_1 SecondChild_{if} + \gamma X_{if} + \mu_f + \epsilon_{if}, \quad (1)$$

where $y_{if}$ is the outcome for student $i$ in family $f$ (national standardized age-five cognitive or noncognitive skills; standardized age-11 test scores, etc.). $SecondChild_{if}$ is switched on for the younger of the two children in the family. $X_{if}$ contains a dummy indicating sex of student $i$, as well as month of birth dummies. $\mu_f$ are family fixed effects. Standard errors are clustered at the family level. The parameter of interest is $\beta_1$, the advantage or penalty for the second child compared to the first. In a set of sensitivity exercises, we probe the robustness of our baseline results by incorporating time-varying control variables in the regression model, including free lunch status and the local poverty index, as well as using alternative transformations of the outcome variables. For comparison, we also present OLS estimates of the secondborn effect. This model includes indicators for free lunch status, sex, month of birth, ethnic background and a local area poverty index.

In order to shed light on heterogeneity by birth spacing, models of the following form are also estimated:

$$y_{if} = \beta_1 SecondChild_{if} + \sum_{k=2}^{5} \beta_2^k SecondChild_{if} * Spacing_f^k + \gamma X_{if} + \mu_f + \epsilon_{if} \quad (2)$$

where $Spacing_f^k$ is a dummy variable indicating that the two siblings are $k$ grades apart from each other. As described in the previous section, in our estimation sample siblings are up to five grades apart. In a robustness exercise, we also report results using spacing measured between siblings in years of age rather than grades. The purpose of equation (2) is to unpack the impact of birth order, conditional on grade (or birth) spacing, rather than an attempt to ascribe any causal relationship between spacing, birth order and the outcome of interest. In all the analysis below, the left hand side variables are either nationally standardized test scores or, in the case of outcomes such as school suspensions, these are de-trended. [18] Note

[18] Note that although cohort effects can be identified in a family fixed effect model when there is variation in birth spacing between siblings (Black et al., 2011), in the current setting cohort effects are perfectly correlated with a spacing interaction term. To see this most clearly, consider the period of early childhood data in this study, 2007 – 2012, and grade spacing between siblings. (In the empirical analysis below it is shown that birth spacing and grade spacing interactions produce very similar results.) In a model with (academic year) cohort effects, $CY_f^i$, which represents a cohort dummy for year $Y$ for sibling $i$, $i \in \{1,2\}$, in family $f$, and an interaction term, $SecondChild_{if} * Spacing_f$, where $Spacing_f$ is the gap in grades between the two siblings.

that the inclusion of month of birth dummies is especially important when examining variation in birth order effects by birth spacing. For closely spaced siblings, for example siblings with birth spacing such that they are one grade apart, the data reveals that younger sibling are more likely to be born in the summer, hence the inclusion of month of birth dummies ensures that seasonal birth patterns are not conflated with the impact of birth order. As noted earlier, grade retention is rare in the English education system, hence students typically progress from one grade to the next each academic year.

Finally, we also examine heterogenous effects to shed light on potential mechanisms underlying our findings. Specifically, we estimate effects by free lunch status, a proxy for socioeconomic background, by ethnic or racial background, and by sibling gender composition.[19]

# 4 The Impact of Birth Order on Age-Five Skills

This section presents results for children's age-five nationally standardized language, numeracy, socioemotional, creative, and physical development outcomes. Results are first reported for the full sample. However, given the large differences in birth order effects for the white British versus minority groups, all subsequent results are presented separately for these two groups.

Table 3 presents estimates using the family fixed effects model (equation (1) above) for each of the five outcomes. Simple OLS estimates are produced in online Appendix Table 2. In most cases, the fixed effects estimates are qualitatively similar to their OLS counterparts. In the results reported below, the focus is on estimates from the preferred family fixed effects model.

Panel A of Table 3 reports results for the full sample. For the full sample, the estimates suggest that the secondborn child performs worse than the firstborn on language skills but enjoys an advantage on the four remaining skill measures. Separate results for the white

In this case the following first differences equation holds: $C08_f^2 - C08_f^1 + 2\{C09_f^2 - C09_f^1\} + 3\{C10_f^2 - C10_f^1\} + 4\{C11_f^2 - C11_f^1\} + 5 * C12_f^2 = SecondChild_{if} * Spacing_f$.

[19] Online Appendix A lays out a simple theoretical framework which motivates our empirical analysis.

British majority population and the minority population are presented in Panels B and C of Table 3, respectively.

Turning to the sample of white British students first, the estimates show that on average the secondborn child performs 0.05 standard deviation worse than the firstborn on language skills. The birth order effect for numeracy skills is substantially smaller, with the secondborn child performing just 0.01 standard deviation worse than the firstborn. The impact on socioemotional skills is not statistically significant. This is a precisely estimated null effect: the 95% confidence interval rules out effects larger than 0.004 standard deviation in absolute terms. There is, however, a small negative secondborn effect for creative development. Meanwhile, the secondborn effect for physical development is positive.[20]

These overall effects mask substantial variation in birth order effects by birth or grade spacing between siblings, as reported in Table 4. Estimates of the impact by grade spacing are undertaken using equation (2), which includes grade spacing interaction terms.[21] Panel A of Table 4 reports results for the white British population. Surprisingly, at grade spacing one, the secondborn effect is *positive* across all five age-five outcomes. At grade spacing two, the secondborn effect is either positive or zero for four of the five outcomes. At spacing of three grades and higher, on the other hand, there are large negative secondborn effects for cognitive skills.[22] Effects for noncognitive skills (socioemotional and creative development), on the other hand, remain small. Two key implications follow from the analysis by grade (and birth) spacing. First, relying solely on the basic model (equation 1) can give a misleading picture of the underlying birth order effects for early childhood outcomes, obscuring substantial heterogeneity and even masking effects of opposite sign. Second, relative to

[20] The finding that secondborns outperform firstborns on age-five physical development is consistent with the established literature which shows that secondborns have better prenatal and birth endowments (Breining et al., 2020; Brenøe and Molitor, 2018; Pruckner et al., 2021), suggesting that those early health advantages persist into early childhood, despite the fact that secondborn children tend to receive less preventive care or breastfeeding relative to firstborns (Black et al., 2016; Brenøe and Molitor, 2018; Lehmann et al., 2018; Pruckner et al., 2021).

[21] Online Appendix Table 3 report results using birth spacing interaction terms (i.e. focusing on spacing measured in years between siblings' birth dates, rather than grade spacing). These birth spacing results are very similar to those reported in Table 4.

[22] Note that the purpose of the birth spacing analysis is to present heterogeneity in birth order effects for age-five skills. These estimates do not necessarily reveal the causal effects of the interaction between birth order and birth spacing. In theory, shorter birth spacing may harm the secondborn child through a physiological mechanism if the mother does not have sufficient time to recover from the previous pregnancy. Conversely, shorter birth spacing may disadvantage firstborns be reducing exclusive parental time. See Buckles and Munnich (2012) for further discussion.

cognitive skills, the effects on noncognitive skills remain small. For example, at higher grade and birth spacings, while the secondborn disadvantage for cognitive skills ranges between 0.05 and 0.16 standard deviation, the secondborn disadvantage for socioemotional skills ranges between 0.01 and 0.03 standard deviation. We can benchmark these effects relative to skill gaps between children from advantaged and disadvantaged families, as proxied by eligibility for free lunch. For age-five cognitive (language and numeracy) and socioemotional skills, this gap is 0.6 standard deviation and 0.5 standard deviation, respectively.

In summary, for age-five white British students, the baseline results point to relatively modest secondborn deficits for language skills and no impact on socioemotional skills. For closely spaced siblings, secondborn children enjoy an advantage. Only at birth spacing of three years or higher do we see substantial negative secondborn effects for cognitive skills; for noncognitive skills, these deficits continue to be very small.

The null or very small effects for noncognitive skills, including socioemotional development, are surprising for several reasons. First, as we demonstrate below, secondborns perform substantially worse than firstborns on adolescent behavioral outcomes. A natural assumption would be that such a disadvantage would show up at age five, particularly for socioemotional skills. Second, the prior evidence relating birth order and adult noncognitive skills and behavioral outcomes points to disadvantages for later-born siblings. For example, Black et al. (2018) exploit Swedish registry data and find substantial later-born disadvantage on adult noncognitive skills, including emotional stability and being less socially outgoing. They report a secondborn deficit, relative to the firstborn sibling, of approximately 0.1 standard deviation on a noncognitive ability index.[23] Similarly, Houmark (2023) exploits Danish administrative data to demonstrate that among adolescent children (from around age 10 and 11), later-born siblings score lower on personality measures such as conscientiousness and emotional stability. It seems unlikely that the small effects on age-five noncognitive skills observed in our setting are driven solely by greater measurement error in the former. First, because these skill measures are outcomes, classical measurement error would not mechanically attenuate the estimated treatment effects, though it could reduce precision. However, reduced precision is not a major concern in our setting given the very large sample

[23] See Table 2 in Black et al. (2018 ). Their sample consists of for adult Swedish males with a valid military draft record , approximately age 18. Below we report results separately for age-five boys and girls and our conclusions remain broadly unchanged.

sizes. Second, the strong predictive relationship between age-five noncognitive skills and adolescent behavioral and academic outcomes, even after conditioning on cognitive skills, suggests that these measures capture meaningful underlying variation in the latent noncognitive skill construct. Third, as we show in the robustness analysis below, our baseline set of results survive a series of sensitivity checks, including various transformations of the outcome measures.

Finally, turning to the results for minority students, the baseline estimates in Panel C of Table 3 demonstrate that there is a large advantage enjoyed by secondborn children on all five skill measures. These are around 0.1 standard deviation for noncognitive skills and physical development and even larger for language and numeracy skills. These effects remain positive at all birth spacings (Panel B, Table 4). Possible explanations for the large secondborn advantage for the minority population include cultural norms as well as differential sibling spillovers for minority families compared to those in white British families. These are discussed in greater detail below.

## *4.1 Robustness Checks*

The robustness of the baseline results for age-five outcomes are probed through a series of sensitivity analyses, including alternative transformations of the outcome variables. First, instead of standardizing scores, scores are expressed as percentile ranks within each cohort. The results are presented in Table 5, columns 1 to 5. All the results for the white British and minority populations are in line with the baseline results reported in Table 3, except for the white British socioemotional and physical development skills outcomes, for which there is now a small negative secondborn effect.

Second, the total age-five skills score is employed as the main outcome of interest. This is the sum of the scores for language, numeracy, socioemotional, creative development, physical development, as well as knowledge and understanding of the world. The results are reported in column 6 of Table 5. This shows that for the white British population, the secondborn effect is - 0.02 standard deviation, whilst for the minority population the impact is 0.16 standard deviation.

Third, two factor-indices are constructed to provide a parsimonious summary of each domain: a noncognitive factor (the first factor from the socioemotional and creative development skill measures) and a cognitive factor (the first factor from language and mathematics skills as well as knowledge and understanding of the world). Each factor is then standardized to mean 0 and unit variance and used as the outcome variable. The results in columns 7 and 8 of Table 5 show that for the white British population the secondborn impact is - 0.01 standard deviation and - 0.05 standard deviation for the noncognitive and cognitive factors, respectively; the results for the minority population are 0.12 standard deviation and 0.18 standard deviation, respectively.

Fourth, in order to assess the importance of differential socioeconomic circumstances experienced by older and younger siblings, in online Appendix Table 4, time-varying control variables are included as controls in the main (family fixed effects) specification. Specifically, the free lunch status and the home location local poverty index are included, both measured at the time of Reception (age four or five) enrollment. The results presented in Appendix Table 4 demonstrate that there is very little change in the estimates compared to the core results presented in Table 3.

Next, the model is re-estimated using a sample restricted to families with three siblings with early childhood outcomes in the 2007 – 2012 period. The main analysis focuses on first- and secondborn siblings, given the relatively narrow time span of availability of the early childhood outcomes on a consistent basis between 2007 and 2012. Given the limited window for the early childhood measures, restricting the sample to families with three siblings yields a smaller sample than in the main analysis. The estimates are reported in online Appendix Table 5. These show that the secondborn effects are broadly in line with the baseline results. The only exceptions are numeracy and creative development skills for white British students, both of which yield small negative effects in the baseline model, but are now statistically insignificant. Overall, these results suggest that the core results in Table 3 are robust to the inclusion of higher born siblings and are not driven by, for example, a 'last born' effect in, say, families of parity two (Ejrnæs and Pörtner, 2016). Finally, it should be noted that, in general, there is a modest rise (in absolute terms) in impact for third-born children, relative to secondborn siblings, for both white British as well as minority families.

In another robustness exercise, normalization of skill measures is undertaken separately for the two main groups: scores are now standardized against that cohort's

national distribution, separately for white British and minority students. This accounts for, for example, the possibility that there is a secular improvement in skills for the minority group relative to the majority one. Results in Appendix Table 6, Panel A demonstrate that for the white British group there is very little change relative to the baseline results, whilst for the minority group the impact of being secondborn remains large and positive, although the estimates are around a third smaller relative to the earlier results.

## *4.2 Heterogeneous Effects*

We provide further heterogeneity analysis by family socioeconomic status (free lunch status and neighborhood deprivation), gender composition of the siblings, as well as detailed minority status in the next subsection. This analysis can potentially shed light on the mechanisms underlying our core results.

Turning first to variation in birth order impact by socioeconomic status, existing evidence provides a mixed set of results. For example, Black et al. (2005) find little variation in education attainment when they stratify their sample by mother's education. Breining et al. (2020), on the other hand, find that heterogeneity in birth order effects for delinquency by socioeconomic status does not follow a uniform pattern, with stronger secondborn disadvantage appearing in lower SES families in some cases and in higher SES families in others, depending on age and geographic location. There is scant evidence on such heterogeneity for children's skills at age five or earlier. If differential parental inputs (resource dilution) is the key mechanism underlying birth order effects, then prior evidence suggests that by relaxing constraints, including capacity to parent, a better family environment (for the firstborn child) may be particularly effective in disadvantaged households (Doyle, 2020). This would point to a relatively higher secondborn penalty in such families. On the other hand, the evidence also suggests that in advantaged families, maternal time inputs may be especially productive (Ginja et al., 2019), in which case a higher secondborn penalty may be observed in more advantaged families.

In order to assess such heterogenous effects, we use two proxies for family background: free lunch status and neighborhood deprivation. Table 6 reports results for the sample stratified by free lunch status. The free lunch dummy for a family is switched on if

either sibling is eligible for free lunch in any given grade. The results in Panel A of Table 6 suggest that for white British children the secondborn disadvantage in age-five cognitive skills is larger in more advantaged families, while the impact on socioemotional skills remains insignificant for both groups. Secondborn siblings perform 0.06 standard deviation lower than firstborns on age-five language skills in families where the siblings are not eligible for free lunch, whilst the secondborn disadvantage is 0.02 standard deviation in eligible families. These results suggest that resource dilution may produce especially pronounced birth order penalties for cognitive skills in more advantaged households.

Heterogenous results by neighborhood deprivation are shown in online Appendix Table 8. Neighborhood deprivation is measured using the Income Deprivation Affecting Children Index, IDACI, an indicator based on the share of children living in disadvantaged families in a local area. The small statistical neighbourhoods used for this metric (Lower-layer Super Output Area, LSOA) average about 650 households. Students are linked to this neighborhood deprivation index via their home postcode. The results from this exercise corroborates the analysis using the free lunch measure: once again, the strongest impact is for language skills, with the largest secondborn penalty faced by students from more advantaged families. For the bottom and the middle tercile of the sample stratified by neighborhood deprivation, secondborn siblings experience a language skill disadvantage of 0.07 and 0.06 standard deviation, respectively, whilst the penalty for secondborns from families residing in the most disadvantaged neighborhood is a more modest 0.02 standard deviation.

Results for minority students by free lunch status in Panel B of Table 6, as well as those in online Appendix Table 8, show that the positive impact for secondborn children is significantly larger for disadvantaged families. This pattern of results is discussed below in the context of the detailed minority-level analysis below.

There may also be important heterogeneity in the birth order effect by gender composition of the siblings. Such an analysis is important both for learning about heterogenous effects but also because it helps establish the robustness of our results if, for example, parents have fertility stopping rules based on sex preferences of children. To investigate this, we report results separately for samples of boy-boy, girl-girl and mixed boy-girl pairs of siblings. Online Appendix Table 9 presents the results. We do not find evidence of any strong variation or consistent patterns in birth order impact on age-five skills by the

gender composition of the sibling pairs for white British students. For example, the secondborn disadvantage for language skills is somewhat larger for girl-girl pairs than for boy-boy pairs, but this pattern is reversed for creative development. Similarly for minority students, across all skill types, the results are similar for all pairs of gender compositions in Appendix Table 9.

### *4.3 Detailed Minority Age-Five Results*

There are several plausible mechanisms for the opposite-signed birth order estimates observed among minority students, compared with white British students. One possible explanation hinges on different cultural norms between the white majority and the minority populations. These include the interaction of birth order with eldest son preference (Jayachandran and Pande, 2017) and varying parenting and educational practices among different immigrant groups (Isungset et al., 2020). Prior evidence on birth order effects for skills and education outcomes for minority versus majority groups is limited to adolescent students and adults. For example, Isungset et al. (2020) focus on lower secondary school test scores. They find that amongst Norwegian-born children of immigrant parents, birth order effects vary by country of origin.[24]

The administrative data allows us to undertake fine-grained analysis across several diverse minority groups to explore the cultural norms hypothesis. We report results for the minorities sample stratified by the six largest ethnic or heritage groups, as follows: African, Afro-Caribbean, Bangladeshi, Indian, Pakistani, and 'white other', which includes Eastern European background children. These groups represent successive large waves of migration to the UK in the post-World War II era. For example, the Afro-Caribbean, Indian and Pakistani groups are among the earliest wave of migrants, whilst those from Eastern Europe are among the most recent. Table 7 presents the results. Strikingly, the table reveals that the large, positive advantage for secondborn siblings across age-five skills is present for *all* these minority groups. The fact that a broadly similar birth order pattern emerges across a diverse set of minority groups suggests that explanations based on cultural practices are unlikely to

[24] For example, Breining et al. (2020) find relatively larger secondborn delinquency disadvantages in immigrant families in Denmark, except for teen incarceration.

be driving these results. A common pattern of results suggests that birth order effects are more likely to reflect mechanisms shared across immigrant families.

Two alternative explanations for the large advantage enjoyed by secondborn minority students relate to improvements in economic conditions for minority families over time, as well greater experience with the education system leading to improved school quality choices for later-born children. The results from the earlier exercise using time-varying measures of economic conditions (free lunch status of each individual child as well as local neighborhood poverty indicator at the time of enrollment in the Reception grade) goes against the former explanation. We discuss the role of differential school quality for siblings in the next subsection.

A further potential explanation for the large secondborn advantage for siblings from minority families is that the younger sibling benefits from positive spillovers in English language skills from their older sibling. This may be especially important for minority families which speak limited or no English at home. Table 8, Panel A reports results for minority students separately for those students who speak English versus those who speak another language at home. The results demonstrate that there are very large differences in birth order effects by home language status. For families where English is the language spoken at home, the benefit for secondborn children varies between 0.04 standard deviation and 0.08 standard deviation for the five age-five skill measures. For families where English is not the language spoken at home, the secondborn impact is dramatically larger: between 0.14 standard deviation and 0.25 standard deviation. Positive effects for secondborn siblings are not confined to language skills alone, but instead are broad-based across all skill domains. This likely reflects the fact that language skills are a critical input for the acquisition of all other skills, especially at early ages. The large difference in impact by home language status for minority students points to potentially important English language spillover effects, which may drive the secondborn advantage among minority families.[25] This would also help explain the pattern of results by free lunch status for minority students reported in Table 6 above: free lunch status and speaking a language other than English at home are correlated, hence

[25] It is noteworthy that the secondborn advantage, although weaker than for other minority groups, is nevertheless present among Afro-Caribbean children. This may reflect non-standard English spoken in some of these families. See, for example, Sebba (1993) and Sebba (2007) for evidence of distinctive language practices among Afro-Caribbean families.

the larger secondborn advantage in families where the siblings are eligible for free lunch. We return to the interpretation of birth order effects for minority students in section 7, where we examine how these effects evolve from age five to adolescence.

## *4.4 Mechanisms: Deferred Enrollment, School Quality, Peers, and Parental Investment*

We next explore potential mechanisms underlying the observed birth order effects by examining within-year deferred enrollment, school quality, peer quality and parental investment. Systematic differences by birth order in these inputs may help explain why firstborn children differ from their secondborn siblings in age-five cognitive and noncognitive outcomes.[26]

In England, over the period of analysis there is within-year variation in Reception grade enrollment of children to the term of entry (fall, spring or summer term). This variation in within-year deferred enrollment is induced by local authority policies in the early years of the analysis, as well as parental choice in later years (Cornelissen and Dustmann, 2019; Campbell, 2023). With respect to parental choice, parents are allowed to defer their child's enrollment to later in the academic year, though not beyond the point at which the child reaches compulsory school age, which is reached at the end of the term in which the child turns five. The mean delayed enrollment rates are 6% and 4% for firstborn and secondborn white British children, respectively; and 9% and 4% for firstborn and secondborn minority children, respectively.[27] We ask whether these systematic differences between siblings can account for birth order effects. Table 9, Panel A reports results from this analysis for the white British sample and online Appendix Table 10 also reports it for the minorities sample. The results show that deferred enrollment is associated with large negative deficits in age-five cognitive,

[26] Andrew et al. (2024) and Ganimian et al. (2024) provide evidence on the importance of preschool quality for early childhood outcomes. Gray-Lobe et al. (2022) shed light on the long-term benefits of pre-school attendance.

[27] We define delayed enrollment as not being enrolled in the current school in September and that there is no change in the student's home postcode in the last 12 months. Note that deferring school enrollment for a whole academic year ('redshirting') and grade retention (repeating a year) is rare in English state schools. Thomson (2024) reports that over the period of our analysis, less than 0.3% of students are enrolled in a grade which does not correspond to their age appropriate grade. By contrast, Bassok and Reardon (2013) document redshirting rates of around 4 to 5.5% in the United States.

noncognitive and physical development skills, for both white British as well as minority children.[28] However, controlling for deferred enrollment does not dramatically affect birth order gaps for white British children, but it does reduce the estimates of the advantage enjoyed by secondborns in minority families by between 10% and 20%. This suggests that at least part of the apparent secondborn advantage in minority families may reflect their lower rates of delayed Reception grade enrollment relative to firstborn siblings.

We next consider whether differences in school quality can help explain the observed birth order effects. For the school quality measure, we use schools' inspection ratings. Prior evidence shows that inspection ratings in England are related to conventional measures of school effectiveness, including test score value added, and that parents respond to these ratings through school choice and investment decisions (Greaves et al., 2021; Greaves and Hussain, 2026). Even if younger and older siblings are enrolled in the same school, there is variation over time in schools' ratings and we exploit such variation across siblings to address the role that school quality may play in mediating birth order effects. Navigating the school choice system is complex and parents' search costs for better schools may decrease over time as their experience with the education system improves. Such a mechanism may be especially important for minority families. In fact the data point to marginally higher quality schools in the Reception grade both for white British and minority secondborn siblings (an advantage of 0.03 and 0.05, respectively, on the 4-point school quality inspection scale). The results in Table 9, Panel B and online Appendix Table 11 show that there is indeed a positive association between school inspection ratings and age-five outcomes. However, the estimated secondborn effect is virtually unchanged after controlling for school inspection ratings. This suggests that school quality, at least as captured by inspection ratings, does not account of the estimated birth order differences in early childhood outcomes.[29]

A further potential mechanism is peer quality. This is distinct from the school quality channel: even if siblings attend schools with similar inspection ratings, they may be exposed to different peer environments, such as socioeconomic composition of their peers. Systematic

[28] These correlational results are in line with those reported in Cornelissen and Dustmann (2019), who exploit variation across English local authorities in Reception grade admission rules to identify the causal effect of school exposure on early childhood outcomes.

[29] Breining et al. (2020) also examine whether differences in school quality mediate birth order effects for delinquency outcomes. Using school accountability grades in Florida, they find little evidence that secondborn boys attend lower quality schools relative to their firstborn siblings, leading them to rule out school quality as a major mechanism in driving their results.

differences in peer exposure could therefore contribute to birth order gaps over and above any differences in measured school quality. For this exercise we focus on gender and socioeconomic (proxied by the proportion of students eligible for free lunch) measures of peer quality (Hoxby, 2000; Sacerdote, 2011). Mean peer quality on each of these measures is calculated for the school cohort, leaving out the student's own gender or free lunch status. We include both measures of peer quality in the birth order estimation equations. Results are reported in Table 9, Panel C and online Appendix Table 12. These show that the proportion of female peers is positively associated with age-five outcomes and the proportion of free lunch eligible students is negatively associated with age-five outcomes. Nevertheless, there is very little change in the estimated secondborn effect relative to the baseline estimates after controlling for these measures of peer quality. These results indicate that birth order gaps in early childhood outcomes operate through channels other than the dimensions of peer quality considered here.

In order to investigate the role of birth order on parental investments further, we use the Millennium Cohort Study, a UK cohort panel study which follows children born in the UK between 2000 and 2002 (UCL Centre for Longitudinal Studies, 2026). These cohorts correspond closely to the cohorts employed in our main analysis. The MCS has detailed information on parental inputs into activities such as reading to the child and help with early childhood skills such as writing and spelling, but lacks information on siblings. Therefore, for this analysis we undertake a cross-family comparison of associations between birth order and parental investments, after controlling for a detailed set of controls, including mothers' fertility and detailed parental and child covariates. Details of the data are provided in Appendix B. The results are presented in online Appendix Table 13. These show that for the white British sample, there is a negative relationship between birth order and the frequency of reading to the child, both at age three and at age five, as well as for help with the alphabet and reading at age three. This is consistent with the existing literature suggesting that parental investments decline with birth order (Black et al., 2018). For minority children, however, there is little evidence of a strong birth order gradient. The coefficient estimates are generally statistically insignificant, with evidence of a negative effect for reading at age 5 at birth order 3, but positive effects at birth order 4 or higher for help with language and numeracy. These results suggest that parental investment behavior by minority parents is

unlikely to explain the strong advantage enjoyed by minority secondborn siblings relative to firstborns on age-five cognitive and noncognitive skills.

# 5 The Evolution of Adolescent Outcomes

This section assesses how the effects of birth order evolves over time for white British students. The next section reports results for minority students.

## *5.1 Test Scores and Exams*

Table 10 reports results for age-11 ("Key Stage 2") nationally standardized reading and mathematics test scores. The first two columns of Panel A show that for white British students, the impact of being secondborn confers a substantial negative penalty, with estimates of - 0.13 standard deviation and - 0.06 standard deviation for standardized reading and mathematics outcomes, respectively. Results for high stakes age-16 exams ("GCSE's") are reported in Panel B.[30] These show that for white British students, there is a large penalty of - 0.15 standard deviation on the nationally standardized total GCSE exams score for secondborn siblings.[31] These are substantial effects, as compared to, for example, the gap in test scores between children by eligibility for free lunch. For age-11 test scores and age-16 exams, this gap is 0.6 standard deviation and 0.7 standard deviation, respectively. Online Appendix Table 14 uses percentile ranks within each cohort, instead of standardized scores. These results are consistent with the core results in Table 10, showing relatively large penalties for secondborn siblings.

The results in Table 10 imply substantial growth in the firstborn advantage in cognitive skills from age five to adolescence.[32] A potential concern with this interpretation is that changes in birth order effects on standardized test scores may simply reflect changes in the

[30] As explained above and in the data appendix, for age-16 outcomes, students are enrolled in grade 11 in the six years 2014 to 2019.

[31] Heterogeneous impact by free lunch status and by gender composition of the siblings are reported in online Appendix Tables 16 and 17, respectively. In both cases, there is little variation in the impact of birth order across the various sub-samples, either for the white British or the minority population. Online Appendix Table 18 also demonstrates that the results are robust to within-group standardization.

[32] Recall that at age five, the secondborn penalty is – 0.05 standard deviation and – 0.01 standard deviation for language and numeracy skills, respectively.

dispersion or scaling of the underlying tests, rather than changes in the underlying impact of birth order on skills. It should be noted however, that, as we showed in Table 2, age-five and age-11 skill measures are comparable in their predictive content: a unit standard deviation increase in either age-five or age-11 cognitive skills predicts a similar increase in the high stakes age-16 exam outcome. Furthermore, as we show below, we also uncover examples of a reversal in the sign of the birth order effect from age five to age 11. Scaling concerns are unlikely on their own to explain such sign reversals.

As in our earlier analysis, we examine mechanisms related to differential schooling inputs as well as the role of deferred enrollment in the Reception grade in driving adolescent birth order effects. In practice, neither deferred enrollment in the Reception grade, nor peer or school quality at age 11 have substantial impact on estimated birth order effects on age-11 education outcomes (see online Appendix Tables 10, 11 and 12). Time-varying controls (free lunch status and local poverty index) also cannot account for the observed birth order differences (online Appendix Table 15).

Although we do not observe changes in family structure such as parental separation or divorce, prior evidence suggests that this mechanism is unlikely to account for birth order effects (see for example, Black et al. 2005; Hotz and Pantano, 2015; Lehman et al., 2018). Furthermore, the national patterns of divorce also lend weight to the argument that changes in whether families remain intact or not are unlikely to account for the small and null birth order effects at early ages followed by substantial impact at age 11. In England and Wales, approximately 31,900 couples whose divorce was finalized in 2006 had an eldest child aged up to 10 when the divorce petition was filed (Office for National Statistics, 2009). For 29% of these couples the eldest child was aged between 0 and 4. This would suggest that if family structure is a key driver of the birth order patterns we observe, we should already observe substantial secondborn penalties by age five. The fact that we do not suggests that this is unlikely to be the driving mechanism.

We also investigate whether our results are sensitive to a proxy for family stability, namely whether or not the older and younger sibling are enrolled in the same school at the time of their Reception grade. Our results show that there is a larger secondborn penalty at age five for cognitive and noncognitive skills in families where there is such mobility between schools. However, the rise in the penalty between age five and age 11 is larger for siblings in families where there is no such mobility. We interpret this finding as suggestive evidence that

family instability is unlikely to account for the rise in the secondborn penalty between age five and 11.

Results on how the influence of birth order on age-11 test scores varies by birth spacing between siblings for white British students is reported in Panel A of Table 4, columns 6 and 7. Recall that for the white British population, at birth spacing of one year, secondborns outperform firstborn siblings on age-five skills. For this group of students, the secondborn advantage at age five reverses sign by age 11, with secondborn siblings experiencing a substantial penalty of 0.1 standard deviation for reading and a more modest penalty of 0.03 standard deviation for mathematics. Similarly, at birth spacing of two years, age-five secondborn effects on skills are either modestly negative or zero; in contrast, age-11 birth order effects at birth spacing of two years are now large and negative. Online Appendix Table 19 shows the impact by birth spacing for age-16 exam scores. For the white British population, the secondborn impact at birth spacing of one year and two years continues to be large and negative.

Taken together, these set of results on the evolution of birth order effects across child development stages for the white British population are revealing. At age five, although the firstborn advantage is detectable along some domains (particularly for language skills), it grows substantially through adolescence. More striking still are the examples where there are outright reversals between age five and adolescence. For closely spaced siblings, the apparent early deficits for firstborn siblings are ultimately reversed, with large shifts in favor of firstborns by adolescence. These findings have important implications for theoretical explanations of birth order effects. The simplest version of the resource dilution model, which emphasizes parental constraints in the earliest stages of child development, would predict a firstborn advantage from early childhood onwards. Similarly, theories relying on sibling and family spillovers must account for why the firstborn premium intensifies markedly during adolescence rather than remaining stable, as well as providing an explanation for the sign change in the estimated birth order effect between early childhood and adolescence.

In addition to these theoretical considerations, it is also noteworthy that while there is substantial heterogeneity in the impact of birth order by birth spacing for early childhood outcomes, heterogeneity by birth spacing is much less pronounced for adolescent outcomes, such as age-16 test scores. This provides suggestive evidence that during early childhood,

birth order effects are more sensitive to the immediate family environment such as the timing of sibling births, but as children age, the influence of birth pacing weakens and more general birth order mechanisms likely assert themselves within families.

## *5.2 School Suspensions*

School suspensions likely capture important behavioral and noncognitive skills, such as self-regulation and social adjustment. The prior evidence demonstrates that school suspensions, as well as related outcomes such as truancy and misbehavior in school, have adverse long-run consequences, including in the labor market and the criminal justice system (Bacher-Hicks et al., 2024; Bertrand and Pan, 2013; Jackson, 2018).

We report the impact of being secondborn on ever having been suspended in Reception grade through grade 6 (i.e. the primary schooling phase); and grade 7 through grade 11 (i.e. the secondary schooling phase). In the early childhood and primary phase, being suspended is relatively rare and hence an extreme outcome. The mean suspension rate for white British students in the Reception grade is one tenth of one percent, while in the whole primary schooling phase it is 1.3%. On the other hand, 10% of students are suspended at least once during the secondary schooling phase. Rates for boys are higher than for girls and results by gender composition of the siblings are reported below.

Results for the white British population are shown in Panel A of Table 11. These show that for children in primary school, the secondborn child is one-tenth of a percentage point more likely to be suspended, representing a 10% increase relative to the mean suspension rate. At the secondary school phase (grades 7 through 11), secondborn children are 1.1 percentage points more likely to be suspended relative to firstborns, representing an 11% increase relative to the mean. Overall, these results demonstrate that the secondborn child is substantially more likely to be suspended from school, an effect which emerges in primary school and continues through secondary school. Our results are qualitatively similar to those reported in Breining et al. (2020) for students in Florida, although our estimates for England are somewhat smaller, which may reflect institutional differences in approaches to disciplinary infractions across the two settings. Using Danish registry data, Houmark (2023) demonstrates that advantages accrue to earlier-born siblings on different personality

measures for children from grade 4 onwards (approximately age 10 and over), which is consistent with the birth order results for adolescent delinquency outcomes.

Online Appendix Table 20 shows the impact by birth spacing. This demonstrates that the impact of being secondborn at lower birth spacing is always positive, both in the primary as well as secondary schooling phases. At higher birth spacing the impact appears to be negative. Online Appendix Table 21 reports results by the gender composition of the siblings. This shows that the impact is somewhat larger for boys than for girls, although relative to the mean suspension rate by gender, it is broadly similar across the different samples. For example, white British secondborn boys are 1.4 percentage points more likely to be suspended in the boy-boy sample, which represents a 10% rise relative to the overall suspension rate for boys. Secondborn girls are 0.8 percentage points more likely to be suspended in the girl-girl sample, representing a 12% rise relative to the overall rate for girls.

# 6 Minority Student Trajectories

For minority students, the results for outcomes at age five demonstrated a substantial advantage for secondborn children relative to firstborns, in contrast to the pattern observed for the majority white British population. We now show that this secondborn advantage either diminishes sharply or reverses entirely by adolescence, as measured by test scores and performance on high stakes exams. Turning first to age-11 test scores, the results in Table 9 (Panel A, columns 3 and 4) show that secondborn minority students underperform relative to firstborns on reading scores, whilst the advantage secondborns enjoy on mathematics is now relatively modest (0.03 standard deviation). By age-16, the results in Panel B of Table 9 show that secondborn minority students underperform by 0.05 standard deviation relative to their firstborn counterparts on the overall standardized GCSE exam score.[33]

This reversal in fortune for secondborn students relative to firstborns is even more stark in those minority families where English is not spoken at home. The secondborn impact on age-11 test scores for minority students stratified by whether English is spoken at home

[33] Estimating the model using age-11 and age-16 scores normalized with respect to each cohort's national distribution, separately for white British and minority students, yields very similar estimates to those reported in Table 9 (see online Appendix Table 18).

are reported in Panel B of Table 8 (columns 1 to 4). These demonstrate that the very large skill advantages enjoyed by secondborn children at age five from families where English is not spoken at home – above 0.2 standard deviation for cognitive skills – either shrink dramatically (mathematics) or become indistinguishable from zero (reading) by age 11. For minority children where English is spoken at home, the relatively small positive secondborn effects at age five also turn negative by age 11. For age-16 exam scores the secondborn impact is negative, both for those minority children who speak English at home, as well as those who do not.

Online Appendix Tables 16 and 17 report heterogeneous impact on adolescent education outcomes for minority students by free lunch status and by gender composition of the siblings, respectively. Broadly speaking, results for minority students across these different sub-groups are quite similar.[34]

The reversal of the secondborn advantage in minority families points to large relative gains for firstborn children between age five and age 11, especially for families where English is not spoken at home. Next, we rule out explanations based on specific cultural factors as well as mechanisms related to school quality, peers, delayed primary school enrollment, and evolving economic circumstances as possible mechanisms driving adolescent outcomes for minority students. Results by detailed minority status for adolescent academic outcomes are reported in online Appendix Table 22. Consistent with the results for the minority group overall, the disaggregated analysis across all minority groups also shows that by age-11 the benefits enjoyed by secondborn children at age-five are either reversed, becoming negative, or are substantially diminished. By age-16, the impact of being secondborn is now either negative or statistically indistinguishable from zero for all minority groups. Results for age-11 outcomes using deferred enrollment, peer characteristics, and school quality as controls are reported in the second panel in online Appendix Tables 10, 11 and 12, respectively. Results using time-varying controls (free lunch status and local poverty index) are reported in the final two columns of online Appendix Table 15.

[34] For the school suspensions outcome, the results in Table 10 show that for minority students, there is no discernible impact of birth order in the primary phase of schooling. In the secondary schooling phase, the secondborn sibling is 1.1 percentage points more likely to be suspended relative to firstborns, representing an 8% increase relative to the mean.

One potential explanation for the relatively rapid progress made by firstborn students from families who do not speak English at home may lie in theories in psychology and economics which point to critical periods in childhood for language acquisition (Bleakley and Chin, 2004; Johnson and Newport, 1989; Lenneberg, 1967). A key point emphasized in this literature is that language acquisition is age-sensitive. For example, Bleakley and Chin (2004) show that immigrant children from non-English speaking origin countries exposed to English in the United Staes at younger ages (up to ages 8 or 9) become fluent in English as adults, while those arriving later tend to have worse adult proficiency. Our results are consistent with this critical period hypothesis. For households where English is not spoken at home, firstborns may enter school with relatively weaker English language skills, whilst secondborns, by contrast, benefit from English language spillovers. This can generate a secondborn advantage at age five. However, schooling may allow firstborns to close these initial gaps relatively quickly, so that more standard firstborn advantages, deriving from greater early parental investment, become more salient. The fact that firstborns from minority families catch up and subsequently outperform secondborns by adolescence is therefore consistent with a process in which early language spillovers initially favor secondborns, while later educational outcomes increasingly reflect conventional firstborn advantages.

# 7 Conclusion

The prior literature has established the importance of birth order effects for adolescent and adult outcomes using large administrative datasets (Black et al., 2005; Breining et al., 2020). This study exploits unique population-level data with unusually rich measures of child development at age five to address whether birth order gradients are present in early childhood. We arrive at the surprising conclusion that among white British students, although firstborns exhibit advantages in early cognitive development, there is no impact on age-five socioemotional skills. In particular, for closely spaced siblings, firstborns experience a *disadvantage* on both cognitive and noncognitive skills, relative to their secondborn siblings. Significantly, this penalty is reversed by middle childhood. For white British siblings overall, the firstborn advantage grows through adolescence, and is large and positive for age-11 test scores, high stakes exams at age-16, and adolescent behavioral outcomes.

Our findings of a growth in the birth order effect, as well as the shift from a secondborn advantage to a firstborn advantage in some cases, are striking. The null or very small effects we observe at age five are unlikely to be explained by simple measurement error or test score scaling in the different stages of schooling. One reason is that the early childhood skill measures we use strongly predict outcomes through adolescence. Another is that the impact of being firstborn not only grows over time, but also flips sign in some cases. One way to interpret birth order effects is through the lens of the dynamic human capital model, a natural approach if the resource dilution hypothesis is the underlying mechanism (Cunha et al., 2010). If the very early stages of child development (the 'first 1,000' days, Almond et al., 2018, Attanasio, 2026) are a critical period, and the degree of substitutability for some early skills declines in later stages of child development, then under the resource dilution hypothesis we would expect a firstborn advantage to have emerged by age five.

Instead, we uncover very modest or even no advantages for firstborns at age five. There are several potential explanations for this result. One possibility is that that there may be a 'sleeper' effect, mirroring a pattern documented in the early childhood interventions literature: even if gains for firstborn siblings from relatively higher parental investment in their first years of life do not appear at age five, they subsequently emerge in adolescent behavioral outcomes and high stakes exams.[35] However, this explanation cannot account for the reversal of the birth order effect that we also document. Hotz and Pantano (2015) discuss nine plausible mechanisms through which birth order gradients might arise.[36] Most of these theories would predict a firstborn advantage at age five and, in their simplest versions, do not predict a sign reversal between early childhood and adolescence. This throws up the intriguing possibility that there may in fact be different birth order mechanisms at play at different stages of development.

For the population of minority students, we find that firstborns experience large penalties on all age-five skills compared to secondborn siblings. These age-five birth order results run counter to the consensus finding in developed country settings for adolescent and

[35] Studies documenting fade-out followed by longer-term impacts include Chetty et al. (2011), Deming (2009), Gray-Lobe et al. (2023) and Heckman et al. (2013). Recent examples of early childhood intervention studies which find evidence of impact on early childhood cognitive skills but more muted effects on socioemotional skills include Carneiro et al. (2025), and Doyle (2020).

[36] These are: parental time and resource dilution; differential genetic endowments; endogenous fertility; the confluence (or sibling spillovers) model; disruption to family structure; differential parental investment for insurance reasons; parental learning; the easing of financial constraints over time; and strategic parenting.

adult outcomes. The emergence of broadly similar age-five birth order patterns across diverse immigrant groups is difficult to reconcile with explanations based narrowly on cultural norms inherited from origin countries, since these groups differ substantially in language, religion, migration history, and family practices. Instead, we find evidence in favor of positive spillovers in English language skills from the older sibling to the younger one. This firstborn penalty is reversed in adolescence, with firstborns overtaking their secondborn siblings on test scores and high stakes exams, suggesting that any early language advantage between siblings from a minority background fades over time. These results provide novel insights into the skills gaps between white British and minority students. The prior literature has noted that these gaps diminish as students age. For example, Dustmann et al. (2010) show that the white British versus minority student achievement gap in England is largest at the start of school but narrows markedly as students move through the school system. Our results suggest a subtle refinement to the standard interpretation of the narrowing of the majority-minority skills gap. Our findings show that an important part of this convergence process occurs *within* minority families. Among families where English is not spoken at home, firstborn children make substantially greater progress from age five to age 11 than their secondborn siblings. This implies that the closing of the majority-minority gap may be driven, at least in part, by especially rapid catch up among firstborn children in non-English speaking households, rather than a uniform rate of improvement across all minority students.

A key lesson from our analysis of birth order effects in the early childhood development phase is the value of large administrative datasets for uncovering patterns that would otherwise remain hidden. The population-level scale of the data make it possible to detect heterogeneous effects across groups, tracing how birth order gradients evolve from early childhood to the adolescent years. Even without unpacking the precise mechanisms driving birth order effects, our results demonstrate that identifying these effects is crucial for a better understanding of the dynamics of early childhood development. From a policy perspective, if birth order systematically shifts the initial stock of cognitive or noncognitive skills, it can fundamentally alter the gains from all subsequent investments under the assumption of dynamic complementarity. Similarly, birth order may interact with policy interventions targeting early childhood and later outcomes, suggesting that birth order should be taken into account when devising early childhood interventions.

## References


Adler, A. (1928). Characteristics of the first, second and third child. *Children* 3, no. 5:14-52.

Aizer, A., & Cunha, F. (2012). The production of human capital: Endowments, investments and fertility (No. w18429). National Bureau of Economic Research.

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360-1446.

Almond, D., & Mazumder, B. (2013). Fetal origins and parental responses. *Annu. Rev. Econ.*, *5*(1), 37-56.

Andrew, A., Attanasio, O. P., Bernal, R., Sosa, L. C., Krutikova, S., & Rubio-Codina, M. (2024). Preschool quality and child development. *Journal of Political Economy*, *132*(7), 2304-2345.

Attanasio, O. (2026). The first 1000 days and beyond: The process of child development. NBER WP.

Averett, S. L., Argys, L. M., & Rees, D. I. (2011). Older siblings and adolescent risky behavior: does parenting play a role?. *Journal of Population Economics*, *24*(3), 957-978.

Bacher-Hicks, A., Billings, S. B., & Deming, D. J. (2024). The school-to-prison pipeline: Long-run impacts of school suspensions on adult crime. *American Economic Journal: Economic Policy*, *16*(4), 165-193.

Barclay, K. J. (2015). Birth order and educational attainment: Evidence from fully adopted sibling groups. Intelligence, 48, 109-122.

Bagger, J., Birchenall, J. A., Mansour, H., & Urzua, S. (2021). Education, birth order and family size. *The Economic Journal*, *131*(633), 33-69.

Bassok, D., & Reardon, S. F. (2013). "Academic redshirting" in kindergarten: Prevalence, patterns, and implications. *Educational Evaluation and Policy Analysis*, *35*(3), 283-297.

Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, *84*(4, Part 2), S143-S162.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1994). Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. *Journal of political economy*, *102*(6), 1131-1174.

Behrman, J. R., & Taubman, P. (1986). Birth order, schooling, and earnings. *Journal of Labor Economics*, *4*(3, Part 2), S121-S145.

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American economic journal: applied economics*, *5*(1), 32-64.

Bharadwaj, P., Eberhard, J. P., & Neilson, C. A. (2018). Health at birth, parental investments, and academic outcomes. *Journal of labor Economics*, *36*(2), 349-394.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., ... & Simonsen, M. (2021). Sibling spillovers. *The Economic Journal*, *131*(633), 101-128.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, *120*(2), 669-700.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Older and wiser? Birth order and IQ of young men. *CESifo Economic Studies*, *57*(1), 103-120.

Black, S. E., Grönqvist, E., & Öckert, B. (2018). Born to lead? The effect of birth order on noncognitive abilities. *Review of Economics and Statistics*, *100*(2), 274-286.

Blake, J. (1981). Family size and the quality of children. *Demography*, *18*(4), 421-442.

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and statistics*, *86*(2), 481-496.

Bleakley, H., & Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American Economic Journal: Applied Economics*, *2*(1), 165-192.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly journal of economics*, *107*(1), 123-150.

Breining, S., Doyle, J., Figlio, D. N., Karbownik, K., & Roth, J. (2020). Birth order and delinquency: Evidence from Denmark and Florida. *Journal of Labor Economics*, *38*(1), 95-142.

Brenøe, A. A., & Molitor, R. (2018). Birth order and health of newborns: what can we learn from Danish registry data?. *Journal of Population Economics*, *31*(2), 363-395.

Buckles, K. S., & Munnich, E. L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources*, *47*(3), 613-642.

Card, D. (2005). Is the new immigration really so bad?. *The Economic Journal*, *115*(507), F300-F323.

Casey, T., & Dustmann, C. (2008). Intergenerational transmission of language capital and economic outcomes. *Journal of Human Resources*, *43*(3), 660-687.

Cornelissen, T., & Dustmann, C. (2019). Early school exposure, test scores, and noncognitive outcomes. *American Economic Journal: Economic Policy*, *11*(2), 35-63.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American economic review*, *97*(2), 31-47.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883-931.

Currie, J., & Almond, D. (2011). Human capital development before age five. In *Handbook of labor economics* (Vol. 4, pp. 1315-1486). Elsevier.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111-134.

Datar, A., Kilburn, M. R., & Loughran, D. S. (2010). Endowments and parental investments in infancy and early childhood. *Demography*, *47*(1), 145-162.

Daysal, N. M., Ding, H., Rossin-Slater, M., & Schwandt, H. (2026). Germs in the Family: The Short-and Long-Term Consequences of Intrahousehold Disease Spread. *American Economic Review*, *116*(7), 2643-2684.

DCSF (2008). Statutory framework for the Early years foundation stage. May 2008, Department for Children, Schools and Families.

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, *29*(4), 576-588.

Doyle, O. (2020). The first 2,000 days and child skills. *Journal of Political Economy*, *128*(6), 2067-2122.

Duncan, G. J., Kalil, A., Mogstad, M., & Rege, M. (2023). Investing in early childhood development in preschool and at home. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 6, pp. 1–91).

Ejrnæs, M., & Pörtner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics*, 86(4), 1008–1019.

Elder, T., & Zhou, Y. (2021). The black-white gap in noncognitive skills among elementary school children. *American Economic Journal: Applied Economics*, *13*(1), 105-132.

Evangelou, M., Sylva, K., Kyriacou, M., Wild, M., & Glenny, G. (2009). Early years learning and development. A Literature Review. *Department of Children, Schools and Families*.

Feng, S., Han, Y., Heckman, J. J., & Kautz, T. (2022). Comparing the reliability and predictive power of child, teacher, and guardian reports of noncognitive skills. *Proceedings of the National Academy of Sciences*, *119*(6), e2113992119.

Figlio, D., Guryan, J., Karbownik, K., & Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, *104*(12), 3921-3955.

Galton, F. (1874). *English men of science: Their nature and nurture*. Routledge, 2018.

Ganimian, A. J., Muralidharan, K., & Walters, C. R. (2024). Augmenting state capacity for child development: Experimental evidence from India. *Journal of Political Economy*, *132*(5), 1565-1602.

Ginja, R., Jans, J., & Karimi, A. (2020). Parental leave benefits, household labor supply, and children's long-run outcomes. *Journal of Labor Economics*, *38*(1), 261-320.

Grätz, M., & Torche, F. (2016). Compensation or reinforcement? The stratification of parental responses to children's early ability. *Demography*, *53*(6), 1883-1904.

Gray-Lobe, G., Pathak, P. A., & Walters, C. R. (2023). The long-term effects of universal preschool in Boston. *The Quarterly Journal of Economics*, *138*(1), 363-411.

Greaves, E., Hussain, I., Rabe, B., & Rasul, I. (2023). Parental responses to information about school quality: Evidence from linked survey and administrative data. *The Economic Journal*, *133*(654), 2334-2402.

Greaves, E., & Hussain, I. (2026). The importance of school quality ratings for school choices: Evidence from a nationwide system. *Journal of Public Economics*, 263, 105787.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689-733.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052-2086.

Hotz, V. J., & Pantano, J. (2015). Strategic parenting, birth order, and school performance. *Journal of population economics*, *28*(4), 911-936.

Houmark, M. A. (2023). First among equals? How birth order shapes child development. WP, Aarhus University.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER WP 7867.

Isungset, M. A., Freese, J., Andreassen, O. A., & Lyngstad, T. H. (2022). Birth order differences in education originate in postnatal environments. *PNAS nexus*, *1*(2), pgac051.

Isungset, M. A., Lillehagen, M., & Ugreninov, E. (2020). One order fits all? Birth order and education in immigrant families. *European Sociological Review*, *36*(1), 16-31.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of political Economy*, *126*(5), 2072-2107.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, *21*(1), 60-99.

Johnston, A. C., Jones, M. R., & Pope, N. G. (2025). Divorce, family arrangements, and children's adult outcomes (No. w33776). National Bureau of Economic Research.

Jayachandran, S., & Pande, R. (2017). Why are Indian children so short? The role of birth order and son preference. *American Economic Review*, *107*(9), 2600-2629.

Lehmann, J. Y. K., Nuevo-Chiquero, A., & Vidal-Fernandez, M. (2018). The early origins of birth order differences in children's outcomes and parental behavior. *Journal of Human Resources*, *53*(1), 123-156.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67.

Muslimova, D., Van Kippersluis, H., Rietveld, C. A., Von Hinke, S., & Meddens, S. F. W. (2026). Gene-environment complementarity in educational attainment. *Journal of Labor Economics*, *44*(3).

Office for National Statistics (2009), "Marriage, Divorce and Adoption Statistics", https://assets.publishing.service.gov.uk/media/5a7c8c2f40f0b62aff6c2701/9787777148227.pdf.

Papageorge, N. W., Ronda, V., & Zheng, Y. (2019). *The economic value of breaking bad: Misbehavior, schooling and the labor market* (No. w25602). National Bureau of Economic Research.

Pavan, R. (2016). On the production of skills and the birth-order effect. *Journal of Human Resources*, *51*(3), 699-726.

Price, J. (2008). Parent-child quality time: Does birth order matter? *Journal of human resources*, *43*(1), 240-265.

Pruckner, G. J., Schneeweis, N., Schober, T., & Zweimüller, M. (2021). Birth order, parental health investment, and health in childhood. *Journal of Health Economics*, *76*, 102426.

QCA (2003). Foundation Stage Profile Handbook. Department for Children, Schools and Families.

Rohrer, J. M., Egloff, B., & Schmukle, S. C. (2015). Examining the effects of birth order on personality. *Proceedings of the National Academy of Sciences*, *112*(46), 14224-14229.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education* (Vol. 3, pp. 249-277). Elsevier.

Sebba, M. (1993), *London Jamaican: Language Systems in Interaction*, Routledge.

Sebba, M. (2007), "Caribbean Creoles and Black English," in D. Britain, *Language in the British Isles*, pp. 276–292, Cambridge University Press.

Siraj-Blatchford, I., Muttock, S., Sylva, K., Gilden, R., & Bell, D. (2002). *Researching effective pedagogy in the early years* (Vol. 356). London: Department for Education and Skills.

Sulloway, F. J. (1996). *Born to rebel: Birth order, family dynamics, and creative lives*. Pantheon Books.

Thomson, D. (2023). The increase in the number of pupils taught out of year group. FFT Education Datalab.

UCL Centre for Longitudinal Studies. (2026). Millennium Cohort Study. [data series]. 18th Release. UK Data Service. SN: 2000031.

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal*, *125*(588), F347-F371.

Zajonc, R. B., & Markus, G. B. (1975). Birth order and intellectual development. *Psychological review*, *82*(1), 74.

# Online Appendices

## Online Appendix A: Theoretical Background

In this Appendix we provide a simple framework which helps motivate the empirical analysis. Under the resource dilution hypothesis, the skill building framework (Cunha and Heckman, 2007) suggests that there are a number of channels through which birth order may influence children's outcomes. This may be illustrated with a simple three-period model, with two siblings in the family. Let $t = 0$ be the pre-school phase, $t = 1$ end of Reception grade (grade 'zero' in primary school for children age five) and $t = 2$ the adolescent phase, at age 11. At $t = 0$, pre-school cognitive and noncognitive outcomes, $\theta_0^C$ and $\theta_0^{NC}$, respectively, are determined by parental investment, $x_0$, as follows: $\theta_0^C = f_0(x_0)$ and $\theta_0^{NC} = g_0(x_0)$. Later skills are determined by skills in the previous period and current school inputs, proxied by school quality:

$$\theta_t^C = f_t(\theta_{t-1}^C, \theta_{t-1}^{NC}, Q_t), \qquad \theta_t^{NC} = g_t(\theta_{t-1}^{NC}, \theta_{t-1}^C, Q_t),$$

where $\theta_t^C$ and $\theta_t^{NC}$ are the cognitive and noncognitive outcomes at time t, respectively, $\theta_{t-1}^C$ and $\theta_{t-1}^C$ are the previous period skills, $Q_t$ is the quality of the child's school, and $t \in \{1,2\}$. We abstract from age-five and age-11 parental inputs by assuming that parental inputs are equal across school-age children.[37] Crucially, however, under the resource dilution hypothesis, pre-school parental inputs, $x_0$, favour the firstborn child, which are assumed to yield higher pre-school human capital, $\theta_0$.

In light of this framework, a birth order effect emanating from higher initial investment, $x_0$, for the firstborn child, operates through multiple channels. The first channel is via self-productivity, which may be expressed mathematically for age-five cognitive skills, as follows: $\frac{\partial \theta_1^C}{\partial x_0} = \frac{\partial \theta_1^C}{\partial \theta_0^C}\frac{\partial \theta_0^C}{\partial x_0}$. Thus, initially higher investment ($x_0$) for a firstborn child, leads to higher pre-school human capital, which persists into later childhood via self-productivity, so that age-five outcomes for firstborn children are higher than for their secondborn counterparts.

A second channel is via dynamic complementarity (or substitutability), where higher levels of pre-school skills influence the productivity of school inputs, $Q_1$, as expressed by: $\frac{\partial^2 \theta_1^C}{\partial Q_1 \partial x_0} = \frac{\partial^2 \theta_1^C}{\partial Q_1 \partial \theta_0^C}\frac{\partial \theta_0^C}{\partial x_0}$.[38] Under dynamic complementarity (substitutability) returns from attending a higher quality school are larger (smaller) for the firstborn sibling thanks to higher pre-school human capital, $\theta_0$.

Finally, there may be an impact from early parental investment, $x_0$, on later outcomes such as $\theta_t^C$ and $\theta_t^{NC}$, which does not operate via $\theta_{t-1}^C$ and $\theta_{t-1}^{NC}$. For example, very early investments ($x_0$) may boost human capital for firstborn children because this is a particularly sensitive phase of childhood development, but the firstborn impact may not be detected at age five, rather, emerging in adolescence. Such a mechanism may account for latent or 'sleeper' birth order effects (Duncan et al., 2023).

[37] We assume that the age gap between siblings is five years or less.

[38] Similar expressions hold for noncognitive skills as well as for cross-productivities, for example, between early noncognitive and later cognitive skills.

## Online Appendix B: Construction of the Siblings Sample

Reception to Grade 6 sample

Consistent measures of children's cognitive and noncognitive skills in the Reception grade are available from 2007 to 2012, whilst the family identifier variable in the administrative data is available for the years 2008, 2013 and 2016. We drop both siblings in a pair if any of the five Reception grade outcomes (language, numeracy, socioemotional, creative development and physical development skills) is missing for either of the siblings. In addition, pairs of siblings are also dropped if information on ethnicity is missing for both siblings. Note that we refer to each academic year by the calendar year in which it ends, for example, the 2006/07 academic year is denoted 2007. The oldest students in the estimation sample are therefore enrolled in the Reception grade in 2007. In 2008 (the first year that the siblings variable is available in the dataset) these students are enrolled in first grade (known as 'Year 1' in England). [39] For this cohort, the sample is restricted to those students who do not have an older sibling in the dataset in a higher grade (up to grade 11, the final year of compulsory formal academic study) in 2008 and who have at least one younger sibling enrolled in the Reception grade in years 2008 to 2012. For these students the 'grade spacing' gap between siblings ranges from one to five years. For the 2007 Reception cohort, students will be incorrectly coded as the oldest sibling if the gap between the younger sibling and their next older sibling is more than ten years.

Similarly, for siblings inferred to be the oldest in their families and enrolled in the Reception grade in 2008, the sample is restricted to students who do not have an older sibling in a higher grade up to grade 11 in 2008 and who have at least one younger sibling enrolled in the Reception grade in years 2009 to 2012. In this case, the grade spacing gap between siblings ranges from one to four years.

With regards to students who are the oldest sibling and enrolled in the Reception grade in 2009, the 2013 sibling variable is employed to identify older siblings. Students enrolled in Reception in 2009 progress to grade 4 in 2013. Students in grade 4 in 2013 with an older sibling in grades 5 through 11 are dropped from the estimation sample, whilst those with at least one younger sibling in grades 1, 2 and 3 are retained. The grade spacing gap between siblings ranges from one to three years. For the 2009 Reception cohort, students will be incorrectly coded as the oldest sibling if the gap between the younger sibling and their next older sibling is more than seven years.

A similar approach is used for the older sibling enrolled in Reception grades in 2010 and 2011. Beyond this, the older sibling in 2012 will have a younger sibling in 2013 or later. The latter are assessed under a new curriculum and hence comparable data are not available to undertake the early childhood analysis, and hence cannot be included in the estimation sample. Note that for the 2009 and 2010 Reception cohorts, students will be incorrectly coded as the oldest sibling if the gap between the younger sibling and their next older sibling is more than eight and nine years, respectively.

In order to assess the robustness of the main analysis to the possibility that the eldest sibling is incorrectly coded, the sensitivity of the main results is assessed by using just the

[39] Grade retention (repeating a year) is exceptionally rare in English state primary and secondary schools. England's school system strongly favours automatic promotion rather than grade retention. Pupils typically progress automatically to the next year group, regardless of academic performance. Schools address learning gaps through differentiation (tailored teaching) rather than retention.

oldest siblings in the 2007 and 2008 Reception cohorts. For these two cohorts of oldest siblings, the oldest sibling will be miscoded if the gap between the younger sibling and their next older sibling is more than ten or eleven years, respectively. Results for the early childhood outcomes are reported in online Appendix Table 7. These show that there is little change in the results when this sample restriction is imposed, allaying concerns that miscoding of the eldest sibling variable leads to biased estimates.

Age-16 (Key Stage 4) Sample

A similar approach is used to construct the sample employed to undertake analysis of outcomes at age 16, i.e. grade 11 (Year 11). These students are enrolled in grade 11 in the six years 2014 to 2019. In 2008 (the first year of availability of the siblings variable) these students are enrolled in Reception through to grade 5. Starting with grade 5 students in 2008, selection is on those grade 5 students with no older sibling in grades 6 to 11. Of these, those with at least one younger sibling in grades 4 through Reception are retained. Similarly, moving on to grade 4 students in 2008, selection is on grade 4 students with no older sibling in grades 5 to 11; those with at least one younger sibling in grades 3 through Reception are retained. This procedure is repeated for grade 3, 2 and 1 students in 2008.

As before, birth order one and two siblings are retained for the estimation sample, although robustness analysis is also undertaken where higher birth order siblings are available. This then constitutes the estimations ample for the age-16 analysis. Furthermore, information on these students, relating to absences, suspensions and special needs status, is also retained for grades 7 through 11.

## Online Appendix C: Parental Investment Data in the Millennium Cohort Study

We employ data from the Millennium Cohort Study (MCS) to investigate associations between birth order and parental investments. The MCS is a nationally representative birth cohort study following children born in the United Kingdom between 2000 and 2002. The study questionnaires include information on parental investments when the child is age three (wave 2 of the MCS) and five (wave 3). Specifically, these measures are as follows: how frequently the parent (carer) reads to the child (at age three and five); whether anyone at home helps the child with the alphabet and reading (at age three); whether anyone at home helps the child with reading, writing or mathematics (at age five); and whether the carer or their partner attended a parents' evening at the child's school (at age five). We recode the frequency with which the parent or carer reads to the child as a binary variable equal to one if the parent reports reading to the child 'several times a week' or 'every day or almost every day', and zero otherwise. We regress each of these outcomes on the birth order of the study child. Control variables include the following: dummy variables for the mother's fertility at the time of the survey; a cubic polynomial for the study child's age in months as well as their sex and year of birth. Further parental controls include main carer characteristics: age in years, sex, religion, schooling level, carer and partner's labour market status, whether the carer and their partner live in the same household, whether English is the only language spoken at home and the main carer's age when study child was born. We exclude singletons from the analysis. We account for the survey's sampling design by applying sampling weights. Our conclusions remain the same if sample weights are not used. The analysis is undertaken

separately for the whole British sample and the minority sample. The latter sample consists of those who have mixed, Indian, Pakistani/Bangladeshi, Black or Other ethnicities.

## Online Appendix D: The Early Years Foundation Stage Profile

The six broad Early Years Foundation Stage (EYFS) categories are divided into 13 sub-categories, or assessment scales: four under Communication, Language and Literacy, three under Mathematical Development, three under Personal, Social and Emotional Development, and one each for Physical Development, Creative Development, and Knowledge and Understanding of the World. Each sub-category is scored separately from 0 to 9 according to the child's demonstrated achievements. This yields a maximum total score of 117 points across all 13 scales. Source: QCA (2003).

**Communication, language and literacy**

Language for Communication and Thinking
1 Listens and responds.
2 Initiates communication with others, displaying greater confidence in more informal contexts.
3 Talks activities through, reflecting on and modifying actions.
4 Listens with enjoyment to stories, songs, rhymes and poems, sustains attentive listening and responds with relevant comments, questions or actions.
5 Uses language to imagine and recreate roles and experiences.
6 Interacts with others in a variety of contexts, negotiating plans and activities and taking turns in conversation.
7 Uses talk to organise, sequence and clarify thinking, ideas, feelings and events, exploring the meanings and sounds of new words.
8 Speaks clearly with confidence and control, showing awareness of the listener.
9 Talks and listens confidently and with control, consistently showing awareness of the listener by including relevant detail. Uses language to work out and clarify ideas, showing control of a range of appropriate vocabulary.

Linking Sounds and Letters
1 Joins in with rhyming and rhythmic activities.
2 Shows an awareness of rhyme and alliteration.
3 Links some sounds to letters.
4 Links sounds to letters, naming and sounding letters of the alphabet.
5 Hears and says sounds in words.
6 Blends sounds in words.
7 Uses phonic knowledge to read simple regular words.
8 Attempts to read more complex words, using phonic knowledge.
9 Uses knowledge of letters, sounds and words when reading and writing independently.

Reading
1 Is developing an interest in books.
2 Knows that print conveys meaning.
3 Recognises a few familiar words.
4 Knows that, in English, print is read from left to right and top to bottom.
5 Shows an understanding of the elements of stories, such as main character, sequence of events and openings.
6 Reads a range of familiar and common words and simple sentences independently.

7 Retells narratives in the correct sequence, drawing on language patterns of stories.
8 Shows an understanding of how information can be found in non-fiction texts to answer questions about where, who, why and how.
9 Reads books of own choice with some fluency and accuracy.

Writing
1 Experiments with mark-making, sometimes ascribing meaning to the marks.
2 Uses some clearly identifiable letters to communicate meaning.
3 Represents some sounds correctly in writing.
4 Writes own name and other words from memory.
5 Holds a pencil and uses it effectively to form recognisable letters, most of which are correctly formed.
6 Attempts writing for a variety of purposes, using features of different forms.
7 Uses phonic knowledge to write simple regular words and make phonetically plausible attempts at more complex words.
8 Begins to form captions and simple sentences, sometimes using punctuation.
9 Communicates meaning through phrases and simple sentences with some consistency in punctuating sentences.

**Mathematical development**

Numbers as Labels and for Counting
1 Says some number names in familiar contexts, such as nursery rhymes.
2 Counts reliably up to three everyday objects.
3 Counts reliably up to six everyday objects.
4 Says number names in order.
5 Recognises numerals 1 to 9.
6 Counts reliably up to 10 everyday objects.
7 Orders numbers, up to 10.
8 Uses developing mathematical ideas and methods to solve practical problems.
9 Recognises, counts, orders, writes and uses numbers up to 20.

Calculating
1 Responds to the vocabulary involved in addition and subtraction in rhymes and games.
2 Recognises differences in quantity when comparing sets of objects.
3 Finds one more or one less from a group of up to five objects.
4 Relates addition to combining two groups.
5 Relates subtraction to taking away.
6 In practical activities and discussion, begins to use the vocabulary involved in adding and subtracting.
7 Finds one more or one less than a number from 1 to 10.
8 Uses developing mathematical ideas and methods to solve practical problems.
9 Uses a range of strategies for addition and subtraction, including some mental recall of number bonds.

Shape, Space and Measures
1 Experiments with a range of objects and materials showing some mathematical awareness.
2 Sorts or matches objects and talks about sorting.
3 Describes shapes in simple models, pictures and patterns.
4 Talks about, recognises and recreates simple patterns.
5 Uses everyday words to describe position.

6 Uses language such as ‘circle’ or ‘bigger’ to describe the shape and size of solids and flat shapes.
7 Uses language such as ‘greater’, ‘smaller’, ‘heavier’ or ‘lighter’ to compare quantities.
8 Uses developing mathematical ideas and methods to solve practical problems.
9 Uses mathematical language to describe solid (3D) objects and flat (2D) shapes.

**Personal, social and emotional development**

Dispositions and Attitudes
1 Shows an interest in classroom activities through observation or participation.
2 Dresses, undresses and manages own personal hygiene with adult support.
3 Displays high levels of involvement in self-chosen activities.
4 Dresses and undresses independently and manages own personal hygiene.
5 Selects and uses activities and resources independently.
6 Continues to be interested, motivated and excited to learn.
7 Is confident to try new activities, initiate ideas and speak in a familiar group.
8 Maintains attention and concentrates.
9 Sustains involvement and perseveres, particularly when trying to solve a problem or reach a satisfactory conclusion.

Social Development
1 Plays alongside others.
2 Builds relationships through gesture and talk.
3 Takes turns and shares with adult support.
4 Works as part of a group or class, taking turns and sharing fairly.
5 Forms good relationships with adults and peers.
6 Understands that there need to be agreed values and codes of behaviour for groups of people, including adults and children, to work together harmoniously.
7 Understands that people have different needs, views, cultures and beliefs that need to be treated with respect.
8 Understands that s/he can expect others to treat her or his needs, views, cultures and beliefs with respect.
9 Takes into account the ideas of others.

Emotional Development
1 Separates from main carer with support.
2 Communicates freely about home and community.
3 Expresses needs and feelings in appropriate ways.
4 Responds to significant experiences, showing a range of feelings when appropriate.
5 Has a developing awareness of own needs, views and feelings and is sensitive to the needs, views and feelings of others.
6 Has a developing respect for own culture and beliefs and those of other people.
7 Considers the consequences of words and actions for self and others.
8 Understands what is right, what is wrong, and why.
9 Displays a strong and positive sense of self-identity and is able to express a range of emotions fluently and appropriately.

**Creative development**

1 Explores different media and responds to a variety of sensory experiences. Engages in representational play.

2 Creates simple representations of events, people and objects and engages in music making.
3 Tries to capture experiences, using a variety of different media.
4 Sings simple songs from memory.
5 Explores colour, texture, shape, form and space in two or three dimensions.
6 Recognises and explores how sounds can be changed. Recognises repeated sounds and sound patterns and matches movements to music.
7 Uses imagination in art and design, music, dance, imaginative and role-play and stories. Responds in a variety of ways to what s/he sees, hears, smells, touches and feels.
8 Expresses and communicates ideas, thoughts and feelings using a range of materials, suitable tools, imaginative and role-play, movement, designing and making, and a variety of songs and musical instruments.
9 Expresses feelings and preferences in response to artwork, drama and music and makes some comparisons and links between different pieces. Responds to own work and that of others when exploring and communicating ideas, feelings and preferences through art, music, dance, role-play and imaginative play.

**Physical Development**

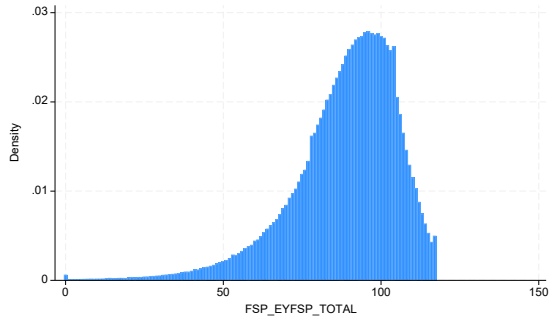
1 Moves spontaneously, showing some control and coordination.
2 Moves with confidence in a variety of ways, showing some awareness of space.
3 Usually shows appropriate control in large- and small-scale movements.
4 Moves with confidence, imagination and in safety. Travels around, under, over and through balancing and climbing equipment. Shows awareness of space, of self and others.
5 Demonstrates fine motor control and coordination.
6 Uses small and large equipment, showing a range of basic skills.
7 Handles tools, objects, construction and malleable materials safely and with basic control.
8 Recognises the importance of keeping healthy and those things which contribute to this. Recognises the changes that happen to her/his body when s/he is active.
9 Repeats, links and adapts simple movements, sometimes commenting on her/his work. Demonstrates coordination and control in large and small movements, and in using a range of tools and equipment.

**Knowledge and Understanding of the World**

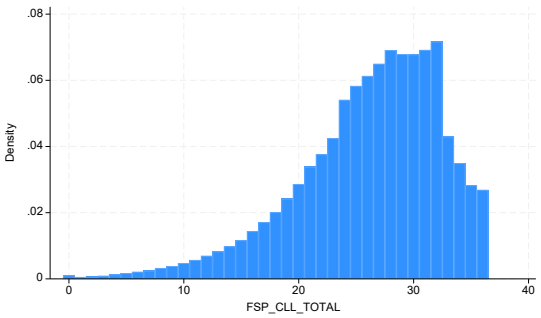
1 Shows curiosity and interest by exploring surroundings.
2 Observes, selects and manipulates objects and materials. Identifies simple features and significant personal events.
3 Identifies obvious similarities and differences when exploring and observing. Constructs in a purposeful way, using simple tools and techniques.
4 Investigates places, objects, materials and living things by using all the senses as appropriate. Identifies some features and talks about those features s/he likes and dislikes.
5 Asks questions about why things happen and how things work. Looks closely at similarities, differences, patterns and change.
6 Finds out about past and present events in own life, and in those of family members and other people s/he knows. Begins to know about own culture and beliefs and those of other people.
7 Finds out about and identifies the uses of everyday technology and uses information and communication technology and programmable toys to support her/his learning.
8 Builds and constructs with a wide range of objects, selecting appropriate resources, tools and techniques and adapting her/his work where necessary.
9 Communicates simple planning for investigations and constructions and makes simple records and evaluations of her/his work. Identifies and names key features and properties, sometimes linking

different experiences, observations and events. Begins to explore what it means to belong to a variety of groups and communities.

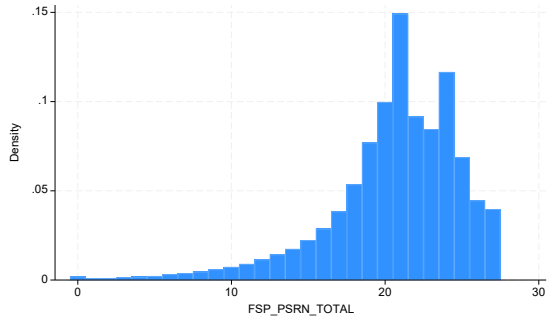
# Figure 1 Distribution of Age-Five Early Years Foundation Stage Skills



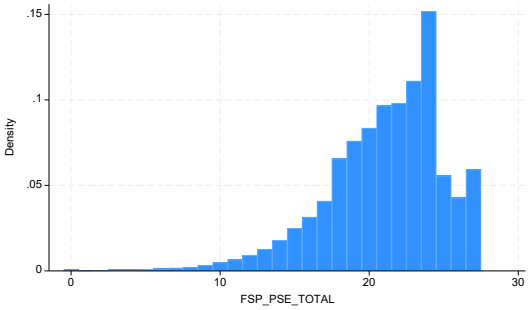

Total EYFS (maximum score: 118)



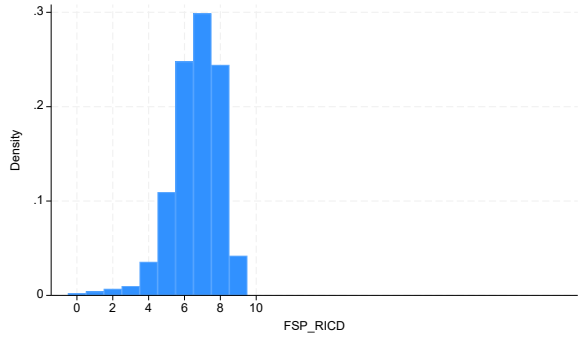

Language (maximum score: 37)



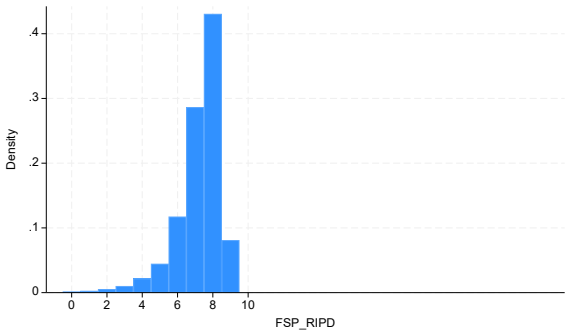

Numeracy (maximum score: 28)




Socioemotional dev. (maximum score: 28)




Creative development (maximum score: 10)




Physical development (maximum: 10)

Notes: Histograms plot for the estimation sample the raw scores for the following Reception age-five Early Years Foundation Stage skills: the total Early Years Foundation Stage profile score; Communication, Language and Literacy ('Language'); Problem Solving, Reasoning and Numeracy ('Numeracy'); Personal, Social, and Emotional Development ('Socioemotional'); Creative Development; and Physical Development. All histogram bins contain a minimum of 50 observations.

# Table 1 Descriptive Statistics

| | Full sample | Estimation sample | White British | Minorities |
|---|---|---|---|---|
| Free lunch | 0.319 | 0.290 | 0.262 | 0.390 |
| | (0.466) | (0.454) | (0.440) | (0.488) |
| Female | 0.487 | 0.487 | 0.486 | 0.490 |
| | (0.500) | (0.500) | (0.500) | (0.500) |
| Minority | 0.267 | 0.217 | 0 | 1 |
| | (0.443) | (0.412) | | |
| Age gap between siblings | | 2.401 | 2.422 | 2.322 |
| | | (0.953) | (0.938) | (1.002) |
| Age-five skills | | | | |
| Language | 0.00931 | 0.103 | 0.157 | -0.0950 |
| | (0.985) | (0.959) | (0.929) | (1.037) |
| | | [0.53] | [0.514] | [0.582] |
| Numeracy | 0.00840 | 0.0963 | 0.159 | -0.130 |
| | (0.984) | (0.950) | (0.908) | (1.058) |
| | | [0.532] | [0.511] | [0.603] |
| Socioemotional | 0.00541 | 0.0876 | 0.137 | -0.0915 |
| | (0.987) | (0.958) | (0.931) | (1.029) |
| | | [0.561] | [0.545] | [0.614] |
| Creative development | 0.00452 | 0.0684 | 0.125 | -0.135 |
| | (0.984) | (0.960) | (0.936) | (1.019) |
| | | [0.584] | [0.571] | [0.628] |
| Physical development | 0.00420 | 0.0683 | 0.108 | -0.0755 |
| | (0.988) | (0.951) | (0.934) | (0.994) |
| | | [0.583] | [0.573] | [0.619] |
| Age-11 test scores | | | | |
| Reading | 0.0196 | 0.0967 | 0.111 | 0.0446 |
| | (0.987) | (0.981) | (0.985) | (0.965) |
| | | [0.514] | [0.517] | [0.502] |
| Mathematics | 0.00774 | 0.0988 | 0.0811 | 0.163 |
| | (0.995) | (0.973) | (0.973) | (0.970) |
| | | [0.502] | [0.506] | [0.485] |
| Observations | 3401862 | 885,686 | 693,632 | 192,054 |
| Families | | 442,843 | 346,816 | 96,027 |

Notes: The full sample consist of students in the Reception grade in the years 2007 to 2012. The estimation sample consists of the two oldest siblings with age-five outcomes from one of the years 2007 to 2012. We drop both siblings in a pair if any of the five Reception grade outcomes is missing for either of the siblings. The free lunch ('free school meal') indicator is set to 1 if the child is eligible for free lunch in any grade between Reception and grade (Year) 6. Age gap is measured in years between the two siblings. Reception age-five Early Years Foundation Stage skills: Communication, Language and Literacy ('Language'); Problem Solving, Reasoning and Numeracy ('Numeracy'); Personal, Social, and Emotional Development ('Socioemotional'); Creative Development; and Physical Development. Age-11 test scores are the Key Stage 2 assessments on Reading comprehension achievement and Mathematics achievement. All age-five and age-11 scores are standardized within each national cohort to have mean zero and unit standard deviation. Standard deviation shown in parentheses. Withn-family standard deviation shown in brackets; calculated as the standard deviation of the residuals from a regression of each skill measure on the family fixed effect.

# Table 2 Relationship Between Age-5, Age-11 and Later Outcomes

| | Age-16 exam performance | | | Ever suspended, Grades 1 - 6 | Ever suspended, Grades 7 - 11 | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Age-5 skills: | | | | | | | |
| Language | 0.359*** | | 0.0918*** | 0.00132** | -0.0167*** | | 0.000793 |
| | (0.00457) | | (0.00373) | (0.000538) | (0.00178) | | (0.00181) |
| Numeracy | 0.163*** | | -0.0270*** | 0.00330*** | -0.0112*** | | 0.000892 |
| | (0.00426) | | (0.00341) | (0.000580) | (0.00172) | | (0.00174) |
| Socioemotional | 0.0732*** | | 0.0589*** | -0.0124*** | -0.0365*** | | -0.0355*** |
| | (0.00353) | | (0.00282) | (0.000589) | (0.00143) | | (0.00142) |
| Creative development | -0.0408*** | | -0.00888*** | 0.00155*** | 0.0100*** | | 0.00805*** |
| | (0.00337) | | (0.00265) | (0.000406) | (0.00128) | | (0.00127) |
| Physical development | -0.0541*** | | -0.0360*** | -0.000217 | 0.0175*** | | 0.0162*** |
| | (0.00305) | | (0.00242) | (0.000437) | (0.00125) | | (0.00125) |
| Age-11 test scores: | | | | | | | |
| Reading | | 0.316*** | 0.296*** | | | -0.0234*** | -0.0210*** |
| | | (0.00206) | (0.00211) | | | (0.00102) | (0.00105) |
| Mathematics | | 0.383*** | 0.362*** | | | -0.0240*** | -0.0216*** |
| | | (0.00211) | (0.00222) | | | (0.00105) | (0.00111) |
| Observations | 246,337 | 246,337 | 246,337 | 246,337 | 246,337 | 246,337 | 246,337 |

*Notes* : Standard errors are clustered at the school level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. This table relates nationally standardized age-16 exam performance ('GCSE' exams) and school suspensions to age-5 and age-11 test scores. 'Evere suspended' is a binary dummy variable indicating whether the student was ever suspended in the respective grades. All estimates are from OLS regressions, which include indicators for free lunch status, female, month and year of birth, and ethnic group, as well as age-five school fixed effects. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 3 Impact of Birth Order on Age-Five Outcomes

| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: All** | | | | | |
| Second child | -0.00546*** | 0.0290*** | 0.0273*** | 0.0109*** | 0.0484*** |
| | (0.00149) | (0.00154) | (0.00163) | (0.00168) | (0.00171) |
| Girl | 0.304*** | 0.103*** | 0.311*** | 0.443*** | 0.260*** |
| | (0.00210) | (0.00217) | (0.00229) | (0.00236) | (0.00241) |
| Observations | 885,686 | 885,686 | 885,686 | 885,686 | 885,686 |
| Sibling pairs | 442,843 | 442,843 | 442,843 | 442,843 | 442,843 |
| **Panel B: White British** | | | | | |
| Second child | -0.0518*** | -0.0136*** | 0.000141 | -0.0128*** | 0.0311*** |
| | (0.00163) | (0.00167) | (0.00178) | (0.00185) | (0.00190) |
| Girl | 0.302*** | 0.0908*** | 0.304*** | 0.444*** | 0.264*** |
| | (0.00228) | (0.00234) | (0.00250) | (0.00259) | (0.00266) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 |
| Families | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 |
| **Panel C: Minorities** | | | | | |
| Second child | 0.161*** | 0.182*** | 0.125*** | 0.0962*** | 0.110*** |
| | (0.00352) | (0.00372) | (0.00382) | (0.00390) | (0.00391) |
| Girl | 0.313*** | 0.152*** | 0.336*** | 0.441*** | 0.247*** |
| | (0.00504) | (0.00533) | (0.00547) | (0.00558) | (0.00560) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 |
| Families | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills in: Communication, Language and Literacy ('Language'); Problem Solving, Reasoning and Numeracy ('Numeracy'); Personal, Social, and Emotional Development ('Socioemotional'); Creative Development; and Physical Development. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 4 Sibling Spacing Interaction

| | Age-Five | | | | | Age-11 | |
|---|---|---|---|---|---|---|---|
| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. | Reading | Maths |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Panel A: White British** | | | | | | | |
| Second child | 0.0434*** | 0.0560*** | 0.0166*** | 0.0226*** | 0.0479*** | -0.102*** | -0.0240*** |
| | (0.00400) | (0.00411) | (0.00439) | (0.00456) | (0.00467) | (0.00455) | (0.00450) |
| Second child | -0.0686*** | -0.0417*** | -0.00848* | -0.0186*** | -0.00158 | -0.0243*** | -0.0242*** |
| x Grade spacing = 2 | (0.00464) | (0.00477) | (0.00510) | (0.00529) | (0.00542) | (0.00526) | (0.00520) |
| Second child | -0.145*** | -0.113*** | -0.0262*** | -0.0629*** | -0.0318*** | -0.0382*** | -0.0550*** |
| x Grade spacing = 3 | (0.00516) | (0.00531) | (0.00568) | (0.00589) | (0.00604) | (0.00585) | (0.00578) |
| Second child | -0.193*** | -0.154*** | -0.0390*** | -0.0752*** | -0.0545*** | -0.0217*** | -0.0568*** |
| x Grade spacing = 4 | (0.00653) | (0.00672) | (0.00718) | (0.00745) | (0.00764) | (0.00737) | (0.00728) |
| Second child | -0.201*** | -0.156*** | -0.0510*** | -0.0758*** | -0.0518*** | -0.0273** | -0.0704*** |
| x Grade spacing = 5 | (0.0104) | (0.0107) | (0.0115) | (0.0119) | (0.0122) | (0.0118) | (0.0116) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 | 638,852 | 638,852 |
| Families | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 | 319,426 | 319,426 |
| **Panel B: Minorities** | | | | | | | |
| Second child | 0.218*** | 0.224*** | 0.122*** | 0.112*** | 0.121*** | -0.0101 | 0.0297*** |
| | (0.00750) | (0.00794) | (0.00815) | (0.00832) | (0.00834) | (0.00734) | (0.00712) |
| Second child | -0.0286*** | -0.0152 | 0.0191* | 0.00348 | 0.00908 | -0.0232** | -0.00444 |
| x Grade spacing = 2 | (0.00935) | (0.00990) | (0.0102) | (0.0104) | (0.0104) | (0.00914) | (0.00887) |
| Second child | -0.119*** | -0.0881*** | -0.0171 | -0.0511*** | -0.0347*** | -0.0313*** | 0.0105 |
| x Grade spacing = 3 | (0.0107) | (0.0113) | (0.0116) | (0.0118) | (0.0118) | (0.0104) | (0.0101) |
| Second child | -0.127*** | -0.107*** | -0.0115 | -0.0366** | -0.0427*** | -0.00717 | 0.0150 |
| x Grade spacing = 4 | (0.0135) | (0.0143) | (0.0147) | (0.0150) | (0.0150) | (0.0132) | (0.0128) |
| Second child | -0.127*** | -0.118*** | 0.00737 | -0.0545** | -0.0412* | -0.00881 | 0.00940 |
| x Grade spacing = 5 | (0.0209) | (0.0222) | (0.0228) | (0.0232) | (0.0233) | (0.0204) | (0.0198) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 | 174,984 | 174,984 |
| Families | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 | 87,492 | 87,492 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. 'Grade spacing = k' is a dummy variable indicating that the two siblings are k grades apart from each other. All regressions control for family fixed effects and include dummy variables for month of birth and gender. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 5 Alternative Measures for Age-Five Outcomes

| | Percentile ranks | | | | | Total age-5 score | Cognitive skills factor | Non-cognitive skills factor |
|---|---|---|---|---|---|---|---|---|
| | Numeracy | Language | Socio-emotional | Creative dev. | Physical dev. | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Panel A: White British** | | | | | | | | |
| Second child | -0.0205*** | -0.00714*** | -0.00420*** | -0.0149*** | -0.00144*** | -0.0228*** | -0.0326*** | -0.00710*** |
| | (0.000480) | (0.000502) | (0.000518) | (0.000527) | (0.000489) | (0.00162) | (0.00169) | (0.00183) |
| Girl | 0.0916*** | 0.0196*** | 0.0894*** | 0.132*** | 0.0624*** | 0.263*** | 0.182*** | 0.421*** |
| | (0.000672) | (0.000703) | (0.000725) | (0.000738) | (0.000685) | (0.00227) | (0.00237) | (0.00256) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 | 693,619 | 693,632 |
| Families | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 |
| **Panel B: Minorities** | | | | | | | | |
| Second child | 0.0385*** | 0.0425*** | 0.0288*** | 0.00911*** | 0.0161*** | 0.164*** | 0.186*** | 0.125*** |
| | (0.000949) | (0.000980) | (0.00102) | (0.000993) | (0.000910) | (0.00358) | (0.00378) | (0.00393) |
| Girl | 0.0848*** | 0.0295*** | 0.0887*** | 0.115*** | 0.0491*** | 0.293*** | 0.225*** | 0.437*** |
| | (0.00136) | (0.00140) | (0.00146) | (0.00142) | (0.00131) | (0.00513) | (0.00541) | (0.00564) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 | 192,046 | 192,054 |
| Families | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. For columns 1 to 5, age-five assessment scores are expressed as percentile ranks within each national cohort. For column 6, the total age-five skills score is the sum of the scores for language, numeracy, socioemotional, creative development, physical development, and knowledge and understanding of the world; the summed score is standardized within each national cohort to have mean zero and unit standard deviation. For columns 7 and 8, two factor-indices are constructed: a non-cognitive factor (the first factor from the socioemotional and creative development skill measures) and a cognitive factor (the first factor from language and mathematics skills as well as knowledge and understanding of the world), respectively. Each factor is standardized to mean 0 and unit variance.

# Table 6 Birth Order and Age-Five Outcomes: Free Lunch Status

| | Language | | Numeracy | | Socioemotional | | Creative dev. | | Physical dev. | |
|---|---|---|---|---|---|---|---|---|---|---|
| | FL = 0 | FL = 1 | FL = 0 | FL = 1 | FL = 0 | FL = 1 | FL = 0 | FL = 1 | FL = 0 | FL = 1 |
| **Panel A: White British** | | | | | | | | | | |
| Second child | -0.0633*** | -0.0222*** | -0.0223*** | 0.00827** | 0.00171 | -0.00424 | -0.0180*** | 0.000424 | 0.0334*** | 0.0244*** |
| | (0.00181) | (0.00350) | (0.00181) | (0.00374) | (0.00200) | (0.00377) | (0.00208) | (0.00389) | (0.00209) | (0.00413) |
| Girl | 0.285*** | 0.347*** | 0.0733*** | 0.137*** | 0.289*** | 0.345*** | 0.437*** | 0.460*** | 0.251*** | 0.297*** |
| | (0.00252) | (0.00493) | (0.00253) | (0.00527) | (0.00279) | (0.00532) | (0.00290) | (0.00549) | (0.00293) | (0.00581) |
| Observations | 502,988 | 190,644 | 502,988 | 190,644 | 502,988 | 190,644 | 502,988 | 190,644 | 502,988 | 190,644 |
| Families | 251,494 | 95,322 | 251,494 | 95,322 | 251,494 | 95,322 | 251,494 | 95,322 | 251,494 | 95,322 |
| **Panel B: Minorities** | | | | | | | | | | |
| Second child | 0.130*** | 0.205*** | 0.151*** | 0.226*** | 0.110*** | 0.146*** | 0.0767*** | 0.123*** | 0.102*** | 0.122*** |
| | (0.00442) | (0.00574) | (0.00460) | (0.00618) | (0.00481) | (0.00620) | (0.00493) | (0.00631) | (0.00487) | (0.00643) |
| Girl | 0.296*** | 0.337*** | 0.128*** | 0.187*** | 0.323*** | 0.356*** | 0.439*** | 0.443*** | 0.236*** | 0.264*** |
| | (0.00629) | (0.00829) | (0.00655) | (0.00893) | (0.00686) | (0.00897) | (0.00702) | (0.00912) | (0.00694) | (0.00929) |
| Observations | 112,188 | 79,866 | 112,188 | 79,866 | 112,188 | 79,866 | 112,188 | 79,866 | 112,188 | 79,866 |
| Families | 56,094 | 39,933 | 56,094 | 39,933 | 56,094 | 39,933 | 56,094 | 39,933 | 56,094 | 39,933 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table reports results separately for students eligible for Free Lunch ('Free School Meal') (FL = 1) versus those not eligible for Free Lunch (FL = 0). Both siblings are classified as belonging to a free lunch status family if the free lunch status variable is equal to one in any grade for either sibling. Outcomes are teacher assessment scores for the Early Years Foundation Stage skills, standardized within each national cohort to have mean zero and unit standard deviation.

# Table 7 Birth order and Early Years: Detailed Minority Status

| | Heritage | | | | | | |
|---|---|---|---|---|---|---|---|
| | White British | African | Afro-Caribbean | Bangladeshi | Indian | Pakistani | White other |
| **Language** | | | | | | | |
| Second child | -0.0545*** | 0.274*** | 0.102*** | 0.228*** | 0.127*** | 0.153*** | 0.122*** |
| | (0.00169) | (0.0106) | (0.0186) | (0.0161) | (0.0103) | (0.00887) | (0.00808) |
| Girl | 0.301*** | 0.338*** | 0.357*** | 0.358*** | 0.306*** | 0.308*** | 0.297*** |
| | (0.00237) | (0.0148) | (0.0260) | (0.0264) | (0.0148) | (0.0131) | (0.0114) |
| Observations | 639,308 | 24,356 | 6,642 | 10,204 | 20,562 | 32,348 | 32,324 |
| Families | 319,654 | 12,178 | 3,321 | 5,102 | 10,281 | 16,174 | 16,162 |
| **Numeracy** | | | | | | | |
| Second child | -0.0156*** | 0.276*** | 0.137*** | 0.259*** | 0.144*** | 0.177*** | 0.146*** |
| | (0.00174) | (0.0112) | (0.0193) | (0.0175) | (0.0106) | (0.00955) | (0.00856) |
| Girl | 0.0902*** | 0.209*** | 0.168*** | 0.197*** | 0.151*** | 0.178*** | 0.100*** |
| | (0.00243) | (0.0157) | (0.0271) | (0.0287) | (0.0152) | (0.0141) | (0.0121) |
| Observations | 639,308 | 24,356 | 6,642 | 10,204 | 20,562 | 32,348 | 32,324 |
| Families | 319,654 | 12,178 | 3,321 | 5,102 | 10,281 | 16,174 | 16,162 |
| **Socioemotional** | | | | | | | |
| Second child | -0.00198 | 0.204*** | 0.113*** | 0.165*** | 0.113*** | 0.0930*** | 0.112*** |
| | (0.00186) | (0.0116) | (0.0208) | (0.0174) | (0.0112) | (0.00946) | (0.00874) |
| Girl | 0.304*** | 0.413*** | 0.383*** | 0.336*** | 0.300*** | 0.304*** | 0.319*** |
| | (0.00260) | (0.0163) | (0.0291) | (0.0285) | (0.0160) | (0.0140) | (0.0123) |
| Observations | 639,308 | 24,356 | 6,642 | 10,204 | 20,562 | 32,348 | 32,324 |
| Families | 319,654 | 12,178 | 3,321 | 5,102 | 10,281 | 16,174 | 16,162 |
| **Creative dev.** | | | | | | | |
| Second child | -0.0146*** | 0.143*** | 0.0882*** | 0.145*** | 0.0864*** | 0.0829*** | 0.0933*** |
| | (0.00192) | (0.0115) | (0.0204) | (0.0179) | (0.0114) | (0.00977) | (0.00904) |
| Girl | 0.443*** | 0.478*** | 0.451*** | 0.458*** | 0.426*** | 0.391*** | 0.442*** |
| | (0.00269) | (0.0162) | (0.0285) | (0.0293) | (0.0164) | (0.0144) | (0.0128) |
| Observations | 639,308 | 24,356 | 6,642 | 10,204 | 20,562 | 32,348 | 32,324 |
| Families | 319,654 | 12,178 | 3,321 | 5,102 | 10,281 | 16,174 | 16,162 |
| **Physical dev.** | | | | | | | |
| Second child | 0.0296*** | 0.152*** | 0.106*** | 0.155*** | 0.0920*** | 0.0909*** | 0.122*** |
| | (0.00197) | (0.0119) | (0.0211) | (0.0177) | (0.0112) | (0.0103) | (0.00900) |
| Girl | 0.263*** | 0.321*** | 0.249*** | 0.208*** | 0.214*** | 0.230*** | 0.243*** |
| | (0.00276) | (0.0167) | (0.0295) | (0.0290) | (0.0160) | (0.0152) | (0.0127) |
| Observations | 639,308 | 24,356 | 6,642 | 10,204 | 20,562 | 32,348 | 32,324 |
| Families | 319,654 | 12,178 | 3,321 | 5,102 | 10,281 | 16,174 | 16,162 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Table reports results separately for students from different minority ethnic heritage groups. All regressions control for family fixed effects and include dummy variables for month of birth. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 8 Impact for Minority Students: English Spoken at Home Status

| | Language | | Numeracy | | Socioemotional | | Creative dev. | | Physical dev. | |
|---|---|---|---|---|---|---|---|---|---|---|
| | English = 1 | English = 0 | English = 1 | English = 0 | English = 1 | English = 0 | English = 1 | English = 0 | English = 1 | English = 0 |
| Panel A: Age 5 | | | | | | | | | | |
| Second child | 0.0475*** | 0.218*** | 0.0803*** | 0.234*** | 0.0574*** | 0.159*** | 0.0339*** | 0.127*** | 0.0724*** | 0.129*** |
| | (0.00572) | (0.00442) | (0.00596) | (0.00470) | (0.00632) | (0.00477) | (0.00642) | (0.00488) | (0.00646) | (0.00489) |
| Girl | 0.331*** | 0.304*** | 0.138*** | 0.159*** | 0.351*** | 0.329*** | 0.458*** | 0.432*** | 0.270*** | 0.235*** |
| | (0.00810) | (0.00637) | (0.00845) | (0.00678) | (0.00895) | (0.00688) | (0.00910) | (0.00704) | (0.00914) | (0.00705) |
| Observations | 63,976 | 127,960 | 63,976 | 127,960 | 63,976 | 127,960 | 63,976 | 127,960 | 63,976 | 127,960 |
| Families | 31,988 | 63,980 | 31,988 | 63,980 | 31,988 | 63,980 | 31,988 | 63,980 | 31,988 | 63,980 |

| Panel B: Age 11 and 16 | KS2 | | | | KS4 | |
|---|---|---|---|---|---|---|
| | Reading | | Mathematics | | Point Score | Point Score |
| | English = 1 | English = 0 | English = 1 | English = 0 | English = 1 | English = 0 |
| Second child | -0.0670*** | -0.00717 | -0.0164*** | 0.0563*** | -0.0715*** | -0.0312*** |
| | (0.00605) | (0.00522) | (0.00600) | (0.00498) | (0.00628) | (0.00441) |
| Girl | 0.198*** | 0.135*** | -0.0647*** | -0.105*** | 0.282*** | 0.248*** |
| | (0.00858) | (0.00754) | (0.00850) | (0.00720) | (0.00885) | (0.00649) |
| Observations | 77,104 | 96,328 | 77,104 | 96,328 | 56,963 | 102,266 |
| Families | 48,891 | 58,527 | 48,891 | 58,527 | 33,094 | 55,899 |

*Notes*: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Table reports results for students from the minority ethnic group, separately for those families where English is spoken at home (English = 1) versus those where English is not spoken at home (English = 0). All regressions control for family fixed effects and include dummy variables for month of birth. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 9 Mechanisms: Delayed School Entry, School Quality and Peers

| | Language | Numeracy | Socio-emotional | Creative dev. | Physical dev. |
|---|---|---|---|---|---|
| **Panel A: Delayed School Entry** | | | | | |
| Second child | -0.0608*** | -0.0207*** | -0.00561*** | -0.0194*** | 0.0254*** |
| | (0.00163) | (0.00168) | (0.00180) | (0.00186) | (0.00191) |
| Dealyed entry | -0.225*** | -0.168*** | -0.162*** | -0.168*** | -0.147*** |
| | (0.00608) | (0.00626) | (0.00669) | (0.00693) | (0.00712) |
| Observations | 682,410 | 682,410 | 682,410 | 682,410 | 682,410 |
| Families | 341,205 | 341,205 | 341,205 | 341,205 | 341,205 |
| **Panel B: School Quality** | | | | | |
| Second child | -0.0493*** | -0.0106*** | 0.00187 | -0.0103*** | 0.0336*** |
| | (0.00161) | (0.00164) | (0.00176) | (0.00183) | (0.00187) |
| Latest school inspection grade | 0.00824*** | 0.0182*** | 0.0297*** | 0.0308*** | 0.0298*** |
| | (0.00241) | (0.00245) | (0.00264) | (0.00273) | (0.00281) |
| Observations | 670,150 | 670,150 | 670,150 | 670,150 | 670,150 |
| Families | 335,075 | 335,075 | 335,075 | 335,075 | 335,075 |
| **Panel C: Peer Quality** | | | | | |
| Second child | -0.0537*** | -0.0131*** | 0.00192 | -0.0109*** | 0.0329*** |
| | (0.00165) | (0.00170) | (0.00182) | (0.00188) | (0.00193) |
| Mean girls in cohort | 0.0762*** | 0.186*** | 0.0825*** | 0.0357** | 0.0711*** |
| | (0.0124) | (0.0128) | (0.0136) | (0.0141) | (0.0145) |
| Mean free lunch in cohort | -0.0615*** | -0.170*** | -0.169*** | -0.213*** | -0.185*** |
| | (0.0156) | (0.0160) | (0.0171) | (0.0177) | (0.0182) |
| Observations | 682,126 | 682,126 | 682,126 | 682,126 | 682,126 |
| Families | 341,063 | 341,063 | 341,063 | 341,063 | 341,063 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Results shown for the white Brtish sample. All regressions control for family fixed effects and include dummy variables for month of birth as well as gender. Regressions in Panel A include an indicator for whether the child experienced delayed enrollment in the Reception grade. Regressions in Panel B include the school's latest inspection grade in the year of Reception grade enrollment for each child. Regressions in Panel C include leave-self-out means for the number of girls in the Reception cohort and the number of free lunch status students in the Reception cohort. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 10 Adolescent Academic Outcomes

| | White British | | Minorities | |
|---|---|---|---|---|
| | Reading | Mathematics | Reading | Mathematics |
| **Panel A:** | | | | |
| **Age-11 test score** | | | | |
| Second child | -0.126*** | -0.0577*** | -0.0278*** | 0.0322*** |
| | (0.00183) | (0.00181) | (0.00343) | (0.00333) |
| Girl | 0.166*** | -0.106*** | 0.158*** | -0.0901*** |
| | (0.00256) | (0.00253) | (0.00492) | (0.00478) |
| Observations | 638,852 | 638,852 | 174,984 | 174,984 |
| Families | 319,426 | 319,426 | 87,492 | 87,492 |
| **Panel B:** | White British | | Minorities | |
| **Age-16, GCSE total points score** | | | | |
| Second child | -0.148*** | | -0.0463*** | |
| | (0.00158) | | (0.00340) | |
| Girl | 0.252*** | | 0.261*** | |
| | (0.00222) | | (0.00493) | |
| Observations | 676,452 | | 160,262 | |
| Families | 338,226 | | 80,131 | |

*Notes*: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table shows the impact of birth order on age-11 and age-16 academic performance. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Table 11 School Suspensions

| | Ever suspended, Grades Reception to 6 | Ever suspended, Grades 7 to 11 |
|---|---|---|
| **Panel A: White British** | | |
| Second child | 0.000725*** | 0.01000*** |
| | (0.000272) | (0.000672) |
| Girl | -0.0203*** | -0.0682*** |
| | (0.000381) | (0.000941) |
| Observations | 693,632 | 676,452 |
| Families | 346,816 | 338,226 |
| **Panel B: Minorities** | | |
| Second child | -0.000940* | 0.00971*** |
| | (0.000511) | (0.00151) |
| Girl | -0.0192*** | -0.0992*** |
| | (0.000733) | (0.00219) |
| Observations | 192,054 | 160,262 |
| Families | 96,027 | 80,131 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table shows the impact of birth order on school suspensions. 'Ever suspended' is a binary dummy variable indicating whether the student was ever suspended in the respective grades.

# Appendix Table 1 Relationship Between Age-5 and Age-11 Outcomes, Family Fixed Effects

| | Reading | | Mathematics | |
|---|---|---|---|---|
| | OLS | Family FE | OLS | Family FE |
| Age-5 skills: | | | | |
| Language | 0.446*** | 0.323*** | 0.375*** | 0.274*** |
| | (0.00232) | (0.00299) | (0.00230) | (0.00292) |
| Numeracy | 0.212*** | 0.146*** | 0.316*** | 0.239*** |
| | (0.00224) | (0.00278) | (0.00224) | (0.00272) |
| Socioemotional | 0.00547*** | 0.00990*** | -0.00566*** | 0.00694*** |
| | (0.00180) | (0.00225) | (0.00179) | (0.00221) |
| Creative development | -0.0329*** | -0.0303*** | -0.0788*** | -0.0544*** |
| | (0.00156) | (0.00195) | (0.00155) | (0.00189) |
| Physical development | -0.0679*** | -0.0255*** | -0.0252*** | 0.000433 |
| | (0.00166) | (0.00200) | (0.00168) | (0.00197) |
| Observations | 813,836 | 813,836 | 813,836 | 813,836 |
| R-squared | 0.312 | 0.165 | 0.321 | 0.185 |
| Families | | 406,918 | | 406,918 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. This table relates nationally standardized age-11 test scores (Key Stage 2) to age-5 Early Years Foundation Stage skills. Age-11 and age-5 scores are standardized within each national cohort to have mean zero and unit standard deviation. Sample consists of sibling pairs. OLS regression estimates include indicators for free lunch status, female, month and year of birth, and ethnic group, as well as age-five school fixed effects. Family fixed effect estimates include a gender dummy and indicators for month of birth.

# Appendix Table 2 Birth Order and Age-Five Outcomes: OLS Estimates

| | Language | Numeracy | Socioemotional | Creative dvpt. | Physical dvpt. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **All** | | | | | |
| Second child | -0.0212*** | 0.0140*** | 0.0145*** | -0.000662 | 0.0380*** |
| | (0.00250) | (0.00250) | (0.00287) | (0.00301) | (0.00282) |
| Girl | 0.312*** | 0.108*** | 0.316*** | 0.449*** | 0.263*** |
| | (0.00194) | (0.00198) | (0.00205) | (0.00226) | (0.00216) |
| Observations | 885,686 | 885,686 | 885,686 | 885,686 | 885,686 |
| **White British** | | | | | |
| Second child | -0.0651*** | -0.0261*** | -0.0112*** | -0.0232*** | 0.0216*** |
| | (0.00252) | (0.00252) | (0.00301) | (0.00317) | (0.00299) |
| Girl | 0.309*** | 0.0946*** | 0.310*** | 0.450*** | 0.267*** |
| | (0.00209) | (0.00210) | (0.00222) | (0.00246) | (0.00236) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 |
| **Minorities** | | | | | |
| Second child | 0.140*** | 0.161*** | 0.108*** | 0.0800*** | 0.0970*** |
| | (0.00508) | (0.00533) | (0.00575) | (0.00585) | (0.00569) |
| Girl | 0.324*** | 0.158*** | 0.336*** | 0.446*** | 0.252*** |
| | (0.00457) | (0.00469) | (0.00470) | (0.00482) | (0.00468) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 |

*Notes* : Standard errors are clustered at the age-five primary school level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. OLS regressions include indicators for free lunch status, gender, month of birth, ethnic group, neighborhood deprivation index and its square, as well as indicators for missing ethnicity (recoded as white British when missing) and neighborhood deprivation index (recoded with mean value when missing). The outcomes are teacher assessment scores for the Early Years Foundation Stage skills in: Communication, Language and Literacy ('Language'); Problem Solving, Reasoning and Numeracy ('Numeracy'); Personal, Social, and Emotional Development ('Socioemotional'); Creative Development; and Physical Development. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 3 Sibling Spacing Interaction (Spacing in Years of Age)

| | Age-Five | | | | | Age-11 | |
|---|---|---|---|---|---|---|---|
| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. | Reading | Maths |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **White British** | | | | | | | |
| Second child | 0.0368*** | 0.0535*** | 0.0152*** | 0.0212*** | 0.0474*** | -0.0991*** | -0.0282*** |
| | (0.00423) | (0.00435) | (0.00465) | (0.00482) | (0.00494) | (0.00483) | (0.00478) |
| Second child | -0.0611*** | -0.0415*** | -0.00644 | -0.0190*** | -0.00493 | -0.0269*** | -0.0206*** |
| x Birth spacing = 2 | (0.00488) | (0.00502) | (0.00537) | (0.00557) | (0.00571) | (0.00556) | (0.00550) |
| Second child | -0.132*** | -0.102*** | -0.0246*** | -0.0560*** | -0.0247*** | -0.0401*** | -0.0442*** |
| x Birth spacing = 3 | (0.00520) | (0.00535) | (0.00572) | (0.00593) | (0.00608) | (0.00590) | (0.00584) |
| Second child | -0.176*** | -0.144*** | -0.0368*** | -0.0707*** | -0.0499*** | -0.0233*** | -0.0505*** |
| x Birth spacing = 4 | (0.00645) | (0.00664) | (0.00709) | (0.00735) | (0.00754) | (0.00730) | (0.00721) |
| Second child | -0.193*** | -0.159*** | -0.0465*** | -0.0743*** | -0.0585*** | -0.0411*** | -0.0816*** |
| x Birth spacing = 5 | (0.00995) | (0.0102) | (0.0109) | (0.0113) | (0.0116) | (0.0112) | (0.0111) |
| Second child | -0.168*** | -0.111*** | 0.0126 | -0.0770** | -0.0393 | 0.00512 | -0.0450 |
| x Birth spacing = 6 | (0.0299) | (0.0308) | (0.0329) | (0.0341) | (0.0350) | (0.0339) | (0.0335) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 | 638,852 | 638,852 |
| Families | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 | 319,426 | 319,426 |
| **Minorities** | | | | | | | |
| Second child | 0.220*** | 0.232*** | 0.133*** | 0.117*** | 0.126*** | -0.0130* | 0.0336*** |
| x Birth spacing = 2 | (0.00757) | (0.00801) | (0.00822) | (0.00839) | (0.00841) | (0.00740) | (0.00718) |
| Second child | -0.0372*** | -0.0325*** | -0.00210 | -0.00798 | -0.00434 | -0.0160* | -0.0114 |
| x Birth spacing = 3 | (0.00931) | (0.00986) | (0.0101) | (0.0103) | (0.0104) | (0.00910) | (0.00883) |
| Second child | -0.102*** | -0.0811*** | -0.0214* | -0.0451*** | -0.0298** | -0.0265*** | 0.0149 |
| x Birth spacing = 4 | (0.0104) | (0.0111) | (0.0113) | (0.0116) | (0.0116) | (0.0102) | (0.00989) |
| Second child | -0.146*** | -0.126*** | -0.0271* | -0.0508*** | -0.0594*** | -0.0243* | -0.0119 |
| x Birth spacing = 5 | (0.0132) | (0.0140) | (0.0143) | (0.0146) | (0.0147) | (0.0129) | (0.0125) |
| Second child | -0.117*** | -0.110*** | 0.0103 | -0.0379* | -0.0332 | 0.0129 | 0.0331* |
| x Birth spacing = 6 | (0.0199) | (0.0211) | (0.0216) | (0.0221) | (0.0221) | (0.0194) | (0.0188) |
| second_born_YrsSpace6 | -0.0841 | -0.0859 | -0.00272 | -0.0815 | 0.00114 | 0.0226 | -0.0207 |
| | (0.0576) | (0.0610) | (0.0626) | (0.0639) | (0.0641) | (0.0560) | (0.0543) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 | 174,984 | 174,984 |
| Families | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 | 87,492 | 87,492 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. 'Birth spacing = k' is a dummy variable indicating that the two siblings are born k years apart from each other. All regressions control for family fixed effects and include dummy variables for month of birth and gender. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 4 Birth Order and Age-5 Outcomes: Time-Varying Controls

| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: White British** | | | | | |
| Second child | -0.0564*** | -0.0178*** | -0.00321* | -0.0166*** | 0.0273*** |
| | (0.00164) | (0.00168) | (0.00180) | (0.00187) | (0.00191) |
| Girl | 0.302*** | 0.0907*** | 0.304*** | 0.443*** | 0.264*** |
| | (0.00227) | (0.00234) | (0.00250) | (0.00259) | (0.00265) |
| Observations | 693,632 | 693,632 | 693,632 | 693,632 | 693,632 |
| Families | 346,816 | 346,816 | 346,816 | 346,816 | 346,816 |
| **Panel B: Minorities** | | | | | |
| Second child | 0.151*** | 0.172*** | 0.117*** | 0.0868*** | 0.104*** |
| | (0.00351) | (0.00372) | (0.00383) | (0.00390) | (0.00392) |
| Girl | 0.312*** | 0.151*** | 0.336*** | 0.440*** | 0.247*** |
| | (0.00500) | (0.00530) | (0.00545) | (0.00556) | (0.00558) |
| Observations | 192,054 | 192,054 | 192,054 | 192,054 | 192,054 |
| Families | 96,027 | 96,027 | 96,027 | 96,027 | 96,027 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects, dummy variables for month of birth, as well as the following time-varying controls: free lunch status and the home location local poverty index are included, both measured at the time of Reception (age four or five) enrollment. Teacher assessment scores for the Early Years Foundation Stage skills are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 5 Impact of Birth Order in Three-Child Families

| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: White British** | | | | | |
| Second child | -0.0310*** | 0.0104 | -0.00181 | -0.00109 | 0.0439*** |
| | (0.00641) | (0.00675) | (0.00692) | (0.00717) | (0.00744) |
| Third child | -0.0932*** | -0.0369*** | -0.0251*** | -0.0186** | 0.0279*** |
| | (0.00653) | (0.00688) | (0.00705) | (0.00730) | (0.00758) |
| Girl | 0.298*** | 0.0911*** | 0.295*** | 0.430*** | 0.254*** |
| | (0.00651) | (0.00686) | (0.00703) | (0.00728) | (0.00756) |
| Observations | 74,175 | 74,175 | 74,175 | 74,167 | 74,172 |
| Families | 24,725 | 24,725 | 24,725 | 24,725 | 24,725 |
| **Panel B: Minorities** | | | | | |
| Second child | 0.199*** | 0.215*** | 0.134*** | 0.122*** | 0.129*** |
| | (0.0112) | (0.0119) | (0.0120) | (0.0122) | (0.0124) |
| Third child | 0.222*** | 0.250*** | 0.166*** | 0.153*** | 0.134*** |
| | (0.0113) | (0.0121) | (0.0122) | (0.0124) | (0.0126) |
| Girl | 0.293*** | 0.129*** | 0.296*** | 0.400*** | 0.220*** |
| | (0.0115) | (0.0123) | (0.0124) | (0.0126) | (0.0128) |
| Observations | 30,150 | 30,150 | 30,150 | 30,142 | 30,147 |
| Families | 10,050 | 10,050 | 10,050 | 10,050 | 10,050 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and dummy variables for month of birth. Sample consists of families with three siblings with early childhood outcomes in the 2007 – 2012 period. Teacher assessment scores for the Early Years Foundation Stage skills are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 6 Age-Five Results: Within-Group Standardization of Outcomes

| **White British** | Language (1) | Numeracy (2) | Socioemotional (3) | Creative dev. (4) | Physical dev. (5) |
|---|---|---|---|---|---|
| Second child | -0.0508*** | -0.0122*** | -0.00201 | -0.0135*** | 0.0279*** |
| | (0.00170) | (0.00177) | (0.00185) | (0.00192) | (0.00194) |
| Girl | 0.316*** | 0.0964*** | 0.316*** | 0.460*** | 0.270*** |
| | (0.00238) | (0.00248) | (0.00260) | (0.00269) | (0.00272) |
| Observations | 693,470 | 693,470 | 693,470 | 693,470 | 693,470 |
| Families | 346,817 | 346,817 | 346,817 | 346,817 | 346,817 |
| **Minorities** | | | | | |
| Second child | 0.108*** | 0.127*** | 0.0926*** | 0.0597*** | 0.0912*** |
| | (0.00328) | (0.00338) | (0.00359) | (0.00369) | (0.00378) |
| Girl | 0.292*** | 0.138*** | 0.317*** | 0.417*** | 0.239*** |
| | (0.00470) | (0.00485) | (0.00515) | (0.00529) | (0.00542) |
| Observations | 192,030 | 192,030 | 192,030 | 192,030 | 192,030 |
| Families | 96,052 | 96,052 | 96,052 | 96,052 | 96,052 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and dummy variables for month of birth. Teacher assessment scores for the Early Years Foundation Stage skills are standardized within each cohort's national distribution, separately for white British and minority students.

# Appendix Table 7 Robustness to potential miscoding of firstborn child

| | Language | Numeracy | Socioemotional | Creative dev. | Physical dev. |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **White British** | | | | | |
| Second child | -0.0769*** | -0.0327*** | -0.00760*** | -0.0230*** | 0.0233*** |
| | (0.00213) | (0.00219) | (0.00235) | (0.00242) | (0.00249) |
| Girl | 0.302*** | 0.0931*** | 0.303*** | 0.436*** | 0.269*** |
| | (0.00301) | (0.00309) | (0.00331) | (0.00342) | (0.00352) |
| Observations | 399,644 | 399,644 | 399,644 | 399,613 | 399,626 |
| Families | 199,822 | 199,822 | 199,822 | 199,822 | 199,822 |
| **Minorities** | | | | | |
| Second child | 0.134*** | 0.158*** | 0.119*** | 0.0829*** | 0.0961*** |
| | (0.00473) | (0.00502) | (0.00515) | (0.00526) | (0.00530) |
| Girl | 0.316*** | 0.155*** | 0.335*** | 0.438*** | 0.253*** |
| | (0.00688) | (0.00729) | (0.00749) | (0.00765) | (0.00770) |
| Observations | 105,846 | 105,846 | 105,846 | 105,832 | 105,842 |
| Families | 52,923 | 52,923 | 52,923 | 52,923 | 52,923 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and dummy variables for month of birth. Sample consists of oldest siblings in the 2007 and 2008 Reception cohorts; for these two cohorts of oldest siblings, the oldest sibling will be miscoded if the gap between the younger sibling and their next older sibling is more than ten or eleven years, respectively. Teacher assessment scores for the Early Years Foundation Stage skills are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 8 Birth Order and Age-Five Outcomes: Poverty Index Tercile

| | Language | | | Numeracy | | | Socioemotional | | | Creative dev. | | | Physical dev. | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | tercile = 1 | tercile = 2 | tercile = 3 | tercile = 1 | tercile = 2 | tercile = 3 | tercile = 1 | tercile = 2 | tercile = 3 | tercile = 1 | tercile = 2 | tercile = 3 | tercile = 1 | tercile = 2 | tercile = 3 |
| **White British** | | | | | | | | | | | | | | | |
| Second child | -0.0666*** | -0.0588*** | -0.0302*** | -0.0225*** | -0.0200*** | 0.00101 | 0.00463 | -0.000721 | -0.00417 | -0.0179*** | -0.0149*** | -0.00598* | 0.0410*** | 0.0317*** | 0.0201*** |
| | (0.00263) | (0.00278) | (0.00302) | (0.00264) | (0.00284) | (0.00319) | (0.00293) | (0.00307) | (0.00326) | (0.00305) | (0.00318) | (0.00337) | (0.00305) | (0.00323) | (0.00355) |
| Girl | 0.281*** | 0.296*** | 0.330*** | 0.0651*** | 0.0820*** | 0.125*** | 0.282*** | 0.299*** | 0.332*** | 0.437*** | 0.444*** | 0.449*** | 0.248*** | 0.266*** | 0.278*** |
| | (0.00368) | (0.00389) | (0.00424) | (0.00368) | (0.00397) | (0.00447) | (0.00409) | (0.00430) | (0.00458) | (0.00426) | (0.00446) | (0.00473) | (0.00426) | (0.00453) | (0.00498) |
| Observations | 230,832 | 230,820 | 230,822 | 230,832 | 230,820 | 230,822 | 230,832 | 230,820 | 230,822 | 230,832 | 230,820 | 230,822 | 230,832 | 230,820 | 230,822 |
| Families | 115,416 | 115,410 | 115,411 | 115,416 | 115,410 | 115,411 | 115,416 | 115,410 | 115,411 | 115,416 | 115,410 | 115,411 | 115,416 | 115,410 | 115,411 |
| **Minorities** | | | | | | | | | | | | | | | |
| Second child | 0.0973*** | 0.181*** | 0.205*** | 0.129*** | 0.195*** | 0.223*** | 0.104*** | 0.130*** | 0.141*** | 0.0721*** | 0.108*** | 0.108*** | 0.103*** | 0.108*** | 0.119*** |
| | (0.00576) | (0.00619) | (0.00630) | (0.00597) | (0.00658) | (0.00675) | (0.00633) | (0.00670) | (0.00680) | (0.00648) | (0.00683) | (0.00693) | (0.00637) | (0.00684) | (0.00707) |
| Girl | 0.296*** | 0.313*** | 0.328*** | 0.113*** | 0.160*** | 0.182*** | 0.337*** | 0.331*** | 0.340*** | 0.455*** | 0.434*** | 0.430*** | 0.256*** | 0.236*** | 0.248*** |
| | (0.00816) | (0.00889) | (0.00912) | (0.00846) | (0.00946) | (0.00976) | (0.00897) | (0.00963) | (0.00983) | (0.00918) | (0.00982) | (0.0100) | (0.00902) | (0.00983) | (0.0102) |
| Observations | 63,908 | 63,916 | 63,898 | 63,908 | 63,916 | 63,898 | 63,908 | 63,916 | 63,898 | 63,908 | 63,916 | 63,898 | 63,908 | 63,916 | 63,898 |
| Families | 31,954 | 31,958 | 31,949 | 31,954 | 31,958 | 31,949 | 31,954 | 31,958 | 31,949 | 31,954 | 31,958 | 31,949 | 31,954 | 31,958 | 31,949 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Table reports results separately families in each tercile of the neighborhood deprivation index distribution. Neighborhood deprivation is measured using the Income Deprivation Affecting Children Index, IDACI, an indicator based on the share of children living in disadvantaged families in the local geographical area (Lower-layer Super Output Area). All regressions control for family fixed effects and include dummy variables for month of birth. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 9 Heterogeneous Impact on Age-Five Outcomes by Gender Composition

| | Language skills | | | Numeracy skills | | | Socioemotional skills | | | Creative dev. | | | Physical dev. | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GG | BB | G/B | GG | BB | G/B | GG | BB | G/B | GG | BB | G/B | GG | BB | G/B |
| **White British** | | | | | | | | | | | | | | | |
| Second child | -0.0648*** | -0.0457*** | -0.0489*** | -0.0291*** | 0.00769** | -0.0175*** | -0.00798** | 0.00584 | 0.000970 | -0.00379 | -0.0261*** | -0.00991*** | 0.00725** | 0.0528*** | 0.0309*** |
| | (0.00311) | (0.00331) | (0.00231) | (0.00318) | (0.00343) | (0.00237) | (0.00338) | (0.00368) | (0.00253) | (0.00348) | (0.00383) | (0.00263) | (0.00346) | (0.00396) | (0.00271) |
| Girl | | | 0.302*** | | | 0.0908*** | | | 0.304*** | | | 0.444*** | | | 0.264*** |
| | | | (0.00228) | | | (0.00234) | | | (0.00250) | | | (0.00259) | | | (0.00268) |
| Observations | 164,004 | 183,916 | 345,712 | 164,004 | 183,916 | 345,712 | 164,004 | 183,916 | 345,712 | 164,004 | 183,916 | 345,712 | 164,004 | 183,916 | 345,712 |
| Families | 82,002 | 91,958 | 172,856 | 82,002 | 91,958 | 172,856 | 82,002 | 91,958 | 172,856 | 82,002 | 91,958 | 172,856 | 82,002 | 91,958 | 172,856 |
| **Minorities** | | | | | | | | | | | | | | | |
| Second child | 0.173*** | 0.171*** | 0.150*** | 0.181*** | 0.205*** | 0.171*** | 0.121*** | 0.135*** | 0.122*** | 0.0998*** | 0.0963*** | 0.0943*** | 0.0972*** | 0.127*** | 0.108*** |
| | (0.00661) | (0.00705) | (0.00511) | (0.00686) | (0.00757) | (0.00540) | (0.00709) | (0.00769) | (0.00556) | (0.00722) | (0.00786) | (0.00568) | (0.00710) | (0.00802) | (0.00567) |
| Girl | | | 0.313*** | | | 0.152*** | | | 0.336*** | | | 0.441*** | | | 0.247*** |
| | | | (0.00505) | | | (0.00535) | | | (0.00550) | | | (0.00561) | | | (0.00561) |
| Observations | 48,364 | 52,110 | 91,580 | 48,364 | 52,110 | 91,580 | 48,364 | 52,110 | 91,580 | 48,364 | 52,110 | 91,580 | 48,364 | 52,110 | 91,580 |
| Families | 24,182 | 26,055 | 45,790 | 24,182 | 26,055 | 45,790 | 24,182 | 26,055 | 45,790 | 24,182 | 26,055 | 45,790 | 24,182 | 26,055 | 45,790 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. Table reports results separately for samples of girl-girl ('GG'), boy-boy ('BB'), and mixed girl-boy ('G/B') pairs of siblings. All regressions control for family fixed effects and include dummy variables for month of birth. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. The outcomes are teacher assessment scores for the Early Years Foundation Stage skills. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 10 Mechanisms: Delayed School Entry

| | Age Five | | | | | Age 11 | |
|---|---|---|---|---|---|---|---|
| | Language | Numeracy | Socio-emotional | Creative dev. | Physical dev. | Reading | Maths |
| **White British** | | | | | | | |
| Second child | -0.0608*** | -0.0207*** | -0.00561*** | -0.0194*** | 0.0254*** | -0.126*** | -0.0580*** |
| | (0.00163) | (0.00168) | (0.00180) | (0.00186) | (0.00191) | (0.00185) | (0.00183) |
| Dealyed entry | -0.225*** | -0.168*** | -0.162*** | -0.168*** | -0.147*** | 0.000466 | 0.00420 |
| | (0.00608) | (0.00626) | (0.00669) | (0.00693) | (0.00712) | (0.00697) | (0.00688) |
| Observations | 682,410 | 682,410 | 682,410 | 682,410 | 682,410 | 633,676 | 633,676 |
| Families | 341,205 | 341,205 | 341,205 | 341,205 | 341,205 | 319,103 | 319,103 |
| **Minorities** | | | | | | | |
| Second child | 0.135*** | 0.158*** | 0.104*** | 0.0737*** | 0.0930*** | -0.0281*** | 0.0316*** |
| | (0.00355) | (0.00376) | (0.00388) | (0.00395) | (0.00397) | (0.00352) | (0.00342) |
| Dealyed entry | -0.266*** | -0.226*** | -0.209*** | -0.217*** | -0.178*** | -0.0115 | -0.00736 |
| | (0.0106) | (0.0112) | (0.0115) | (0.0118) | (0.0118) | (0.0106) | (0.0103) |
| Observations | 186,920 | 186,920 | 186,920 | 186,920 | 186,920 | 172,729 | 172,729 |
| Families | 93,460 | 93,460 | 93,460 | 93,460 | 93,460 | 87,404 | 87,404 |

Notes: Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth as well as gender. Regressions also include an indicator for whether the child experienced delayed enrollment in the Reception grade. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 11 Mechanisms: School Quality

| | Age Five | | | | | Age 11 | |
|---|---|---|---|---|---|---|---|
| | Language | Numeracy | Socio-emotional | Creative dev. | Physical dev. | Reading | Maths |
| **White British** | | | | | | | |
| Second child | -0.0493*** | -0.0106*** | 0.00187 | -0.0103*** | 0.0336*** | -0.124*** | -0.0556*** |
| | (0.00161) | (0.00164) | (0.00176) | (0.00183) | (0.00187) | (0.00216) | (0.00213) |
| Latest school inspection grade | 0.00824*** | 0.0182*** | 0.0297*** | 0.0308*** | 0.0298*** | 0.00254*** | 0.00263*** |
| | (0.00241) | (0.00245) | (0.00264) | (0.00273) | (0.00281) | (0.000759) | (0.000748) |
| Observations | 670,150 | 670,150 | 670,150 | 670,150 | 670,150 | 559,841 | 559,841 |
| Families | 335,075 | 335,075 | 335,075 | 335,075 | 335,075 | 300,032 | 300,032 |
| **Minorities** | | | | | | | |
| Second child | 0.165*** | 0.185*** | 0.126*** | 0.0963*** | 0.112*** | -0.0260*** | 0.0353*** |
| | (0.00348) | (0.00366) | (0.00376) | (0.00384) | (0.00386) | (0.00396) | (0.00384) |
| Latest school inspection grade | 0.0106** | 0.0144*** | 0.0288*** | 0.0320*** | 0.0304*** | -4.69e-05 | 0.00440*** |
| | (0.00493) | (0.00518) | (0.00533) | (0.00543) | (0.00546) | (0.00142) | (0.00137) |
| Observations | 184,676 | 184,676 | 184,676 | 184,676 | 184,676 | 157,432 | 157,432 |
| Families | 92,338 | 92,338 | 92,338 | 92,338 | 92,338 | 83,553 | 83,553 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth as well as gender. Regressions also include the school's latest inspection grade in the year of Reception grade enrollment for each child. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 12 Mechanisms: Peer Quality

| | Age Five | | | | | Age 11 | |
|---|---|---|---|---|---|---|---|
| | Language | Numeracy | Socio-emotional | Creative dev. | Physical dev. | Reading | Maths |
| **White British** | | | | | | | |
| Second child | -0.0537*** | -0.0131*** | 0.00192 | -0.0109*** | 0.0329*** | -0.127*** | -0.0589*** |
| | (0.00165) | (0.00170) | (0.00182) | (0.00188) | (0.00193) | (0.00184) | (0.00182) |
| Mean girls in cohort | 0.0762*** | 0.186*** | 0.0825*** | 0.0357** | 0.0711*** | 0.0571*** | 0.167*** |
| | (0.0124) | (0.0128) | (0.0136) | (0.0141) | (0.0145) | (0.0142) | (0.0140) |
| Mean free lunch in cohort | -0.0615*** | -0.170*** | -0.169*** | -0.213*** | -0.185*** | -0.0682*** | -0.0867*** |
| | (0.0156) | (0.0160) | (0.0171) | (0.0177) | (0.0182) | (0.0201) | (0.0199) |
| Observations | 682,126 | 682,126 | 682,126 | 682,126 | 682,126 | 638,807 | 638,807 |
| Families | 341,063 | 341,063 | 341,063 | 341,063 | 341,063 | 319,426 | 319,426 |
| **Minorities** | | | | | | | |
| Second child | 0.150*** | 0.170*** | 0.116*** | 0.0854*** | 0.103*** | -0.0298*** | 0.0317*** |
| | (0.00351) | (0.00371) | (0.00382) | (0.00389) | (0.00391) | (0.00362) | (0.00351) |
| Mean girls in cohort | 0.274*** | 0.429*** | 0.339*** | 0.262*** | 0.238*** | 0.0568* | 0.102*** |
| | (0.0319) | (0.0337) | (0.0347) | (0.0353) | (0.0355) | (0.0310) | (0.0301) |
| Mean free lunch in cohort | -0.105*** | -0.138*** | -0.182*** | -0.214*** | -0.173*** | -0.0638* | -0.0191 |
| | (0.0291) | (0.0308) | (0.0317) | (0.0323) | (0.0325) | (0.0343) | (0.0333) |
| Observations | 186,866 | 186,866 | 186,866 | 186,866 | 186,866 | 174,980 | 174,980 |
| Families | 93,433 | 93,433 | 93,433 | 93,433 | 93,433 | 87,492 | 87,492 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth as well as gender. Regressions also include leave-self-out means for the number of girls in the Reception cohort and the number of free lunch status students in the Reception cohort. Sample consists of the oldest two sibling with age-five outcomes from one of the years 2007 to 2012. Age-five skills are from the Early Years Foundation Stage and age-11 test scores are from the Key Stage 2 assessments. Scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 13 Birth Order and Parental Investment

| | How often do you read to | | If anyone helps child with | | If the carer or their partner | |
|---|---|---|---|---|---|---|
| | White British | Minority | White British | Minority | White British | Minority |
| **Panel A: Wave 2** | | | | | | |
| **(Study child is 3 years old)** | | | | | | |
| Birth order = 2 | -0.043*** | 0.013 | -0.066*** | -0.040 | | |
| | (0.011) | (0.038) | (0.012) | (0.028) | | |
| Birth order = 3 | -0.090*** | 0.017 | -0.072*** | 0.023 | | |
| | (0.025) | (0.063) | (0.026) | (0.047) | | |
| Birth order = 4 or higher | -0.054 | 0.121 | -0.118*** | 0.143** | | |
| | (0.050) | (0.094) | (0.045) | (0.071) | | |
| N | 8090 | 1631 | 8090 | 1631 | | |
| mean of outcome | 0.789 | 0.541 | 0.780 | 0.818 | | |
| **Panel B: Wave 3** | | | | | | |
| **(Study child is 5 years old)** | | | | | | |
| Birth order = 2 | -0.076*** | -0.008 | -0.001 | -0.003 | 0.011 | 0.006 |
| | (0.010) | (0.039) | (0.001) | (0.006) | (0.007) | (0.020) |
| Birth order = 3 | -0.102*** | -0.098** | -0.007 | -0.009 | 0.017 | -0.024 |
| | (0.022) | (0.042) | (0.005) | (0.007) | (0.013) | (0.029) |
| Birth order = 4 or higher | -0.116*** | -0.091 | 0.000 | 0.033** | -0.022 | 0.017 |
| | (0.043) | (0.066) | (0.009) | (0.014) | (0.022) | (0.045) |
| N | 8543 | 1602 | 8424 | 1585 | 7801 | 1538 |
| mean of outcome | 0.806 | 0.654 | 0.990 | 0.986 | 0.940 | 0.930 |

Note: Data source: Millennoium Cohort Study. The estimation sample excludes families where the cohort study child is an only child at the time when the outcome was recorded, as well as study children with a twin. Estimation accounts for the survey design weights. The birth order of the study child is calculated relative to the number of natural children of their carer (usually the mother). All rgressions inlcuid eth following control variables: are age of the study child in days (as a cubic), sex, year they were born, the number of natural children of their carer, parents' age (in years), sex , religion, educational level, working status, whether they are present at home, whether they speak other language at home and age when the study child was born.

# Appendix Table 14 Adolescent Academic Outcomes: Percentile Ranks

| | White British | | Minorities | |
|---|---|---|---|---|
| | Reading | Mathematics | Reading | Mathematics |
| **Age-11 test score** | | | | |
| Second child | -0.0386*** | -0.0174*** | -0.0105*** | 0.00968*** |
| | (0.000534) | (0.000528) | (0.00100) | (0.000980) |
| Girl | 0.0485*** | -0.0387*** | 0.0456*** | -0.0335*** |
| | (0.000748) | (0.000739) | (0.00144) | (0.00141) |
| Observations | 638,852 | 638,852 | 174,984 | 174,984 |
| Families | 319,426 | 319,426 | 87,492 | 87,492 |

| | White British | Minorities |
|---|---|---|
| **Age-16 total points score** | Total points score | Total points score |
| Second child | -0.0418*** | -0.0147*** |
| | (0.000462) | (0.000975) |
| Girl | 0.0755*** | 0.0779*** |
| | (0.000647) | (0.00141) |
| Observations | 676,452 | 160,262 |
| Families | 338,226 | 80,131 |

*Notes* :Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table shows the impact of birth order on age-11 and age-16 academic performance. Age-11 reading and mathematics test scores are the grade 6, 'Key Stage 2,' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are expressed as percentile ranks within each national cohort.

# Appendix Table 15 Adolescent Academic Outcomes: Time-Varying Controls

| | White British | | Minorities | |
|---|---|---|---|---|
| | Reading | Mathematics | Reading | Mathematics |
| **Age-11 test score** | | | | |
| Second child | -0.126*** | -0.0578*** | -0.0279*** | 0.0318*** |
| | (0.00185) | (0.00182) | (0.00345) | (0.00335) |
| Girl | 0.166*** | -0.106*** | 0.158*** | -0.0902*** |
| | (0.00256) | (0.00253) | (0.00492) | (0.00478) |
| Observations | 638,852 | 638,852 | 174,984 | 174,984 |
| Families | 319,426 | 319,426 | 87,492 | 87,492 |

| | White British | Minorities |
|---|---|---|
| **Age-16 total points score** | Total points score | Total points score |
| Second child | -0.148*** | -0.0459*** |
| | (0.00159) | (0.00343) |
| Girl | 0.252*** | 0.261*** |
| | (0.00222) | (0.00493) |
| Observations | 676,452 | 160,262 |
| Families | 338,226 | 80,131 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects, dummy variables for month of birth, as well as the following time-varying controls: free lunch status and the home location local poverty index are included, both measured at the time of Reception (age four or five) enrollment. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 16 Adolescent Academic Outcomes: Free Lunch Status

| | White British | | | | Minorities | | | |
|---|---|---|---|---|---|---|---|---|
| | Reading | | Mathematics | | Reading | | Mathematics | |
| | FL = 0 | FL = 1 | FL = 0 | FL = 1 | FL = 0 | FL = 1 | FL = 0 | FL = 1 |
| **Age-11 test score** | | | | | | | | |
| Second child | -0.131*** | -0.111*** | -0.0613*** | -0.0479*** | -0.0503*** | 0.00399 | 0.0139*** | 0.0579*** |
| | (0.00204) | (0.00397) | (0.00199) | (0.00401) | (0.00431) | (0.00560) | (0.00406) | (0.00562) |
| Girl | 0.171*** | 0.154*** | -0.108*** | -0.101*** | 0.157*** | 0.161*** | -0.102*** | -0.0722*** |
| | (0.00284) | (0.00559) | (0.00278) | (0.00565) | (0.00615) | (0.00811) | (0.00579) | (0.00814) |
| Observations | 470,814 | 168,038 | 470,814 | 168,038 | 102,828 | 72,156 | 102,828 | 72,156 |
| Families | 235,407 | 84,019 | 235,407 | 84,019 | 51,414 | 36,078 | 51,414 | 36,078 |

| | White British | | Minorities | |
|---|---|---|---|---|
| | FL = 0 | FL = 1 | FL = 0 | FL = 1 |
| **Age-16 total points score** | | | | |
| Second child | -0.158*** | -0.121*** | -0.0446*** | -0.0475*** |
| | (0.00184) | (0.00312) | (0.00459) | (0.00506) |
| Girl | 0.243*** | 0.275*** | 0.232*** | 0.294*** |
| | (0.00257) | (0.00440) | (0.00659) | (0.00741) |
| Observations | 491,988 | 184,482 | 85,600 | 74,666 |
| Families | 245,994 | 92,241 | 42,800 | 37,333 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table reports results separately for students eligible for Free Lunch ('Free School Meal') (FL = 1) versus those not eligible for Free Lunch (FL = 0). Both siblings are classified as belonging to a free lunch status family if the free lunch status variable is equal to one in any grade for either sibling. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 17 Adolescent Academic Outcomes: Gender Composition

| | Age-11 test score | | | | | |
|---|---|---|---|---|---|---|
| | Reading | | | Maths | | |
| | GG | BB | G/B | GG | BB | G/B |
| White British | | | | | | |
| Second child | -0.122*** | -0.138*** | -0.122*** | -0.0465*** | -0.0609*** | -0.0614*** |
| | (0.00363) | (0.00363) | (0.00260) | (0.00357) | (0.00356) | (0.00259) |
| Girl | | | 0.166*** | | | -0.106*** |
| | | | (0.00257) | | | (0.00256) |
| Observations | 153,722 | 166,834 | 318,296 | 153,722 | 166,834 | 318,296 |
| Families | 76,861 | 83,417 | 159,148 | 76,861 | 83,417 | 159,148 |
| Minorities | | | | | | |
| Second child | 0.00163 | -0.0497*** | -0.0311*** | 0.0564*** | 0.0214*** | 0.0253*** |
| | (0.00661) | (0.00669) | (0.00501) | (0.00645) | (0.00651) | (0.00484) |
| Girl | | | 0.158*** | | | -0.0900*** |
| | | | (0.00495) | | | (0.00478) |
| Observations | 44,992 | 46,856 | 83,136 | 44,992 | 46,856 | 83,136 |
| Families | 22,496 | 23,428 | 41,568 | 22,496 | 23,428 | 41,568 |

| | Age-16 total points score | | | | | |
|---|---|---|---|---|---|---|
| | White British | | | Minorities | | |
| | GG | BB | G/B | GG | BB | G/B |
| Second child | -0.159*** | -0.144*** | -0.145*** | -0.0424*** | -0.0393*** | -0.0526*** |
| | (0.00316) | (0.00306) | (0.00228) | (0.00641) | (0.00652) | (0.00506) |
| Girl | | | 0.252*** | | | 0.260*** |
| | | | (0.00225) | | | (0.00502) |
| Observations | 159,656 | 180,454 | 336,342 | 40,896 | 44,486 | 74,880 |
| Families | 79,828 | 90,227 | 168,171 | 20,448 | 22,243 | 37,440 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table reports results separately for samples of girl-girl ('GG'), boy-boy ('BB'), and mixed girl-boy ('G/B') pairs of siblings. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 18 Adolescent Academic Outcomes: Within-Group Standardization

| | White British | | Minorities | |
|---|---|---|---|---|
| | Reading | Mathematics | Reading | Mathematics |
| **Age-11 test score** | | | | |
| Second child | -0.126*** | -0.0393*** | -0.0454*** | -0.0109*** |
| | (0.00184) | (0.00182) | (0.00337) | (0.00330) |
| Girl | 0.168*** | -0.107*** | 0.155*** | -0.0892*** |
| | (0.00258) | (0.00254) | (0.00484) | (0.00473) |
| Observations | 638,700 | 638,700 | 174,964 | 174,964 |
| Families | 319,427 | 319,427 | 87,516 | 87,516 |

| | White British | Minorities |
|---|---|---|
| **Age-16 total points score** | Total points score | Total points score |
| Second child | -0.131*** | -0.0720*** |
| | (0.00162) | (0.00329) |
| Girl | 0.257*** | 0.252*** |
| | (0.00227) | (0.00477) |
| Observations | 676,452 | 160,262 |
| Families | 338,226 | 80,131 |

*Notes* : Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each cohort's national distribution, separately for white British and minority students.

# Appendix Table 19 Age-16 Academic Outcomes: Sibling Spacing Interaction

| | White British | Minorities |
|---|---|---|
| Second child | -0.105*** | -0.0362*** |
| | (0.00389) | (0.00700) |
| Second child | -0.0362*** | -0.00548 |
| x Grade spacing = 2 | (0.00453) | (0.00893) |
| Second child | -0.0651*** | -0.0164 |
| x Grade spacing = 3 | (0.00500) | (0.0102) |
| Second child | -0.0714*** | -0.0295** |
| x Grade spacing = 4 | (0.00623) | (0.0129) |
| Second child | -0.0671*** | -0.0368* |
| x Grade spacing = 5 | (0.00973) | (0.0200) |
| Observations | 676,452 | 160,262 |
| Families | 338,226 | 80,131 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. 'Grade spacing = k' is a dummy variable indicating that the two siblings are k grades apart from each other. All regressions control for family fixed effects and include dummy variables for month of birth. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.

# Appendix Table 20 School Suspensions: Sibling Spacing Interaction

| | White British | | Minorities | |
|---|---|---|---|---|
| | Susp., Reception to Grade 6 | Susp., Grade 7 to Grade 11 | Susp., Reception to Grade 6 | Susp., Grade 7 to Grade 11 |
| Second child | 0.00607*** | 0.0269*** | 0.000727 | 0.0235*** |
| | (0.000709) | (0.00174) | (0.00110) | (0.00311) |
| Second child | -0.00614*** | -0.0162*** | -0.000692 | -0.0127*** |
| x Grade spacing = 2 | (0.000820) | (0.00202) | (0.00135) | (0.00392) |
| Second child | -0.00604*** | -0.0217*** | -0.00279* | -0.0195*** |
| x Grade spacing = 3 | (0.000873) | (0.00214) | (0.00152) | (0.00444) |
| Second child | -0.00736*** | -0.0247*** | -0.00349* | -0.0274*** |
| x Grade spacing = 4 | (0.00108) | (0.00262) | (0.00192) | (0.00558) |
| Second child | -0.00599*** | -0.0305*** | -0.0109*** | -0.0438*** |
| x Grade spacing = 5 | (0.00167) | (0.00395) | (0.00290) | (0.00847) |
| Observations | 693,632 | 676,452 | 192,054 | 160,262 |
| Families | 346,816 | 338,226 | 96,027 | 80,131 |

Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table shows the impact of birth order on school suspensions. 'Ever suspended' is a binary dummy variable indicating whether the student was ever suspended in the respective grades. 'Grade spacing = k' is a dummy variable indicating that the two siblings are k grades apart from each other.

# Appendix Table 21 School Suspensions by Gender Composition

| | Ever suspended, Grades Reception to 6 | | | Ever suspended, Grades 7 to 11 | | |
|---|---|---|---|---|---|---|
| | GG | BB | G/B | GG | BB | G/B |
| **White British** | | | | | | |
| Second child | 0.000194 | 0.000460 | 0.00113*** | 0.00800*** | 0.0127*** | 0.00958*** |
| | (0.000296) | (0.000669) | (0.000390) | (0.00115) | (0.00144) | (0.000963) |
| Girl | | | -0.0203*** | | | -0.0682*** |
| | | | (0.000385) | | | (0.000952) |
| Observations | 164,004 | 183,916 | 345,712 | 159,656 | 180,454 | 336,342 |
| Families | 82,002 | 91,958 | 172,856 | 79,828 | 90,227 | 168,171 |
| **Minorities** | | | | | | |
| Second child | -0.000615 | -0.000364 | -0.00141* | 0.00613*** | 0.0132*** | 0.00961*** |
| | (0.000554) | (0.00122) | (0.000763) | (0.00235) | (0.00323) | (0.00227) |
| Girl | | | -0.0192*** | | | -0.0992*** |
| | | | (0.000754) | | | (0.00225) |
| Observations | 48,364 | 52,110 | 91,580 | 40,896 | 44,486 | 74,880 |
| Families | 24,182 | 26,055 | 45,790 | 20,448 | 22,243 | 37,440 |

Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table shows the impact of birth order on school suspensions. Table reports results separately for samples of girl-girl ('GG'), boy-boy ('BB'), and mixed girl-boy ('G/B') pairs of siblings.

# Appendix Table 22 Adolescent Academic Outcomes: Detailed Minority Status

| | Heritage | | | | | | |
|---|---|---|---|---|---|---|---|
| | White British | African | Afro-Caribbean | Bangladeshi | Indian | Pakistani | White other |
| **Age-11 Reading test score** | | | | | | | |
| Second child | -0.127*** | 0.0546*** | -0.00165 | 0.0514*** | -0.0534*** | -0.0288*** | -0.0636*** |
| | (0.00203) | (0.0111) | (0.0255) | (0.0179) | (0.0110) | (0.00971) | (0.0122) |
| Girl | 0.169*** | 0.161*** | 0.208*** | 0.0992*** | 0.151*** | 0.103*** | 0.173*** |
| | (0.00284) | (0.0156) | (0.0357) | (0.0294) | (0.0157) | (0.0143) | (0.0170) |
| Observations | 578,731 | 21,515 | 5,760 | 8,778 | 18,657 | 27,862 | 26,438 |
| Families | 319,426 | 13,164 | 3,819 | 5,442 | 10,868 | 16,399 | 19,900 |
| **Age-11 Mathematics test score** | | | | | | | |
| Second child | -0.0584*** | 0.0913*** | 0.0118 | 0.123*** | 0.0319*** | 0.0523*** | -0.00351 |
| | (0.00200) | (0.0108) | (0.0256) | (0.0179) | (0.00997) | (0.00977) | (0.0115) |
| Girl | -0.104*** | -0.0432*** | 0.00305 | -0.168*** | -0.0868*** | -0.150*** | -0.114*** |
| | (0.00280) | (0.0151) | (0.0359) | (0.0294) | (0.0143) | (0.0144) | (0.0160) |
| Observations | 578,731 | 21,515 | 5,760 | 8,778 | 18,657 | 27,862 | 26,438 |
| Families | 319,426 | 13,164 | 3,819 | 5,442 | 10,868 | 16,399 | 19,900 |
| **Age-16 total points score** | | | | | | | |
| Second child | -0.148*** | 0.0163 | -0.0653*** | -0.0384*** | -0.0141 | -0.0529*** | -0.0943*** |
| | (0.00164) | (0.0108) | (0.0179) | (0.0134) | (0.00986) | (0.00754) | (0.0106) |
| Girl | 0.251*** | 0.326*** | 0.323*** | 0.179*** | 0.189*** | 0.263*** | 0.255*** |
| | (0.00230) | (0.0151) | (0.0248) | (0.0232) | (0.0141) | (0.0114) | (0.0149) |
| Observations | 651,442 | 18,962 | 7,147 | 10,625 | 19,552 | 31,540 | 25,076 |
| Families | 338,226 | 10,473 | 4,289 | 5,635 | 10,546 | 16,714 | 16,803 |

Notes: Standard errors are clustered at the family level, shown in parentheses. *, **, *** denote significance at the 10%, 5% and 1% level, respectively. All regressions control for family fixed effects and include dummy variables for month of birth. Table reports results separately for students from different minority ethnic heritage groups. Age-11 reading and mathematics test scores are the grade-6 'Key Stage 2' assessments at the end of primary school. Age-16 exam points score is the total points score on all GCSE exams. Age-11 and age-16 scores are standardized within each national cohort to have mean zero and unit standard deviation.