THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 22/13

Right Peer, Right Now? Endogenous Peer Effects and Achievement in Victorian Primary Schools

Duncan McVicar, Julie Moschion and Chris Ryan

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Right Peer, Right Now? Endogenous Peer Effects and Achievement in Victorian Primary Schools*

Duncan McVicar, Julie Moschion and Chris Ryan
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 22/13

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-4312-2

June 2013

* This research was commissioned by the Victorian Department of Education and Early Childhood Development (DEECD). This paper uses unit record data from the NAPLAN administrative data collection, provided by DEECD. Thanks also to Mike Helal for help with the data and to seminar participants at the Melbourne Institute and at the Journées de Microéconomie Appliquée 2013 meeting for helpful comments on an earlier draft. The findings and views reported in this paper are those of the authors and should not be attributed to DEECD or any other branch of the Victorian or Australian Commonwealth government.
Corresponding author: Chris Ryan, <ryan.c@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper presents estimates of endogenous peer effects in pupils' school achievement using data on national test scores, across multiple subjects and cohorts, for the population of primary school pupils in Years 3 and 5 (aged 7/8 and 9/10 years) in the Australian state of Victoria. Identification is achieved via school-grade fixed effects and instrumental variables (IV), exploiting plausibly random differences in the age distribution of peers and their gender mix across cohorts. The results provide strong evidence for the existence of endogenous peer effects across all subjects, with the IV estimates close in magnitude to the corresponding fixed-effects estimates, although less precisely estimated. In reading, for example, a one point increase in peers' average test scores leads to between a .14 and .39 point increase in own test score, with similar ranges across other subjects.

JEL classification: I21, I24, J24

Keywords: Endogenous peer effects, school achievement, education, Australia

1. Introduction

Peer effects refer to externalities in which the actions or characteristics of a reference group affect an individual's behaviour or outcomes.[1] Such effects are possible across a wide range of behaviours and social contexts. Ever since the seminal Coleman (1966) report on US schools, however, the effects of peers on individuals' educational outcomes have attracted particular attention in the literature. Sacerdote (2011) provides a recent review.

Peer effects in educational achievement might work through a number of mechanisms, including direct learning between peers, competition between peers, classroom disruption, and the influence of peers on the pace and level at which the teacher can teach the class. Since Manksi (1993), peer effects that operate via current peer behaviour or outcomes (e.g. individuals study harder and/or perform better in tests if their peers study harder and/or perform better in tests) have become known as *endogenous peer effects*, whereas those that result directly from peer background characteristics have become known as *exogenous peer effects*, or *contextual effects*.

Both kinds of peer effects, if they are of sufficient magnitude, have critical implications for parents and policy makers. For example, parents may be able to improve their child's expected educational outcomes by selecting a school with a more advantaged intake (an *exogenous peer effect*). This, in turn, may lead to widening educational inequalities across the school system. Further, *endogenous peer effects* imply social multipliers whereby interventions that improve the educational achievement of some pupils in a group will have positive spill-over effects on the educational achievement of other pupils in the group.[2]

Establishing the existence and magnitude of peer effects, however, is beset by practical difficulties (see Manski, 1993, 2000; Moffitt, 2001). One such difficulty is separately identifying peer effects from correlation between peers' outcomes driven by shared

[1] A more general term for such externalities is 'social interactions effects', but 'peer effects' is the term more commonly used in the education literature where individuals are generally studied in the context of similarly-aged reference groups with whom they regularly interact, e.g. the school class.

[2] If peer effects are non-linear in nature, e.g. where students in different parts of the ability distribution are impacted in different ways by high-ability peers and low-ability peers, then there are additional policy implications, including the possibility that tracking within schools can lead to overall improvements in educational performance in addition to distributional effects. Partly for this reason, non-linear peer effects have been attracting growing attention in the academic literature in recent years (e.g. see Hoxby and Weingarth (2005) and Sacerdote (2011) for a review).

unobservable characteristics due to common shocks (e.g. a year with a great teacher) or non-random sorting into and within schools (e.g. if the most motivated families send their children to a particular school, or even a particular class within a particular school). Another difficulty – one of the implications of what has become known as Manski's *reflection problem* – is separately identifying the *endogenous peer effect* from the impacts of other peer characteristics, even where such characteristics are observed.

Much of the economics literature on peer effects in educational performance has concentrated on exploiting random group assignment or some other plausibly exogenous source of variation in peer characteristics to tackle the first problem, while not attempting to separately identify peer effects due to peer achievement from those due to other peer characteristics (e.g. race, gender). In other words, many studies have examined the evidence for peer effects *writ large* – simply whether peers impact on individuals – without trying to separate endogenous from exogenous peer effects (e.g. Hoxby, 2000; Angrist and Lang, 2004; Sanbonmatsu et al. 2006; Kling et al., 2007; Lavy and Schlosser, 2011; Ammermueller and Pischke, 2009; Duflo et al., 2011; Black et al., 2013). The evidence from this literature – at least where it looks at linear-in-means models of peer effects[3] – is somewhat mixed, but on balance suggests that peers do have a smallish impact on educational outcomes (Sacerdote, 2011).

Some studies have taken the linear-in-means model further, exploiting additional data or making additional identifying assumptions that arguably help to isolate the endogenous peer effect in educational outcomes from the impacts of other peer characteristics (e.g. Hanushek et al., 2003; Lefgren, 2004; Goux and Maurin, 2007; Atkinson et al., 2008; Boucher et al., 2012; Lin, 2010; Gibbons and Telhaj, 2012; Lavy et al., 2012). Although results vary, on balance these studies also generally point to smallish (positive) impacts from peer achievement on own achievement.[4]

[3] The linear-in-means peer effects model assumes that an individual's outcomes are linearly related to the average characteristics and/or outcomes of his/her peers. It is the workhorse model of peer effects in the literature (see Sacerdote, 2011).

[4] The distinction between endogenous and exogenous peer effects can be a little fuzzy in this literature, particularly where peer outcomes are lagged. For example, lagged peer test scores can be interpreted as proxying for peers' prior ability (an exogenous peer effect) as well as their behaviour (an endogenous peer effect). Parallel strands in the literature seek to identify endogenous peer effects in other behaviours, e.g. in adolescent substance use, where the distinction is clearer (see Gaviria and Raphael, 2001; Powell et al., 2005; Lundborg, 2006; Clark and Lohéac, 2007; Fletcher, 2010; McVicar, 2011).

This paper estimates endogenous peer effects in school achievement using administrative data on the test scores of primary school pupils at public schools in the Australian state of Victoria, and so belongs in the second group of studies. Our data come from the National Assessment Program – Literacy and Numeracy (NAPLAN) conducted across Australia, which provides test scores for five achievement domains – numeracy, reading, spelling, grammar, and writing – for all pupils in grades 3, 5, 7, and 9. The national testing system which generates these data was introduced across Australia in 2008, with tests taking place each year in the specific grades, so we are able to exploit data across four cohorts. We present linear-in-means estimates of the endogenous peer effect in pupils' school achievement for four of the five subject areas.[5] Identification is achieved via school-grade fixed effects and instrumental variables (IV), exploiting plausibly random differences in the age distribution of peers and their gender mix across cohorts, under the assumption that peer age and peer gender have no direct contextual effects on own achievement. Specifically, we use school-grade fixed effects to address selection into schools (and other grade-level unobservables), and IV to address the reflection problem.

Our approach builds most closely on an approach of Goux and Maurin (2007), specifically on their age-based IV estimates of endogenous peer effects. But where Goux and Maurin (2007) use survey data on peer test scores for 7,500 grade 3 pupils in a single cohort across a sample of French primary schools,[6] we use administrative data for the population of pupils across all public primary schools in the state of Victoria, across two grades and four consecutive cohorts, drawing on around 160,000 observations for each grade. Because these data allow us to focus on within-school grade differences across cohorts, and because they cover enough children to generate reasonably precise IV estimates at least in some cases, our chances of correctly identifying the endogenous peer effect using this IV approach are improved.

We focus on Years 3 and 5 – the two primary school grades in Victoria covered by the NAPLAN tests – where pupils are aged 7/8 years and 9/10 years old. Our focus on primary school children reflects a number of factors. First, a substantially higher proportion of the population attend public schools at primary school level in Victoria than at the secondary or

[5] The estimates for writing are omitted because the nature of the writing task was changed substantially after the second year of tests, so the scaled scores are not strictly comparable over time. The results from analysing this domain are available from the authors on request.

[6] These are not the main estimates presented by Goux and Maurin (2007), but they are the closest in their approach to those presented in the current paper.

high school level.[7] Second, grade cohorts arguably make more sense as a reference group in primary schools than in secondary schools because primary schools are smaller and children in a given grade are likely to be in closer contact with their peers in the grade than would be the case in a larger secondary school.[8] Third, there is almost no streaming by ability in primary schools in Victoria – endogenous sorting *within* school grades – whereas there is some streaming by ability at secondary school level.[9] Fourth, the *ex ante* case for our instrument based on the date of birth of peers is stronger for primary school pupils than for secondary school pupils, by which time age-at-starting-school effects may have dissipated (e.g. see Elder and Lubotsky, 2009; Crawford et al., 2010).

The remainder of this paper is set out as follows. Section 2 briefly reviews the existing literature. Section 3 sets out our empirical approach, briefly describing the data and discussing identification. Section 4 presents and discusses the main estimates. Section 5 presents and discusses further results: extending estimates to secondary school grades, examining evidence for heterogeneous peer effects by gender, socio-economic status (SES), and school size, and examining evidence for non-linear peer effects. Section 6 concludes.

2. Literature

The workhorse model of peer effects in school achievement is the linear-in-means model (1), where individual outcomes A_{igst}, for individual *i*, in group *g*, in school *s*, at time *t* – usually but not always measured by test scores – are assumed to be a function of individual observable characteristics X_{igst}, peer mean observable characteristics $\overline{Z}_{-igst}$ (excluding the individual, and possibly including factors that are not in the vector X), mean peer achievement $\overline{A}_{-igst}$ (again measured excluding the individual), and an error term ε_{igst} that captures unobserved influences.

$$A_{igst} = \theta_1 + \theta_2 \overline{A}_{-igst} + X'_{igst}\theta_3 + \overline{Z}'_{-igst}\theta_4 + \varepsilon_{igst} \quad (1)$$

[7] The public school share of students in Victoria was 67% in primary schools in 2011, compared with 58% in secondary schools.
[8] The average secondary school in Victoria had more than three times the number of students in 2011 than the average primary school.
[9] For example, the vast majority (>90%) of Australian primary schools in 2003 did not stream 4th grade students in mathematics, while among 8th graders, 58% of student respondents were streamed (Thomson et al., 2003).

The parameter θ_2 denotes the impact of peer achievement on own achievement, i.e. the endogenous peer effect. The parameter vector θ_4 denotes the impact of observable peer characteristics on own achievement, i.e. the contextual effects. If peer groups have been together for more than the current year, then such parameters can be thought of as capturing the cumulative effects of peers throughout their time together (e.g. Black et al., 2013).

The two identification problems mentioned in the previous section (correlated effects and the reflection problem) mean that the error term ε_{igst} is likely to be correlated with peer achievement (endogeneity, likely leading to upwards-biased estimates of the parameter θ_2),[10] and that the parameters in θ_2 and θ_4 cannot be separately identified, at least at the 'whole group' level, without making additional, and generally very strong, identifying assumptions.

Peer effects *writ large* can be estimated from a reduced-form version of (1) as follows:

$$A_{igst} = \phi_1 + X'_{igst}\phi_2 + \bar{Z}'_{-igst}\phi_3 + e_{igst} \quad (2)$$

There are various different versions of (2) in the literature (for a review see Sacerdote, 2011). Often there is one particular peer characteristic or peer variable – gender balance in the case of Lavy and Schlosser (2011), being part of the Metco scheme in the case of Angrist and Lang (2004), number of books at home in the case of Ammermueller and Pischke (2009) – that is the focus, with the critical issue for identification being the extent to which this particular variable (and other peer characteristics if included) can be treated as exogenous. But all share the interpretation that if the reduced-form parameter (or parameter vector) ϕ_3 is non-zero, this implies that at least one of the underlying structural parameters in θ_2 and θ_4 is non-zero, i.e. that peers impact on own educational achievement.[11]

[10] A further source of endogeneity and therefore likely bias arises because, with non-zero peer effects, an individual's unobserved influences on own performance will be correlated with peer performance through the individual's influence on peer outcomes. This is the second implication of Manski's reflection problem.

[11] Hoxby (2000, p8) sets this out explicitly in the case of peer gender balance: *"If students are influenced by their peers' achievement, then the cohort's gender composition would affect males' achievement. Second, the prevalence of females could have some effect on achievement that does not operate though its effect on peer achievement. Females might, for instance, have a general effect on classroom culture."*

A variant of this approach is a reduced form where only peer *achievement*, usually lagged, is included on the right hand side of (2) in place of peer characteristics, with own achievement sometimes expressed in value-added terms (e.g. Lefgren, 2004; Atkinson et al., 2008). OLS estimates from these models are also generally interpreted as providing evidence of peer effects writ large.[12] But under strong assumptions – in particular if one is prepared to assume that there are no contextual effects, i.e. that only peer *outcomes* influence own outcomes – the parameter on peer achievement can be interpreted as the endogenous peer effect (although strictly speaking this would require contemporaneous peer achievement on the right hand side rather than lagged peer achievement). Unfortunately, the 'no contextual effects' assumption gets at best mixed support in the educational peer effects literature (see e.g. Weingarth and Hoxby, 2005; Lin, 2010; Boucher et al., 2012), although it is more commonly assumed in the parallel substance-use peer effects literature (e.g. Gaviria and Raphael, 2001; Powell et al., 2005).

Several papers estimate versions of (1), including both lagged peer achievement and peer characteristics (to control for contextual effects) on the right hand side (e.g. Hanushek et al. 03; Vigdor and Nechyba, 2007; Clark and Lohéac, 2007 (in the substance use literature)). This helps to sidestep the reflection problem because prior peer outcomes cannot be determined by current own outcomes. On the other hand, peer outcomes are likely to be correlated over time so that lagged peer scores proxy for current peer scores and are therefore still subject to the reflection problem (Manski, 1993).[13] Lagged peer outcomes will also not capture subsequent shocks to peer outcomes which might impact on own outcomes (Hanushek et al., 2003; Lavy et al., 2012). Hanushek et al. (2003) therefore suggest that the resulting estimates should be interpreted as lower bounds on the endogenous peer effect.

An alternative approach to identifying the endogenous peer effect is to exploit plausibly exogenous influences on current peer outcomes in an IV set up (see e.g. Boozer and Cacciola, 2001; Goux and Maurin, 2007; Fletcher, 2010 (in the substance use literature)).[14] Of these

[12] Hoxby (2000, p6) states that *"...the baseline [linear-in-means with no contextual effects] model does not assert that there is a single channel for peer effects: it asserts that mean peer achievement is a sufficient statistic for the multiple channels."*

[13] Two recent papers (Gibbons and Telhaj, 2012; Lavy et al., 2012) solve this problem by exploiting the transition between primary and secondary school and the fact that most secondary school peers attended different primary schools, although neither paper claims explicitly to be identifying the endogenous peer effect.

[14] There are also studies that instrument for lagged peer outcomes (e.g. Lefgren, 2004; Atkinson et al., 2008; and one of the models of Goux and Maurin, 2007). The resulting estimates may also be interpretable as capturing

studies, the closest to the current paper is the model of Goux and Maurin (2007) where they instrument for peer test scores in grade 3 using information on the birth dates of peers within the school class taken from a survey of 7,500 French pupils. They find that a one standard deviation increase in the average test score of peers increases a pupil's score by .36 of a standard deviation. We estimate a similar IV model here, but using data that substantially boost the chances of this IV approach correctly identifying the endogenous peer effect. Where our own IV estimates are most comparable to those of Goux and Maurin (2007) – for Year 3 test scores, based on the peer-age instrument – they generally suggest peer effects that are smaller in magnitude.

A new strand of the literature exploits network data on friendship links within schools to more explicitly identify the endogenous peer effect (e.g. Lin, 2010; Boucher et al., 2012). This approach uses the property that nominated friendship groups do not perfectly overlap to sidestep the reflection problem. Such data remain rare, however, and despite the advantages of this approach, it only allows identification of the endogenous peer effect between friends rather than between classmates or grade-mates.

Finally, to the best of our knowledge, ours is the first such study of peer effects for school achievement using Australian data. The closest Australian studies are probably Foster and Frijters (2009) and Foster and Frijters (2010), both of which use survey data to analyse how *beliefs* about the existence of peer effects and peers' inputs affect undergraduate students' effort levels.

3. Data and Identification

NAPLAN scores for students in grades 3, 5, 7 and 9 are measured on a scale from 0 to 1000, and are designed to measure absolute (rather than relative) competence in the subject concerned.[15] In addition to test score data, we can identify each pupil's school and grade (but

endogenous peer effects, although following the argument of Hanushek et al. (2003) and Lavy et al. (2012) they are likely to provide lower-bound estimates of the underlying structural parameter in (1).

[15] Average NAPLAN scores in a particular subject area are therefore higher in higher school grades. Note that a small number of pupils in each case (always less than 5% at grades 3 and 5) are not entered for, or otherwise miss, NAPLAN tests. This leads to slight variation between subjects in sample composition and sample size in a way that could plausibly be correlated with our IVs. A simple sensitivity check – introducing controls for the proportion of boys and the proportion of 'old' that sit the test in the school at date t – suggests that this does not affect our estimates.

not class), which we use to construct peer test scores and other peer measures. We also have information on gender, date of birth, an indicator of Indigenous status, an indicator for whether English is the main language spoken at home, and information on mothers' and fathers' occupations and education levels. These are included in our models as controls at the individual level. Table 1 gives summary statistics. Peer averages of these observed characteristics are also included as contextual effects (with the exception of proportion male and peer dates of birth) in the IV models.

Our starting point for estimation is an extended version of (1) to allow for school and time fixed effects, denoted by η_s and τ_t respectively:

$$A_{igst} = \theta_1 + \theta_2 \overline{A}_{-igst} + X'_{igst}\theta_3 + \overline{Z}'_{-igst}\theta_4 + \eta_s + \tau_t + \varepsilon_{igst} \quad (3)$$

The coefficient of interest is θ_2 – the *endogenous peer effect* – which captures the association between peer test scores and own test scores. Equation (3) is treated separately by grade (so the school fixed effects are interpretable as school-grade fixed effects) and separately by subject or learning domain.

The extent to which we can interpret θ_2 as capturing a causal effect of peer educational achievement on own achievement depends in part on the extent to which the controls in (3) account for common unobserved influences on own and peer scores (what Manski (1993) calls *correlated effects*). The main concern here is that there may be unobservable influences at the school level, e.g. because of endogenous sorting into schools or differences in teaching quality across schools, which if not accounted for will lead to correlation between own and peer scores even in the absence of peer effects. By including school fixed effects in (3), however, we control for any such school-level unobservables, at least to the extent that they are time-invariant.[16] Identification of the peer effect in (3) therefore relies on cohort-level differences in test scores from the school-grade means. In relying on such cross-cohort variation to help identify peer effects we follow a number of earlier studies, all of which

[16] Use of the school fixed effects strategy in conjunction with grade-level peer groupings allows us to deal effectively with sorting into schools. Our approach could not deal with sorting within schools if the peer grouping was specified more finely. So, while a closer identification of an individual's peers might be attractive (e.g. to examine non-linear peer effects within classrooms), the within school sorting implied would induce additional selection biases.

make the claim that such variation can be treated as exogenous (e.g. Hoxby, 2000; Ammermueller and Pischke, 2009; Lavy and Schlosser, 2011).

Even with school fixed effects, however, identification issues remain in (3) because I am a peer of my peers (Manski's *reflection problem*). First, if there are non-zero peer effects, then my own educational achievement will affect my peers' educational achievement, and unobserved determinants of my own achievement may also be correlated with peer test scores through this channel. Intuitively, however, the empirical importance of the resulting bias will depend on the size of the reference group (and of course the strength of the peer effect). In this case, because we treat the whole school grade as the reference group (with an average size of 58 pupils), it seems reasonable to assume that any simultaneity bias is small and that causality runs almost entirely from peers to the individual.[17] Second, and more critically, if peer characteristics directly influence own educational achievement, i.e. if there are non-zero contextual effects, then Manski (1993) and Moffitt (2001) show that in a linear model such as (3) we cannot separately identify the impact of peer achievement on own educational achievement from that of peer characteristics on own educational achievement.

Following Goux and Maurin (2007), however, if we can find some exogenous driver of variation in peer achievement that itself has no direct impact on own achievement, e.g. an exogenous peer characteristic that has no contextual effect, then we can use an IV approach to estimate θ_2 in (3). In principle, such an IV approach solves both identification issues stemming from the reflection problem as well as purging the estimated peer effects of time-varying correlated effects that may remain after the within-school transformation (e.g. the effect of a particularly good teacher). Further, the assumption of no contextual effects from one particular peer characteristic is much weaker than the blanket assumption of no contextual effects, period.

Here we use two instruments for peer test scores, one based on the gender balance of peers, and the other based on peers' age (the latter following Goux and Maurin (2007)).[18] Specifically, we use the share of boys in the grade (minus the individual) and the share of the

[17] Bramoullé et al. (2009) make a similar point.

[18] Both are, of course, omitted from $\overline{Z}_{ist}$ in (3). Roughly two percent of pupils at the primary school level are in single-sex schools (mostly girls). We include an additional control for being in a single-sex school.

grade (minus the individual) born in the first trimester of the school-eligibility year.[19] We provide IV estimates using each of the two instruments separately, but our preferred IV estimates use both instruments together.[20] The first-stage regression is therefore given by (4):

$$\overline{A}_{-igst} = \lambda_1 + X'_{igst}\lambda_2 + \overline{Z}'_{-igst}\lambda_3 + \gamma_1 P^{gender}_{-igst} + \gamma_2 P^{age}_{-igst} + \psi_s + \omega_t + u_{-igst}, \quad (4)$$

where P^{gender}_{-igst} and P^{age}_{-igst} denote the proportion of *i*'s reference group that are male and that are born in the first trimester, respectively. We estimate (3) and (4) by two stage least squares (2SLS).

In practice, of course, the extent to which such an approach will work depends on the extent to which these two instruments are validly excludable from (3) and sufficiently correlated with peer test scores in (4). First, consider our instrument based on peers' gender. There are well-established gender differences in educational achievement across countries (e.g. see Hoxby, 2000; Jacob, 2002; Charles and Luoh, 2003; Lavy and Schlosser, 2011), and Victorian primary school children are no exception. Table 2 (and Tables A1-A3 in the Appendix) show that girls have significantly higher average test scores for reading, grammar and spelling than boys, and significantly lower average test scores for numeracy, in both Years 3 and 5. We therefore expect a statistically significant γ_1 in (4).

But can the gender balance of one's peers be validly excluded from (3)? There are two main ways in which this might not be the case: endogenous sorting into schools that is correlated with gender, and direct effects from peer gender on own achievement. Gender-correlated sorting into schools is certainly plausible, but is credibly controlled for in (3) by the inclusion of school fixed effects. Supporting evidence for this is given by the lack of significant correlations between peer gender balance and observable individual characteristics in Table 3

[19] Children can enter primary school in Victoria in late January (the start of the school year) if they turn five years old in the year up to 30th of April. The oldest entrants to the school year are therefore born in May or shortly after, with the youngest born just before or in April. In practice splitting the school-eligibility year into trimesters means separately identifying those pupils born from May to July, August to October, November to January and February to April. Our date-of-birth instrument is therefore the proportion of the school grade born in between May and July. Results are reasonably robust to alternative cuts of the date-of-birth data, e.g. by semester, although there is some variation in precise magnitudes.

[20] The estimates that use both instruments are preferred because they are more precise and the approach allows us to assess the validity of the instruments via over-identification tests.

(and Tables A4-A6 in the Appendix).[21] Own gender also appears to be largely orthogonal to other observed individual characteristics (see Table 2 and Tables A1-A3) and peer characteristics (see Table 4 and Tables A7-A9).[22] We can also draw on earlier studies to support the claim that peer gender balance within-school, in particular, can be treated as plausibly exogenous. For example, Hoxby (2000, p6) states the following: *"adjacent cohorts in a grade in a particular public school are a potential source of non-suspect variation [compared to differences across schools and across classes within schools]...Even within a school that has an entirely stable population of families, biological variation in the...timing and gender of births would create idiosyncratic variation in the share of 6 year olds, say, who were female...and so on."* She goes on to state (Hoxby, 2000 p7): *"There is little reason to suspect that variation between cohorts in gender composition, within a grade within a public school, is correlated with unobserved determinants of achievement."*

It is more difficult, however, to rule out a direct influence from the gender balance of peers on an individuals' educational performance, at least *ex ante*. For example, boys may behave differently in a reference group dominated by girls than they would in a reference group dominated by boys, regardless of the educational achievement of peers. Our conjecture is that any such direct effects are swamped by the indirect effect working through peer test scores, at least at the primary school level. In support of this conjecture, as we will see in the following section, over-identification tests fail to reject the excludability of our instruments in any of the domains. Even if our conjecture does not hold, having data across several domains allows us to plausibly estimate a lower bound on the endogenous peer effect in at least one domain. Specifically, because we would expect any contextual effect from the proportion of boys in the grade to take a *negative* sign (Hoxby, 2000; Lavy and Schlosser, 2011), and because the proportion of boys in the grade is *positively* correlated with peer average *numeracy* scores in

[21] Table 3 and Tables A4-A6 give regression coefficients and associated standard errors for each of the instruments in regressions of each individual observed characteristic on the instrument and school-grade fixed effects. Widespread significant correlations with observables may signal similar correlations with unobservables, i.e. that the instrument is not validly excludable from (3). In reading, for example, the only correlation that is significant at the 95% level is with year. This is picking up overall demographic differences in gender balance between cohorts, which are controlled for in (3) by the cohort dummies. There are three further correlations that appear marginally statistically significant – with one of the mother's occupation dummies, one of the mother's education dummies, and one of the father's education dummies – but all are very small in magnitude, and we would expect a smattering of marginally significant correlations purely at random. For numeracy, again we pick up cohort differences (controlled for in (3)) and two small-in-magnitude correlations (with one of the dummies for mother's occupation and one of the dummies for father's education).

[22] Where there are statistically significant correlations they are all small in magnitude. Lavy and Schlosser (2011) similarly demonstrate that the within-school variation in the proportion of female students is not related to within-school variations in student background characteristics.

(4), then the implied omitted variables bias on the peer test scores coefficient in (3) would take a *negative* sign. IV estimates based on the gender-balance instrument would therefore be interpretable as a lower bound on the endogenous peer effect in numeracy.[23]

Second, consider our instrument based on peers' dates of birth. Here too there are well-established differences in educational achievement by age within school years, with older children tending to score more highly in school tests than younger children (e.g. see Bedard and Dhuey, 2006; Elder and Lubotsky, 2009; Crawford et al., 2010). Elder and Lubotsky (2009) argue that the strength of this effect diminishes over time since school entry, and our focus on primary school children is in part motivated by this. Tables 2 and A1-A3 show that older children within the school grade have significantly higher average test scores than younger children for all four subjects in both Years 3 and 5. We therefore expect a statistically significant γ_2 in (4).

What of excludability? Again we are not concerned with age-correlated sorting into schools because we are able to specify school fixed effects.[24] As for the gender-based instrument, supporting evidence for this is given by the lack of significant correlations between peer age and observable individual characteristics in Tables 3 and A4-A6.[25] Note, however, that own date of birth appears to be less orthogonal to other observed individual characteristics than is the case for own gender (see Tables 2 and A1-A3), although again where there are statistically significant correlations they are small in magnitude.[26] There are also a handful of statistically significant correlations between own date of birth and peers' observed characteristics, although again all are small in magnitude.

[23] Lavy and Schlosser (2011) argue that it is unlikely that all reduced-form gains in own achievement due to peer gender balance are generated solely via peer achievement, in part because they find *positive* reduced-form effects of the proportion of girls on own achievement even in mathematics where girls have *lower* achievement than boys. (If there were no contextual effects from gender balance then one would expect a negative or zero reduced-form coefficient on their proportion of girls variable, which is what we find when we estimate the reduced form.)

[24] Note that the above Hoxby (2000) quote also argues that variations in the timing of births across cohorts can be treated as exogenous.

[25] For example, for both reading and numeracy, only six of the 31 individual observable controls are correlated with peers' date of birth at the 90% level or above – we would expect three at random – and again all of these correlations are very small in magnitude.

[26] One possible explanation is that some parents try to time conception so as to influence the age at which their children start school. The negative correlation between higher levels of parental education and date of birth, for example, could signal that higher-SES parents prefer their children to be at the younger end of the age distribution within the school year. Black et al. (2011) provide evidence that children who start school younger tend to perform slightly better in subsequent testing, given age.

Can we also rule out a direct influence from the age of peers on an individuals' educational achievement *ex ante*? Again, it seems plausible that younger children may behave differently in a grade dominated by older children than they might otherwise behave (and *vice versa*), regardless of the educational achievement of peers, and that such differences in behaviour could impact on own educational achievement. But again it seems unlikely that any such effects would be large, certainly relative to the indirect effect of peer age through peer test scores. We can again take some encouragement from the fact that over-identification tests fail to reject excludability in any domain. In this case we can also draw on Goux and Maurin (2007) to support the excludability of the age-based instrument, with the arguments presented there strengthened further by the use of school fixed effects.

If we are prepared to make the stronger assumption of no contextual effects *at all* – i.e. that $\theta_4 = 0$ in (3) – then we can also estimate the endogenous peer effect directly without the need for instruments, from a simplified version of (3) as follows:

$$A_{igst} = \theta_1 + \theta_2 \overline{A}_{-igst} + X'_{igst}\theta_3 + \eta_s + \tau_t + \varepsilon_{igst} \quad (5)$$

The assumption of no contextual effects is not uncommon in the literature, at least in the literature on peer effects in adolescent substance use (e.g. Gaviria and Raphael, 2001; Powell et al., 2005; Lundborg, 2006; McVicar, 2011). Gaviria and Raphael (2001) argue that school pupils are not generally exposed to the family background of their school peers, or at least are less exposed to the family background of school peers than they would be to the family background of close neighbourhood peers. Even if this is true, however, school pupils are likely to be exposed to the behavioural *consequences* of their peers' family backgrounds at school, some of which may have direct impacts on own achievement other than through peer achievement. Further, sorting into schools is likely to engender correlation between peer characteristics and own outcomes, although our fixed-effects specification arguably differences away this potential channel for contextual effects (e.g. see Fletcher, 2010).

Ultimately, whether this assumption is reasonable in this particular context is an empirical question. Fortunately, we can test it directly, if somewhat informally, by examining the joint significance of all the observable contextual effects in the IV models, i.e. testing whether λ_3

=0 in (3).[27] We do this for each of the four subjects, for both Years 3 and 5, using our preferred IV models (i.e. with both instruments included). In all eight cases, the contextual effects are jointly insignificant at the 95% level.[28] Of course, even in the absence of contextual effects, the fixed-effects estimate of θ_2 from (5) is still likely to be subject to simultaneity bias and time-varying correlated effects.

Nevertheless, our motivation for providing the fixed-effects estimates in addition to the IV estimates is threefold. First, as is often the case, the IV estimates are sometimes imprecise; the fixed-effects estimates provide a useful point of reference for these IV estimates. Second, the instruments (particularly the age-based instrument) are only weakly correlated with peer achievement at the secondary school level, so we are unable to learn much about peer effects for older children using this IV approach; the fixed-effects estimates allow us to exploit the full range of NAPLAN data to examine peer effects at the secondary school level. Third, examining the evidence for heterogeneous and non-linear peer effects – extensions we explore in Section 5 – is more straightforward using the fixed-effects approach than the IV approach (where the first stage can become unwieldy).

4. Main Results and Discussion

First consider peer effects in reading. Table 5 gives the key estimates, by grade, first for (5) estimated as a single-equation fixed-effects model, then for (3) and (4) estimated by 2SLS. In the latter case we present estimates for each of the instruments applied separately in addition to our preferred estimates where both instruments are used together. The fixed-effects estimates are both positive, of plausible magnitude, and highly statistically significant. The estimated magnitudes are very similar for the two grade levels, with a one point increase in

[27] Of course there could be other unmeasured peer characteristics that impact on own outcomes directly, but we have detailed peer-level controls here, and if we find zero contextual effects from our observed peer variables then this provides a good indication that there are unlikely to be significant contextual effects, period.

[28] The p-values are as follows: reading level 3 (level 5) = .09 (.73); numeracy level 3 (5) = .86 (.76); grammar level 3 (5) = .41 (.17); spelling level 3 (5) = .10 (.20). In the marginal case of reading level 3, where the contextual effects appear jointly significant at the 90% level but not the 95% level, a closer look at the estimates suggests the result is driven by a statistically significant coefficient on one of the peer parental education dummies (mothers' Year 12 or above). The coefficient is very small in magnitude, however, and takes a counterintuitive sign, suggesting that a 10% increase in the proportion of peers whose mothers have at least Year 12 leads to a .01 of a standard deviation fall in own test score. Similarly, in the marginal case of spelling level 3, the marginally significant result appears to be driven by a statistically significant coefficient on the proportion of peers for whom English is not the main language spoken at home, which is of a similar magnitude (although it takes the expected sign).

the average test scores of peers associated with between a .221 and .250 point increase in own test score. In normalised terms, a one standard deviation increase in peers' average test scores is associated with a .1 standard deviation increase in own test score at Year 3 and a .09 standard deviation increase in own test score at Year 5. These estimates are somewhat smaller in magnitude than the equivalent OLS estimates of Goux and Maurin (2007).

Now consider the IV estimates. Our expectation is that, if anything, the IV estimates are likely to be slightly lower than the fixed-effects estimates, because any remaining biases in the fixed-effects estimates are most likely positive.[29] The first stages for the gender-based instrument and for the age-based instrument are both encouraging. F-stats are above the usual rule of thumb level of 10 in each case (and well above 10 for the gender instrument). The instruments also take the correct signs (and their coefficients are large in magnitude) in both cases, so that the average peer test score is positively related to the proportion of older children in the peer group, and negatively related to the proportion of boys in the peer group. It follows that the first stage is also encouraging for the two instruments together, with F-stats above 20 and correct signs. Note that the two instruments appear orthogonal to one another – the first-stage coefficients do not change when the other instrument is included. Also note the encouraging (or at least not discouraging) over-identification test results in each case.

For each grade, the IV estimates are very similar whether we use the gender instrument, the age-based instrument, or both together. For Year 3, all three IV estimates are positive and statistically significant, despite the loss of precision. The magnitude of the IV estimate of the peer effect ranges from .335 to .392 (.14 to .16 in normalised terms), i.e. close to but *larger* than the corresponding fixed-effects estimate. For Year 5, all three IV estimates are positive, but given their imprecision, are no longer statistically significant. The point estimates are again close to, but in this case *smaller*, than the corresponding fixed-effects estimate, ranging from .143 to .203 (.06 to .08 in normalised terms). In other words, the IV estimates fall in a narrow range either side of the fixed-effects estimates.[30] There is certainly nothing in the IV estimates that leads us to question the earlier conclusion, based on the fixed-effects estimates, of positive peer effects.

[29] We cannot rule out IV estimates that are higher than the equivalent OLS estimates, however, in part because NAPLAN can be thought of as measuring peer achievement with error, which may lead to attenuation bias in the fixed-effects estimates.
[30] The equivalent IV estimates of Goux and Maurin (2007) are larger than their corresponding OLS estimates.

Our estimated peer effects in grammar and spelling follow similar patterns to those for reading (see Tables 6 and 7). In all cases the fixed-effects estimates suggest positive and highly statistically significant endogenous peer effects, with broadly similar estimated magnitudes to those for reading. The IV first stages are also passable in each case, with F-stats everywhere above 10, and the instruments always taking the correct signs. Over-identification tests also continue to fail to reject that our instruments may be excludable. Our preferred IV estimates – using both instruments together – suggest peer effects that are in all cases close to, but generally slightly smaller than, the corresponding fixed-effects estimates.

Now consider peer effects in numeracy. Table 8 gives the key estimates. Both fixed-effects estimates are positive, large in magnitude, and highly statistically significant, with a one-point increase in the average test scores of peers associated with between a .408 and a .454 point increase in own test score. In normalised terms, a one standard deviation increase in peers' average test scores is associated with a .19 standard deviation increase in own test score in Year 3 and a .18 standard deviation increase in own test score in Year 5.

Unfortunately our instruments perform less well in the case of numeracy than in the case of reading. For Year 3, both instruments appear weak, with F-stats well below 10. For Year 5, the first stages are more encouraging, with F-stats above 10 in each case, and with the instruments taking the correct signs, although the IV estimates are again imprecise. Also notice that the two instruments appear orthogonal to one another, as for reading, with similarly encouraging over-identification test results. The IV estimate of the Year 5 peer effect also very much depends on which instrument is used[31], but both are well below the corresponding fixed-effects estimates. The gender-based instrument suggests a very small peer effect – close to zero and nowhere near statistical significance. The age-based instrument suggests a larger peer effect of .271, but given the imprecision it too is statistically insignificant. Our preferred IV estimate using both instruments together suggests a peer effect of .117 (a normalised estimated peer effect of .05), but again it is not statistically significant. The bottom line for numeracy is that we are less sure about the magnitude of peer effects – the various estimates fall in a wider range – but the evidence still suggests that plausible peer effects exist.

[31] Recall from Section 3 that IV estimates based on the gender-balance instrument might give a lower bound estimate of the endogenous peer effect if we cannot entirely rule out direct contextual effects from gender balance.

Note that the IV estimates across all domains are highly robust to inclusion or exclusion of the observed peer average characteristics variables.[32] We also test the sensitivity of the age-based IV estimates to other cuts of the data including splitting peer ages by semester rather than trimester and replacing the trimester variable with a continuous date of birth measure. These versions of the IV tend to perform less well in the first stage, but the peer effects estimates are at least qualitatively robust, with all estimates suggesting positive peer effects of plausible magnitude.[33] We also test the sensitivity of the IV estimates to replacing the IVs specified as continuous proportions with IVs specified in discrete intervals. Again these models tend to perform less well in the first stage, but peer effects estimates remain qualitatively robust.

5. Extensions

5.1 Peer Effects Estimates for Years 7 and 9

Although we focus on Years 3 and 5 for the reasons set out in the Introduction, the NAPLAN data also cover Years 7 and 9. Here we discuss fixed-effects (but not IV[34]) estimates for these grades (see Table 9). The interpretation of these fixed-effects estimates as capturing endogenous peer effects relies on the same identification assumptions as for the fixed-effects estimates for Years 3 and 5. There is an additional identification issue at the secondary school level, however, stemming from the possibility of streaming by ability within schools (Thomson et al., 2006). Specifically, although our focus on grade-level peer effects rather than class-level peer effects means that streaming within grades will not directly impact on our estimated peer effects, sorting of teachers of different quality across cohorts within schools in a way that is correlated with cohort prior average ability could bias the estimates. For example, if the best teachers are systematically allocated to the best cohorts then this would impart an upward bias on the estimated peer effect. Our view, however, is that this is unlikely.

[32] Lin (2010) similarly estimates endogenous peer effects that are highly robust to inclusion or exclusion of contextual effects.

[33] For numeracy Year 3, the semester-based peer age instrument actually does a little better than the trimester version presented in Table 8, with a first-stage F-statistic of 8, suggesting a (normalised) peer effect of .28.

[34] Recall the argument that the *ex ante* case for the age-based instrument weakens as one moves up through school grades. This is reflected in poor first-stage explanatory power at the secondary school level: the F-statistics for the joint significance of the age-based and gender-based instruments in the first-stage regressions are below 10 in all but one subject-grade (numeracy, Year 7).

Secondary school grade cohorts also tend to be larger than primary school grades, with pupils spread between many more classes – on average they are over three times the size in our data, with 176 pupils compared to 58 pupils – so that regular interaction between grade-mates in secondary school is less likely than in primary school. Atkinson et al. (2008) suggest that studies treating the school class as the reference group at secondary school level may therefore be preferable, although this introduces additional identification problems if there is non-random sorting within school grades, and secondary school pupils may also be part of several (different but potentially overlapping) subject-based classes at any one time.[35]

Nevertheless, if the biases affecting the fixed effects estimates are small (as appears to be the case in primary school), then the fixed effects estimates presented in Table 9 should give us a reasonable idea of the magnitude of peer effects in secondary schools. First, consider the estimated peer effects for pupils in Year 7, which in Victoria is the first year of secondary schooling. Because secondary-school intakes generally come from multiple primary schools – in our data only 14.7% of Year 7 peers also shared the same primary school – and because NAPLAN tests take place early in the school year, most current peers have had less opportunity to influence the individual than is the case in other grades.[36] Peer effects between Year 7 grade-mates should be smaller than those for other grades as a result. In other words, we can interpret the Year 7 estimates as something like a placebo test (although we still expect some positive peer effect, just one that is substantially smaller in magnitude). Comparing fixed-effects estimates across grades shows this to be the case for all four subjects. Further, for three of the four subjects the point estimate in Year 7 is very close to zero and not statistically significant. The exception is numeracy, where the Year 7 estimate retains statistical significance, although it is just over half the magnitude of the equivalent estimates for the other grades. Note that the fixed-effects estimates for numeracy at Years 3 and 5 were also larger in magnitude than those for other subjects.

[35] If grade-level average test scores are acting as proxies for class-level average test scores, then there may also be an associated attenuation bias in peer effects estimates using grade-level peer scores. On the other hand, adolescents may be particularly susceptible to peer influences, which would suggest larger peer effects than those found for primary school pupils.

[36] Lavy et al. (2012) and Gibbons and Telhaj (2012) both exploit this aspect of the primary to secondary school transition to help identify peer effects from peer test scores measured at the end of primary school on own test scores measured 2-3 years into secondary school.

Second, consider the estimated peer effects for pupils in Year 9. These are large and statistically significant in all four subjects, providing further evidence that our interpretation of the Year 7 estimates above is the correct one. Also note that the estimated peer effects are no larger than those we estimate for Year 3 and 5, but neither are they consistently smaller.

5.2 Do Peer Effects Vary by Gender, Socio-economic Status, and Grade Size?

Some earlier studies in both the peer effects writ large and endogenous peer effects strands of the literature have examined evidence for heterogeneous peer effects by characteristics like gender and race in addition to homogenous peer effects (e.g. Hoxby, 2000; Angrist and Lang, 2004; Lavy et al., 2012). We can do the same here, by extending (5) to include interactions between peer test scores and own characteristics. We estimate three different extended versions of (5) as follows: including interactions between peer scores and (i) a binary dummy for being a girl; (ii) a binary dummy for being from a high-SES family (defined as mother having at least Year 12 education); (iii) a binary dummy for being in a large reference group (reference group size is above the median, within grade). The resulting estimates are presented in Tables 10 (Year 3) and 11 (Year 5).

The results clearly show stronger peer effects for boys than girls across both grades and all subjects, with the magnitude of the peer effect for girls generally around 80 percent of that for boys. In contrast, Lavy et al. (2012) find a slightly larger negative impact of having peers in the bottom 5% of the ability distribution for girls than for boys, and significant positive impacts of having peers in the top 5% of the ability distribution for girls but not for boys, both of which point to stronger peer effects for girls than for boys. Hoxby (2000) finds little difference in the peer effects writ large impact of peer gender balance on outcomes for boys or girls. Peer effects also tend to be slightly smaller for high-SES pupils than for low-SES pupils, although the differences are smaller and less clear-cut than in the case of gender. Finally, there appear to be no differences in the magnitude of the peer effect by reference group size: peer effects are similar in small school grades and large school grades. This is consistent with the similarity in the magnitude of the estimated peer effects at smaller primary schools and larger secondary schools discussed in 5.1.

5.3 Non-linear Peer Effects

One limitation of linear-in-means models of peer effects is that they tell us nothing about the merits of ability streaming (tracking) within schools (e.g. Sacerdote, 2011). A simple

extension to the linear-in-means model, however, can do so. Specifically, following Hanushek et al. (2003) and Vigdor and Nechyba (2004), we include the standard deviation of peer scores alongside the mean in (5) to examine the extent to which dispersion of peer scores is associated with own peer score. A significant negative association, interpreting grade-level dispersion as a proxy for class-level dispersion within grade, would point to potential efficiency gains from tracking. Niether Hanushek et al. (2003) (grade-level) nor Vigdor and Nechyba (2004) (grade-level and class-level) find significant association between own scores and the standard deviation of peer scores.

Table 12 presents the key estimates in our case for primary school grades 3 and 5. All but one of the eight estimated coefficients on the standard deviation of peer scores is negative and statistically significant, suggesting potential benefits from tracking. The magnitudes of these effects are quite small, with the largest coefficient (the -0.16 in spelling at level 3) suggesting that a one-standard deviation increase in the standard deviation of peer scores would decrease own score by .02 of a standard deviation. The effects are smaller still and generally statistically insignificant at secondary school grades 7 (our placebo test) and 9. Our explanation for the relative lack of tracking effects at level 9 is that grade-level dispersion is a weaker proxy for class-level dispersion at the secondary school level, where just over half of Victorian schools track at least for some subjects, than at the primary school level, where fewer than one in ten schools track.

6. Conclusions

This paper provides strong evidence for the existence of linear-in-means endogenous peer effects across multiple learning domains and school grades, using a combination of IV and fixed-effects for identification. It is one of very few papers to provide credible estimates of the endogenous peer effect in school achievement. Such endogenous peer effects matter most for policy makers: only endogenous peer effects imply social multipliers whereby interventions that improve the educational performance of some pupils in a group will have positive spill-over effects on the educational performance of other pupils in the group.

Our IV approach is closest in spirit to Goux and Maurin (2007), but our data arguably allow more credible identification because we can exploit within-school grade variation in peer ages and peer gender balance. The resulting estimates are smaller in magnitude than the

relevant estimates in Goux and Maurin (2007), but are well within the range of estimates in the wider school achievement peer effects literature. In reading, for example, a one point increase in peers' average test scores leads to between a .14 and .39 point increase in own test score. In normalised terms, this equates to a one standard deviation increase in peer test scores leading to between a .06 and .16 standard deviation increase in own test scores. We also find significant peer effects for numeracy, grammar, and spelling.

The endogenous peer effect appears slightly larger in magnitude for boys than for girls, but with no consistent patterns by SES or school size. We also present evidence that the dispersion of peer scores is negatively associated with own score, suggesting potential efficiency gains from tracking within schools.

References

Ammermueller, A. and Pischke, J.S. (2009). 'Peer effects in European primary schools: evidence from PIRLS.' *Journal of Labor Economics*, 27 (3), 315-348.

Angrist, J.D. and Lang, K. (2004). 'Does school integration generate peer effects? Evidence from Boston's Metco Program.' *American Economic Review*, 94 (5), 1613-1634.

Atkinson, A., Burgess, S., Gregg, P., Propper, C. and Proud, S. (2008). 'The impact of classroom peer groups on pupil GCSE results.' Working Paper 08/187, Centre for Market and Public Organisation, University of Bristol.

Bedard, K. and Dhuey, E. 2006. 'The persistence of early childhood maturity: international evidence of long-run age effects.' *The Quarterly Journal of Economics*, 121 (4), 1437-1472.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2011). 'Too Young to Leave the Nest? The Effects of School Starting Age.' *Review of Economics and Statistics*, 93 (2), 455-467.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2013). 'Under pressure? The effect of peers on outcomes of young adults.' *Journal of Labor Economics*, 31 (1), 119-153.

Boozer, M.A. and Cacciola, S.E (2001). 'Inside the black box of Project Star: estimation of peer effects using experimental data.' Economic Growth Center Discussion Paper No. 832, Yale University.

Boucher, V., Bramoullé, Y., Djebbari, H. and Fortin, B. (2012). 'Do peers affect student achievement? Evidence from Canada using group size variation.' *Journal of Applied Econometrics.*

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). 'Identification of peer effects through social networks.' Journal of Econometrics, 150, 41-55.

Charles, K. K. and M-C. Luoh, (2003): "Gender differences in completed schooling", *Review of Economics and Statistics*, 85 (3), 559-577.

Clark, A.E. and Lohéac, Y. (2007). 'It wasn't me, it was them! Social influence in risky behaviour of adolescents.' *Journal of Health Economics*, 26, 763-84.

Coleman, J. et al (1966). *Equality of Educational Opportunity*, Washington DC, US: Government Printing Office.

Crawford, C., Dearden, L. and Meghir, C. (2010), *When You Are Born Matters: The Impact of Date of Birth on Educational Outcomes in England*, Institute for Fiscal Studies Working Paper No: W10/06, London.

Duflo, E., Dupas, P. and Kremer, M. (2011). 'Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya.' *American Economic Review*, 101 (5), 1739-1774.

Elder, T. and Lubotsky, D. (2009). 'Kindergarten entrance age and children's achievement: impacts of state policies, family background, and peers.' *Journal of Human Resources* 44 (3), 641-683.

Fletcher, J.M. (2010). 'Social interactions and smoking: evidence using multiple student cohorts, instrumental variables and school fixed effects.' *Health Economics*, 19: 466-484.

Foster, G. and Frijters, P. (2009). 'Are there complementarities in educational peer effects?' UNSW Working Paper.

Foster, G. and Frijters, P. (2010). 'Students' beliefs about peer effects.' *Economics Letters*, 108, 260-263.

Gaviria, A. and Raphael, S. (2001). 'School-based peer effects and juvenile behaviour.' *Review of Economics and Statistics*, 83: 257-68.

Gibbons, S. and Telhaj, S. (2012). 'Peers effects: evidence from secondary school transition in England.' IZA Discussion Paper No. 6455.

Goux, D. and Maurin, E. (2007). 'Close neighbours matter: neighbourhood effects on early performance at school.' *Economic Journal*, 117 (523), 1193-1215.

Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G. (2003). 'Does peer ability affect student achievement?' *Journal of Applied Econometrics*, 18 (5), 527-544.

Hoxby, C.M. (2000). 'Peer effects in the classroom: learning from gender and race variation.' NBER Working Paper No. 7867.

Hoxby, C.M. and Weingarth, G. (2005). 'Taking race out of the equation: school reassignment and the structure of peer effects.' Department of Economics, Harvard University.

Jacob, B. A., (2002): "Where the boys aren't: Non-cognitive skills, returns to schooling and the gender gap in higher education." *Economics of Education Review*, 21 (6), 589-598.

Kling, J.R., Liebman, J. and Katz, L. (2007). Experimental analysis of neighbourhood effects.' *Econometrica*, 75 (1), 83-119.

Lavy, V., Silva, O. and Weinhardt, F. (2012). 'The good, the bad, and the average: evidence on ability peer effects in schools.' *Journal of Labor Economics*, 30 (2), 367-414.

Lavy, V. and Schlosser, A. (2011). 'Mechanisms and impacts of gender peer effects at school.' *American Economic Journal: Applied*, 3 (2), 1-33.

Lefgren, L. (2004). 'Educational peer effects and the Chicago public schools.' *Journal of Urban Economics*, 56, 169-191.

Lin, X. (2010). 'Identifying peer effects in student achievement by spatial autoregressive models with group unobservables.' *Journal of Labor Economics*, 28, 825-860.

Lundborg, P. (2006). 'Having the wrong friends? Peer effects in adolescent substance use.' *Journal of Health Economics*, 25: 214-233.

Manski, C. (1993). 'Identification of endogenous social effects: the reflection problem.' *Review of Economic Studies*, 60, 531-42.

Manski, C. (2000). 'Economic analysis and social interactions.' *Journal of Economic Perspectives*, 14(3), 115-136.

McVicar, D. (2011). 'Estimates of peer effects in adolescent smoking across twenty six European countries' *Social Science and Medicine,* 73: 1186-1193.

Moffitt, R. (2001). 'Policy interventions, low-level equilibria, and social interactions,' in *Social Dynamics*, S. Durlauf and P. Young (eds.), Cambridge: MIT Press.

Powell, L.M., John, A. and Hana, R. (2005). 'The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behaviour.' *Journal of Health Economics*, 24: 950-68.

Sacerdote, B. (2011). 'Peer effects in education: how might they work, how big are they and how much do we know thus far?' In Hanushek, E., Machin, S. and Weossman. L. (eds.) *Handbook of the Economics of Education Volume 3*, Amsterdam: Elsevier.

Sanbonmatsu, L., Kling, J.R., Duncan, G.J. and Brooks-Gunn, J. (2006). 'Neighbourhoods and academic achievement: results from the Moving to Opportunity Experiment.' *Journal of Human Resources,* 41 (4), 649-691.

Thomson, S., Lokan, J., Lamb, S. and Ainley, J., (2003) "Lessons from the Third International Mathematics and Science Study", Australian Council for Educational Research, http://research.acer.edu.au/timss_monographs/9

Vigdor, J. and Nechyba, T. (2007). 'Peer effects in North Carolina public schools.' in *Schools and the equal opportunity problem,* L. Woessmann and P. Peterson (eds.), Cambridge: MIT Press.

Table 1: Sample Means & Standard Deviations, Grades 3 and 5, by Subject

	Reading	Numeracy	Grammar	Spelling
Average Score in Y3	425.328	412.978	428.970	412.389
	(86.517)	(75.491)	(90.627)	(76.939)
Average Score in Y5	498.196	495.745	505.954	489.750
	(77.762)	(71.195)	(82.137)	(69.209)
Average Score	461.886	454.491	467.594	451.202
	(89.950)	(84.234)	(94.652)	(82.759)
Boys	0.510	0.510	0.509	0.509
	(0.500)	(0.500)	(0.500)	(0.500)
Born in trimester 1	0.252	0.252	0.252	0.252
	(0.434)	(0.434)	(0.434)	(0.434)
Proportion in 2011	0.248	0.249	0.249	0.249
	(0.432)	(0.432)	(0.432)	(0.432)
Proportion in 2010	0.246	0.246	0.246	0.246
	(0.431)	(0.431)	(0.431)	(0.431)
Proportion in 2009	0.251	0.250	0.251	0.251
	(0.433)	(0.433)	(0.433)	(0.433)
Proportion in 2008	0.255	0.255	0.255	0.255
	(0.436)	(0.436)	(0.436)	(0.436)
Proportion in Y3	0.498	0.498	0.498	0.498
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion in Y5	0.502	0.502	0.502	0.502
	(0.500)	(0.500)	(0.500)	(0.500)
Indigenous	0.015	0.015	0.015	0.015
	(0.121)	(0.120)	(0.121)	(0.121)
Non-English language background	0.245	0.246	0.246	0.246
	(0.430)	(0.431)	(0.430)	(0.430)
Father's occupation				
senior management & qual. professionals	0.136	0.136	0.136	0.136
	(0.343)	(0.343)	(0.343)	(0.343)
Other business managers, arts/media/sportspersons and associate professionals	0.188	0.188	0.187	0.187
	(0.390)	(0.390)	(0.390)	(0.390)
Tradesmen/women, clerks and skilled office, sales and service staff	0.200	0.200	0.200	0.200
	(0.400)	(0.400)	(0.400)	(0.400)
Machine operators, hospitality staff, assistants, labourers and related workers	0.194	0.194	0.194	0.194
	(0.395)	(0.395)	(0.395)	(0.395)
Not in paid work in last 12 months	0.110	0.110	0.110	0.110
	(0.313)	(0.313)	(0.313)	(0.313)
Mother's occupation				
senior management & qual. professionals	0.111	0.111	0.111	0.111
	(0.314)	(0.314)	(0.314)	(0.314)
Other business managers, arts/media/sportspersons and associate professionals	0.143	0.143	0.143	0.143
	(0.350)	(0.350)	(0.350)	(0.350)
Tradesmen/women, clerks and skilled office, sales and service staff	0.163	0.163	0.163	0.163
	(0.369)	(0.369)	(0.369)	(0.369)
Machine operators, hospitality staff, assistants, labourers and related workers	0.166	0.166	0.166	0.166
	(0.372)	(0.372)	(0.372)	(0.372)
Not in paid work in last 12 months	0.390	0.390	0.390	0.390
	(0.488)	(0.488)	(0.488)	(0.488)

Father's school education				
Year 9 or equivalent or below	0.069	0.069	0.069	0.069
	(0.253)	(0.253)	(0.253)	(0.253)
Year 10 or equivalent	0.151	0.151	0.151	0.151
	(0.358)	(0.358)	(0.358)	(0.358)
Year 11 or equivalent	0.157	0.157	0.157	0.157
	(0.364)	(0.364)	(0.364)	(0.364)
Year 12 or equivalent	0.401	0.401	0.401	0.401
	(0.490)	(0.490)	(0.490)	(0.490)
Mother's school education				
Year 9 or equivalent or below	0.069	0.069	0.069	0.069
	(0.253)	(0.253)	(0.253)	(0.253)
Year 10 or equivalent	0.144	0.143	0.144	0.144
	(0.351)	(0.351)	(0.351)	(0.351)
Year 11 or equivalent	0.185	0.185	0.185	0.185
	(0.388)	(0.388)	(0.388)	(0.388)
Year 12 or equivalent	0.527	0.527	0.527	0.527
	(0.499)	(0.499)	(0.499)	(0.499)
Father's non-school education				
No non-school qualification	0.211	0.210	0.211	0.211
	(0.408)	(0.408)	(0.408)	(0.408)
Certificate I to IV (incl. trade certificate)	0.247	0.247	0.247	0.247
	(0.431)	(0.431)	(0.431)	(0.431)
Advanced diploma/Diploma	0.084	0.084	0.084	0.084
	(0.277)	(0.277)	(0.278)	(0.278)
Bachelor degree or above	0.180	0.180	0.180	0.180
	(0.384)	(0.384)	(0.384)	(0.384)
Mother's non-school education				
No non-school qualification	0.316	0.316	0.316	0.316
	(0.465)	(0.465)	(0.465)	(0.465)
Certificate I to IV (incl. trade certificate)	0.195	0.195	0.195	0.195
	(0.396)	(0.396)	(0.396)	(0.396)
Advanced diploma/Diploma	0.122	0.122	0.122	0.122
	(0.327)	(0.328)	(0.327)	(0.327)
Bachelor degree or above	0.212	0.213	0.212	0.212
	(0.409)	(0.409)	(0.409)	(0.409)
Unisex school	0.000	0.000	0.000	0.000
	(0.020)	(0.020)	(0.019)	(0.019)
Size of peer group	58.459	58.445	58.467	58.467
	(30.323)	(30.315)	(30.321)	(30.321)
N	321094	320301	321200	321200

Table 2: Regression-based Differences in Observed Individual Characteristics by Own Gender and Age, Grades 3 & 5, Reading

	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	418.239	432.656	-14.89***	431.475	423.269	8.584***
	(87.670)	(84.689)	(0.444)	(87.448)	(86.105)	(0.496)
Average Score in Y5	492.005	504.660	-13.26***	502.859	496.624	6.588***
	(77.907)	(77.082)	(0.395)	(78.637)	(77.402)	(0.426)
Average Score	455.339	468.688	-14.07***	467.371	460.042	7.581***
	(90.738)	(88.615)	(0.297)	(90.472)	(89.699)	(0.328)
Proportion in 2011	0.248	0.248	0.000	0.249	0.248	0.001
	(0.432)	(0.432)	(0.002)	(0.432)	(0.432)	(0.002)
Proportion in 2010	0.244	0.248	-0.003**	0.244	0.247	-0.002
	(0.430)	(0.432)	(0.002)	(0.430)	(0.431)	(0.002)
Proportion in 2009	0.252	0.249	0.002	0.250	0.251	0.000
	(0.434)	(0.433)	(0.002)	(0.433)	(0.433)	(0.002)
Proportion in 2008	0.256	0.254	0.001	0.257	0.254	0.002
	(0.436)	(0.435)	(0.002)	(0.437)	(0.436)	(0.002)
Proportion in Y3	0.497	0.500		0.497	0.499	
	(0.500)	(0.500)		(0.500)	(0.500)	
Proportion in Y5	0.503	0.500		0.503	0.501	
	(0.500)	(0.500)		(0.500)	(0.500)	
Indigenous	0.014	0.016	-0.002***	0.015	0.015	0.001
	(0.117)	(0.124)	(0.000)	(0.123)	(0.120)	(0.001)
Non-English language background	0.247	0.244	0.003***	0.238	0.248	-0.007***
	(0.431)	(0.430)	(0.001)	(0.426)	(0.432)	(0.001)
Father's occupation						
senior management & qual. professionals	0.138	0.134	0.001	0.134	0.137	-0.001
	(0.345)	(0.341)	(0.001)	(0.341)	(0.344)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.190	0.185	0.003**	0.186	0.188	-0.001
	(0.392)	(0.389)	(0.001)	(0.389)	(0.391)	(0.002)
Tradesmen/women, clerks and skilled office, sales and service staff	0.200	0.199	0.001	0.201	0.200	0.000
	(0.400)	(0.400)	(0.001)	(0.400)	(0.400)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.192	0.195	-0.001	0.193	0.194	-0.003
	(0.394)	(0.397)	(0.001)	(0.395)	(0.396)	(0.002)
Not in paid work in last 12 months	0.110	0.110	0.001	0.109	0.110	-0.001
	(0.313)	(0.313)	(0.001)	(0.312)	(0.313)	(0.001)
Mother's occupation						
senior management & qual. professionals	0.113	0.109	0.002*	0.111	0.111	0.001
	(0.317)	(0.312)	(0.001)	(0.314)	(0.314)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.145	0.142	0.002	0.143	0.143	0.000
	(0.352)	(0.349)	(0.001)	(0.350)	(0.350)	(0.001)
Tradesmen/women, clerks and skilled office, sales and service staff	0.163	0.163	0.001	0.162	0.163	-0.002
	(0.369)	(0.369)	(0.001)	(0.368)	(0.369)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.166	0.166	0.002	0.168	0.166	0.002
	(0.372)	(0.372)	(0.001)	(0.374)	(0.372)	(0.002)
Not in paid work in last 12 months	0.386	0.395	-0.006***	0.389	0.391	-0.003
	(0.487)	(0.489)	(0.002)	(0.488)	(0.488)	(0.002)

Father's school education						
Year 9 or equivalent or below	0.068	0.069	-0.001	0.069	0.068	0.000
	(0.251)	(0.254)	(0.001)	(0.254)	(0.252)	(0.001)
Year 10 or equivalent	0.150	0.152	-0.002	0.154	0.150	0.002
	(0.357)	(0.359)	(0.001)	(0.361)	(0.357)	(0.002)
Year 11 or equivalent	0.157	0.157	0.001	0.159	0.156	0.001
	(0.364)	(0.363)	(0.001)	(0.365)	(0.363)	(0.001)
Year 12 or equivalent	0.405	0.396	0.006***	0.392	0.403	-0.008***
	(0.491)	(0.489)	(0.002)	(0.488)	(0.491)	(0.002)
Mother's school education						
Year 9 or equivalent or below	0.068	0.070	-0.002*	0.069	0.069	0.000
	(0.251)	(0.255)	(0.001)	(0.254)	(0.253)	(0.001)
Year 10 or equivalent	0.142	0.145	-0.002*	0.146	0.143	0.001
	(0.349)	(0.352)	(0.001)	(0.353)	(0.350)	(0.001)
Year 11 or equivalent	0.185	0.185	0.001	0.190	0.183	0.005***
	(0.388)	(0.388)	(0.001)	(0.392)	(0.387)	(0.002)
Year 12 or equivalent	0.529	0.524	0.001	0.520	0.529	-0.007***
	(0.499)	(0.499)	(0.002)	(0.500)	(0.499)	(0.002)
Father's non-school education						
No non-school qualification	0.211	0.211	0.001	0.212	0.210	0.001
	(0.408)	(0.408)	(0.001)	(0.409)	(0.407)	(0.002)
Certificate I to IV (incl. trade	0.246	0.248	-0.001	0.249	0.246	0.002
certificate)	(0.431)	(0.432)	(0.002)	(0.433)	(0.431)	(0.002)
Advanced diploma/Diploma	0.085	0.083	0.001	0.081	0.085	-0.003***
	(0.278)	(0.276)	(0.001)	(0.273)	(0.279)	(0.001)
Bachelor degree or above	0.183	0.176	0.003***	0.175	0.181	-0.003**
	(0.387)	(0.381)	(0.001)	(0.380)	(0.385)	(0.001)
Mother's non-school education						
No non-school qualification	0.314	0.319	-0.003*	0.320	0.315	0.003*
	(0.464)	(0.466)	(0.002)	(0.467)	(0.464)	(0.002)
Certificate I to IV (incl. trade	0.195	0.195	0.001	0.195	0.195	-0.001
certificate)	(0.396)	(0.396)	(0.001)	(0.396)	(0.396)	(0.002)
Advanced diploma/Diploma	0.123	0.121	0.002**	0.121	0.122	-0.001
	(0.329)	(0.326)	(0.001)	(0.326)	(0.328)	(0.001)
Bachelor degree or above	0.214	0.210	0.000	0.209	0.213	-0.002
	(0.410)	(0.408)	(0.001)	(0.407)	(0.410)	(0.002)
Unisex school	0.000	0.000	0.000	0.000	0.000	0.000*
	(0.022)	(0.018)	(0.000)	(0.022)	(0.019)	(0.000)
Size of peer group	58.479	58.438	0.002	58.376	58.486	-0.022
	(30.339)	(30.306)	(0.029)	(30.209)	(30.361)	(0.034)
N	163614	157480	321094	80769	240325	321094

Notes: Columns 1 and 2 give means separately by (own) gender. Column 3 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 4 and 5 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 6 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 3: Regression-based Differences in Observed Individual Characteristics by Peers' Gender Balance and Age, Grades 3 & 5, Reading

	Peers' sex	Peers' DOB
Average Score	-2.178	3.177
	(2.165)	(2.576)
Proportion in 2011	-0.002	0.034
	(0.058)	(0.066)
Proportion in 2010	-0.124**	-0.093
	(0.059)	(0.067)
Proportion in 2009	0.078	-0.015
	(0.058)	(0.069)
Proportion in 2008	0.049	0.073
	(0.059)	(0.067)
Indigenous	0.003	-0.003
	(0.003)	(0.003)
Non-English language background	-0.003	-0.005
	(0.008)	(0.009)
Father's occupation		
senior managmnt & qual. professionals	-0.005	0.021***
	(0.007)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	0.006	0.008
	(0.009)	(0.010)
Tradesmen/women, clerks and skilled office, sales and service staff	0.010	0.000
	(0.008)	(0.010)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.009	-0.007
	(0.008)	(0.010)
Not in paid work in last 12 months	0.001	-0.023***
	(0.007)	(0.008)
Mother's occupation		
senior managmnt & qual. professionals	-0.002	0.005
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	-0.005	0.029***
	(0.007)	(0.009)
Tradesmen/women, clerks and skilled office, sales and service staff	0.013*	-0.002
	(0.007)	(0.009)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.008	-0.018*
	(0.008)	(0.010)
Not in paid work in last 12 months	-0.001	-0.007
	(0.011)	(0.012)
Father's school education		
Year 9 or equivalent or below	-0.006	-0.014**
	(0.006)	(0.006)
Year 10 or equivalent	0.003	0.009
	(0.008)	(0.009)
Year 11 or equivalent	-0.001	0.006
	(0.008)	(0.009)
Year 12 or equivalent	-0.012	-0.004
	(0.011)	(0.012)
Mother's school education		
Year 9 or equivalent or below	0.001	0.005
	(0.005)	(0.006)
Year 10 or equivalent	0.001	-0.012
	(0.008)	(0.009)
Year 11 or equivalent	-0.014*	-0.010
	(0.008)	(0.010)
Year 12 or equivalent	-0.009	0.014
	(0.012)	(0.014)

Father's non-school education		
No non-school qualification	-0.017*	-0.001
	(0.009)	(0.010)
Certificate I to IV (incl. trade certificate)	0.005	-0.002
	(0.009)	(0.011)
Advanced diploma/Diploma	0.004	0.001
	(0.006)	(0.006)
Bachelor degree or above	-0.003	-0.003
	(0.007)	(0.008)
Mother's non-school education		
No non-school qualification	-0.013	-0.008
	(0.010)	(0.012)
Certificate I to IV (incl. trade certificate)	0.001	0.010
	(0.009)	(0.010)
Advanced diploma/Diploma	0.009	0.005
	(0.007)	(0.008)
Bachelor degree or above	-0.011	-0.006
	(0.009)	(0.010)
Unisex school	0.004	0.004*
	(0.003)	(0.002)
Size of peer group	0.060	-0.839
	(1.061)	(1.260)
N	321,094	321,094

Notes: Columns 2 and 3 give the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and the relevant instrument (gender balance of peer group, proportion of peer group born in the first trimester). Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 4: Regression-based Differences in Observed Peer Characteristics by Own Gender and Age, Grades 3 & 5, Reading

	All	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	425.328	425.445	425.207	-0.103	425.211	425.367	0.148
	(34.916)	(34.960)	(34.870)	(0.088)	(34.811)	(34.951)	(0.105)
Average Score in Y5	498.196	498.420	497.962	-0.018	497.989	498.265	0.023
	(30.832)	(30.964)	(30.693)	(0.082)	(30.811)	(30.840)	(0.089)
Average Score	461.886	462.147	461.614	-0.060	461.809	461.912	0.085
	(49.110)	(49.204)	(49.012)	(0.060)	(49.030)	(49.137)	(0.069)
Proportion of boys	0.510	0.511	0.508	-0.006***	0.510	0.510	0.000
	(0.096)	(0.096)	(0.096)	(0.000)	(0.096)	(0.096)	(0.000)
Proportion of old	0.252	0.252	0.251	0.000	0.253	0.251	-0.007***
	(0.081)	(0.081)	(0.081)	(0.000)	(0.082)	(0.081)	(0.000)
Average number of days	182.724	182.779	182.667	0.036	182.896	182.666	-1.340***
since birthday (at 30th April)	(19.737)	(19.822)	(19.648)	(0.064)	(19.917)	(19.676)	(0.092)
Indigenous	0.015	0.015	0.015	0.000	0.015	0.015	0.000
	(0.036)	(0.037)	(0.035)	(0.000)	(0.035)	(0.036)	(0.000)
Non-English language	0.245	0.245	0.246	0.000	0.243	0.246	0.000
background	(0.235)	(0.235)	(0.235)	(0.000)	(0.235)	(0.235)	(0.000)
Father not in paid work in	0.110	0.110	0.111	0.000	0.109	0.110	-0.001***
last 12 months	(0.106)	(0.106)	(0.106)	(0.000)	(0.105)	(0.106)	(0.000)
Mother not in paid work in	0.390	0.389	0.392	0.000	0.391	0.390	0.000
last 12 months	(0.159)	(0.159)	(0.159)	(0.000)	(0.159)	(0.159)	(0.000)
Father graduated Y12 or	0.401	0.402	0.399	0.000	0.398	0.401	0.000
equivalent	(0.203)	(0.204)	(0.202)	(0.000)	(0.203)	(0.204)	(0.000)
Mother graduated Y12 or	0.527	0.528	0.525	0.000	0.525	0.527	0.000
equivalent	(0.196)	(0.197)	(0.196)	(0.000)	(0.195)	(0.197)	(0.000)
N	321094	163614	157480	321094	80769	240325	321094

Notes: Column 1 gives means and standard deviations for each observed variable across the whole sample. Columns 2 and 3 give means separately by (own) gender. Column 4 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 5 and 6 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 7 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 5: Average Peer Effects in Reading by Grade, Coefficient (Standard Error)

		Level 3	Level 5
OLS	Peers' score	0.250***	0.221***
	se	(0.021)	(0.022)
IV: sex	First stage	-20.686***	-15.576***
	se	(3.042)	(2.733)
	F	46	32
	Peers' score	0.335***	0.203
	se	(0.104)	(0.142)
IV: DOB trim.	First stage	13.473***	10.526***
	se	(3.668)	(3.069)
	F	13	12
	Peers' score	0.392**	0.143
	se	(0.182)	(0.260)
IV: sex & DOB trim.	First stage sex	-20.740***	-15.510***
	se	(3.035)	(2.726)
	First stage DOB	13.584***	10.390***
	se	(3.630)	(3.045)
	F	29	23
	Peers' score	0.349***	0.188
	se	(0.090)	(0.125)
	Hansen stat	0.067847	0.043092
	Hansen P-value	0.794498	0.835552
	n	160001	161093

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. Peer group controls (in the IV models only) are as follows (proportion in the peer group): not speaking English at home, indigenous, parents out of labour force, parents graduated from Year 12. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 6: Average Peer Effects in Grammar by Grade, Coefficient (Standard Error)

		Level 3	Level 5
OLS	Peers' score	0.328***	0.262***
	se	(0.022)	(0.025)
IV: sex	First stage	-26.05***	-24.74***
	se	(3.405)	(3.073)
	F	59	65
	Peers' score	0.165	0.097
	se	(0.110)	(0.115)
IV: DOB trim.	First stage	12.712***	11.629***
	se	(3.978)	(3.364)
	F	10	12
	Peers' score	0.322	0.347*
	se	(0.221)	(0.209)
IV: sex & DOB trim.	First stage sex	-26.11***	-24.70***
	se	(3.400)	(3.062)
	First stage DOB	12.891***	11.523***
	se	(3.925)	(3.332)
	F	34	38
	Peers' score	0.188*	0.131
	se	(0.100)	(0.103)
	Hansen stat	0.323977	0.785122
	Hansen P-value	0.569227	0.375579
	n	160049	161151

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. Peer group controls (in the IV models only) are as follows (proportion in the peer group): not speaking English at home, indigenous, parents out of labour force, parents graduated from Year 12. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 7: Average Peer Effects in Spelling by Grade, Coefficient (Standard Error)

		Level 3	Level 5
OLS	Peers' score	0.244***	0.127***
	se	(0.021)	(0.022)
IV: sex	First stage	-20.2***	-18.0***
	se	(2.863)	(2.449)
	F	50	54
	Peers' score	0.155	0.020
	se	(0.123)	(0.135)
IV: DOB trim.	First stage	14.117***	8.739***
	se	(3.348)	(2.690)
	F	18	11
	Peers' score	0.415***	0.265
	se	(0.152)	(0.240)
IV: sex & DOB trim.	First stage sex	-20.31***	-18.0***
	se	(2.856)	(2.443)
	First stage DOB	14.256***	8.662***
	se	(3.297)	(2.677)
	F	34	34
	Peers' score	0.223**	0.056
	se	(0.098)	(0.118)
	Hansen stat	1.371788	0.555747
	Hansen P-value	0.241505	0.455979
	n	160049	161151

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. Peer group controls (in the IV models only) are as follows (proportion in the peer group): not speaking English at home, indigenous, parents out of labour force, parents graduated from Year 12. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 8: Average Peer Effects in Numeracy by Grade, Coefficient (Standard Error)

		Level 3	Level 5
OLS	Peers' score	0.454***	0.408***
	se	(0.016)	(0.019)
IV: sex	First stage	6.462**	12.759***
	se	(3.166)	(2.638)
	F	4	23
	Peers' score	-0.626	0.038
	se	(0.788)	(0.198)
IV: DOB trim.	First stage	8.579**	10.685***
	se	(3.518)	(3.118)
	F	6	12
	Peers' score	0.059	0.271
	se	(0.391)	(0.225)
IV: sex & DOB trim.	First stage sex	6.404**	12.797***
	se	(3.159)	(2.628)
	First stage DOB	8.519**	10.748***
	se	(3.520)	(3.111)
	F	5	18
	Peers' score	-0.241	0.117
	se	(0.378)	(0.150)
	Hansen stat	0.713161	0.516443
	Hansen P-value	0.398396	0.472363
	n	159647	160654

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. Peer group controls (in the IV models only) are as follows (proportion in the peer group): not speaking English at home, indigenous, parents out of labour force, parents graduated from Year 12. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 9: Fixed Effects Estimates of Average Peer Effects in Grades 7 & 9, by Subject, Coefficient (Standard Error)

		Level 7	Level 9
Reading	Peers' score	0.040	0.226***
	se	(0.045)	(0.043)
	n	137825	136683
Numeracy	Peers' score	0.262***	0.403***
	se	(0.035)	(0.033)
	n	137692	136867
Grammar	Peers' score	0.007	0.259***
	se	(0.042)	(0.039)
	n	138059	137591
Spelling	Peers' score	0.000	0.135***
	se	(0.043)	(0.044)
	n	138059	137591

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 10: Peer Effects across Subgroups, Fixed Effects Estimates, Level 3, by Subject, Coefficient (Standard Error)

		Gender	SES	PG Size
Reading	Peers' score	0.271***	0.269***	0.250***
	se	(0.022)	(0.024)	(0.021)
	P * subpop	-0.042***	-0.028*	-0.001
	se	(0.012)	(0.015)	(0.003)
Numeracy	Peers' score	0.509***	0.468***	0.454***
	se	(0.016)	(0.018)	(0.016)
	P * subpop	-0.111***	-0.019	-0.001
	se	(0.011)	(0.013)	(0.002)
Grammar	Peers' score	0.363***	0.364***	0.328***
	se	(0.024)	(0.025)	(0.022)
	P * subpop	-0.070***	-0.055***	0.002
	se	(0.012)	(0.016)	(0.003)
Spelling	Peers' score	0.291***	0.282***	0.244***
	se	(0.022)	(0.023)	(0.021)
	P * subpop	-0.096***	-0.059***	0.001
	se	(0.014)	(0.016)	(0.003)

Notes: These models are extended version of (4) including interactives between peer scores and (i) a binary dummy for being a girl; (ii) a binary dummy for being from a high-SES family (defined as mother having at least Year 12 education); (iii) a binary dummy for being in a large reference group (reference group size is above the median, within grade). Otherwise the models are as usual: individual control variables are sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mothers and fathers school & non-school education. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 11: Peer Effects across Subgroups, Fixed Effects Estimates, Level 5, by Subject, Coefficient (Standard Error)

		Gender	SES	PG Size
Reading	Peers' score	0.221***	0.227***	0.221***
	Se	(0.023)	(0.025)	(0.022)
	P * subpop	0.000	0.005	0.001
	Se	(0.012)	(0.015)	(0.002)
Numeracy	Peers' score	0.459***	0.403***	0.418***
	Se	(0.020)	(0.021)	(0.019)
	P * subpop	-0.083***	0.025*	0.002
	Se	(0.011)	(0.013)	(0.002)
Grammar	Peers' score	0.284***	0.279***	0.262***
	Se	(0.026)	(0.027)	(0.025)
	P * subpop	-0.045***	-0.013	0.000
	Se	(0.012)	(0.015)	(0.002)
Spelling	Peers' score	0.178***	0.165***	0.127***
	Se	(0.023)	(0.024)	(0.022)
	P * subpop	-0.104***	-0.061***	0.000
	Se	(0.015)	(0.018)	(0.002)

Notes: These models are extended version of (4) including interactives between peer scores and (i) a binary dummy for being a girl; (ii) a binary dummy for being from a high-SES family (defined as mother having at least Year 12 education); (iii) a binary dummy for being in a large reference group (reference group size is above the median, within grade). Otherwise the models are as usual: individual control variables are sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Table 12: Benefits from Tracking? Fixed Effects Estimates of Impact of Dispersion in Peer Scores, Grades 3 & 5, by Subject, Coefficient (Standard Error)

		Level 3	Level 5
Reading	Peers' mean score	0.260***	0.229***
	se	(0.021)	(0.022)
	St. dev. peers' score	-0.125***	-0.096***
	se	(0.023)	(0.025)
	n	160001	161093
Numeracy	Peers' mean score	0.473***	0.428***
	se	(0.015)	(0.019)
	St. dev. peers' score	-0.153***	-0.120***
	se	(0.020)	(0.020)
	n	159647	160654
Grammar	Peers' mean score	0.326***	0.266***
	se	(0.022)	(0.025)
	St. dev. peers' score	-0.064***	-0.033
	se	(0.019)	(0.021)
	n	160049	161151
Spelling	Peers' mean score	0.242***	0.128***
	se	(0.021)	(0.022)
	St. dev. peers' score	-0.155***	-0.072***
	se	(0.024)	(0.025)
	n	160049	161151

Notes: Individual control variables are as follows: sex, date of birth, indigenous, language spoken at home, mother's and father's occupation, mother's and father's school & non-school education. The models also control for time dummies and school fixed effects. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

Appendix: Additional Tables

A1: Regression-based Differences in Observed Individual Characteristics by Own Gender and Age, Grades 3 & 5, Numeracy

	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	418.538	407.206	10.970***	418.907	410.990	8.222***
	(79.370)	(70.781)	(0.414)	(76.402)	(75.078)	(0.421)
Average Score in Y5	502.993	488.160	14.239***	499.700	494.413	5.601***
	(74.594)	(66.616)	(0.391)	(71.837)	(70.928)	(0.387)
Average Score	460.983	447.726	12.609***	459.497	452.810	6.906***
	(87.825)	(79.761)	(0.288)	(84.434)	(84.100)	(0.288)
Proportion in 2011	0.249	0.249	0.000	0.249	0.249	0.001
	(0.432)	(0.432)	(0.002)	(0.432)	(0.432)	(0.002)
Proportion in 2010	0.245	0.247	-0.003*	0.244	0.247	-0.003
	(0.430)	(0.432)	(0.002)	(0.430)	(0.431)	(0.002)
Proportion in 2009	0.251	0.249	0.002	0.250	0.250	0.000
	(0.434)	(0.432)	(0.002)	(0.433)	(0.433)	(0.002)
Proportion in 2008	0.256	0.255	0.001	0.257	0.255	0.002
	(0.436)	(0.436)	(0.002)	(0.437)	(0.436)	(0.002)
Proportion in Y3	0.497	0.499		0.498	0.499	
	(0.500)	(0.500)		(0.500)	(0.500)	
Proportion in Y5	0.503	0.501		0.502	0.501	
	(0.500)	(0.500)		(0.500)	(0.500)	
Indigenous	0.014	0.016	-0.002***	0.015	0.014	0.001
	(0.117)	(0.124)	(0.000)	(0.123)	(0.119)	(0.001)
Non-English language background	0.247	0.245	0.003**	0.238	0.248	-0.007***
	(0.431)	(0.430)	(0.001)	(0.426)	(0.432)	(0.001)
Father's occupation						
Senior management & qual. professionals	0.138	0.135	0.001	0.134	0.137	-0.001
	(0.345)	(0.341)	(0.001)	(0.341)	(0.344)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.190	0.186	0.002*	0.186	0.188	-0.001
	(0.392)	(0.389)	(0.001)	(0.389)	(0.391)	(0.002)
Tradesmen/women, clerks and skilled office, sales and service staff	0.200	0.200	0.001	0.201	0.200	0.000
	(0.400)	(0.400)	(0.001)	(0.400)	(0.400)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.192	0.195	-0.001	0.193	0.194	-0.002
	(0.394)	(0.396)	(0.001)	(0.394)	(0.395)	(0.002)
Not in paid work in last 12 months	0.110	0.110	0.002	0.109	0.111	-0.001
	(0.313)	(0.313)	(0.001)	(0.312)	(0.314)	(0.001)
Mother's occupation						
Senior management & qual. professionals	0.113	0.109	0.002*	0.111	0.111	0.001
	(0.317)	(0.312)	(0.001)	(0.315)	(0.314)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.145	0.142	0.002	0.143	0.143	0.000
	(0.352)	(0.349)	(0.001)	(0.350)	(0.350)	(0.001)
Tradesmen/women, clerks and skilled office, sales and service staff	0.163	0.163	0.001	0.162	0.163	-0.002
	(0.369)	(0.369)	(0.001)	(0.369)	(0.369)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.166	0.166	0.002	0.168	0.166	0.002
	(0.372)	(0.372)	(0.001)	(0.374)	(0.372)	(0.002)
Not in paid work in last 12 months	0.386	0.394	-0.006***	0.389	0.390	-0.002
	(0.487)	(0.489)	(0.002)	(0.487)	(0.488)	(0.002)

Father's school education						
Year 9 or equivalent or below	0.068	0.069	-0.001	0.069	0.068	0.000
	(0.252)	(0.254)	(0.001)	(0.254)	(0.252)	(0.001)
Year 10 or equivalent	0.150	0.152	-0.002	0.154	0.150	0.002
	(0.357)	(0.359)	(0.001)	(0.361)	(0.357)	(0.002)
Year 11 or equivalent	0.157	0.156	0.001	0.159	0.156	0.001
	(0.364)	(0.363)	(0.001)	(0.365)	(0.363)	(0.001)
Year 12 or equivalent	0.405	0.397	0.005***	0.393	0.404	-0.008***
	(0.491)	(0.489)	(0.002)	(0.488)	(0.491)	(0.002)
Mother's school education						
Year 9 or equivalent or below	0.068	0.070	-0.001*	0.069	0.068	0.001
	(0.251)	(0.255)	(0.001)	(0.254)	(0.253)	(0.001)
Year 10 or equivalent	0.142	0.145	-0.002*	0.145	0.143	0.001
	(0.349)	(0.352)	(0.001)	(0.352)	(0.350)	(0.001)
Year 11 or equivalent	0.185	0.184	0.001	0.190	0.183	0.005***
	(0.388)	(0.388)	(0.001)	(0.392)	(0.387)	(0.002)
Year 12 or equivalent	0.529	0.525	0.001	0.520	0.530	-0.007***
	(0.499)	(0.499)	(0.002)	(0.500)	(0.499)	(0.002)
Father's non-school education						
No non-school qualification	0.211	0.210	0.002	0.212	0.210	0.001
	(0.408)	(0.407)	(0.001)	(0.409)	(0.407)	(0.002)
Certificate I to IV (incl. trade	0.246	0.248	-0.002	0.250	0.246	0.001
certificate)	(0.431)	(0.432)	(0.002)	(0.433)	(0.431)	(0.002)
Advanced diploma/Diploma	0.085	0.083	0.001	0.082	0.085	-0.003**
	(0.279)	(0.276)	(0.001)	(0.274)	(0.279)	(0.001)
Bachelor degree or above	0.183	0.177	0.003**	0.176	0.181	-0.003**
	(0.387)	(0.381)	(0.001)	(0.381)	(0.385)	(0.001)
Mother's non-school education						
No non-school qualification	0.313	0.318	-0.003*	0.320	0.314	0.003
	(0.464)	(0.466)	(0.002)	(0.466)	(0.464)	(0.002)
Certificate I to IV (incl. trade	0.195	0.195	0.001	0.195	0.195	-0.001
certificate)	(0.396)	(0.396)	(0.001)	(0.396)	(0.396)	(0.002)
Advanced diploma/Diploma	0.123	0.121	0.002*	0.122	0.123	0.000
	(0.329)	(0.326)	(0.001)	(0.327)	(0.328)	(0.001)
Bachelor degree or above	0.214	0.211	0.000	0.209	0.214	-0.002
	(0.410)	(0.408)	(0.001)	(0.407)	(0.410)	(0.002)
Unisex school	0.000	0.000	0.000	0.000	0.000	0.000*
	(0.022)	(0.018)	(0.000)	(0.022)	(0.019)	(0.000)
Size of peer group	58.459	58.431	0.004	58.350	58.477	-0.018
	(30.325)	(30.304)	(0.029)	(30.207)	(30.351)	(0.033)
N	163456	156845	320301	80561	239740	320301

Notes: Columns 1 and 2 give means separately by (own) gender. Column 3 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 4 and 5 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 6 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A2: Regression-based Differences in Observed Individual Characteristics by Own Gender and Age, Grades 3 & 5, Grammar

	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	418.270	440.023	-22.145***	434.707	427.048	8.317***
	(91.655)	(88.199)	(0.453)	(90.870)	(90.464)	(0.521)
Average Score in Y5	495.001	517.373	-22.944***	510.463	504.433	6.480***
	(82.893)	(79.756)	(0.400)	(82.468)	(81.970)	(0.448)
Average Score	456.853	478.741	-22.546***	472.808	465.841	7.252***
	(95.412)	(92.547)	(0.302)	(94.655)	(94.586)	(0.344)
Proportion in 2011	0.249	0.249	0.000	0.249	0.249	0.001
	(0.432)	(0.432)	(0.002)	(0.433)	(0.432)	(0.002)
Proportion in 2010	0.244	0.248	-0.003**	0.244	0.246	-0.002
	(0.430)	(0.432)	(0.002)	(0.430)	(0.431)	(0.002)
Proportion in 2009	0.252	0.249	0.002	0.250	0.251	0.000
	(0.434)	(0.433)	(0.002)	(0.433)	(0.433)	(0.002)
Proportion in 2008	0.255	0.254	0.001	0.256	0.254	0.002
	(0.436)	(0.435)	(0.002)	(0.436)	(0.435)	(0.002)
Proportion in Y3	0.497	0.499		0.497	0.499	
	(0.500)	(0.500)		(0.500)	(0.500)	
Proportion in Y5	0.503	0.501		0.503	0.501	
	(0.500)	(0.500)		(0.500)	(0.500)	
Indigenous	0.014	0.016	-0.002***	0.016	0.015	0.001
	(0.118)	(0.124)	(0.000)	(0.124)	(0.120)	(0.001)
Non-English language background	0.247	0.244	0.003**	0.238	0.248	-0.007***
	(0.431)	(0.430)	(0.001)	(0.426)	(0.432)	(0.001)
Father's occupation						
senior management & qual. professionals	0.138	0.134	0.001	0.134	0.137	-0.002
	(0.345)	(0.341)	(0.001)	(0.340)	(0.344)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.190	0.185	0.003**	0.186	0.188	-0.001
	(0.392)	(0.389)	(0.001)	(0.389)	(0.391)	(0.002)
Tradesmen/women, clerks and skilled office, sales and service staff	0.200	0.199	0.001	0.200	0.200	0.000
	(0.400)	(0.400)	(0.001)	(0.400)	(0.400)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.192	0.195	-0.001	0.193	0.194	-0.002
	(0.394)	(0.396)	(0.001)	(0.395)	(0.396)	(0.002)
Not in paid work in last 12 months	0.110	0.110	0.001	0.109	0.111	-0.001
	(0.313)	(0.313)	(0.001)	(0.312)	(0.314)	(0.001)
Mother's occupation						
senior management & qual. professionals	0.113	0.109	0.002*	0.111	0.111	0.001
	(0.317)	(0.312)	(0.001)	(0.314)	(0.314)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.145	0.142	0.002	0.143	0.143	0.000
	(0.352)	(0.349)	(0.001)	(0.350)	(0.350)	(0.001)
Tradesmen/women, clerks and skilled office, sales and service staff	0.163	0.162	0.001	0.162	0.163	-0.002
	(0.369)	(0.369)	(0.001)	(0.368)	(0.369)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.166	0.166	0.002	0.168	0.166	0.002
	(0.372)	(0.372)	(0.001)	(0.374)	(0.372)	(0.002)
Not in paid work in last 12 months	0.386	0.395	-0.007***	0.389	0.391	-0.003
	(0.487)	(0.489)	(0.002)	(0.487)	(0.488)	(0.002)

Father's school education						
Year 9 or equivalent or below	0.068	0.069	-0.001	0.069	0.068	0.001
	(0.252)	(0.254)	(0.001)	(0.254)	(0.252)	(0.001)
Year 10 or equivalent	0.150	0.152	-0.001	0.154	0.150	0.002
	(0.357)	(0.359)	(0.001)	(0.361)	(0.357)	(0.002)
Year 11 or equivalent	0.157	0.156	0.002	0.158	0.156	0.001
	(0.364)	(0.363)	(0.001)	(0.365)	(0.363)	(0.001)
Year 12 or equivalent	0.405	0.396	0.005***	0.392	0.404	-0.008***
	(0.491)	(0.489)	(0.002)	(0.488)	(0.491)	(0.002)
Mother's school education						
Year 9 or equivalent or below	0.068	0.070	-0.001	0.069	0.069	0.000
	(0.252)	(0.255)	(0.001)	(0.254)	(0.253)	(0.001)
Year 10 or equivalent	0.142	0.145	-0.003**	0.146	0.143	0.002
	(0.349)	(0.353)	(0.001)	(0.353)	(0.350)	(0.001)
Year 11 or equivalent	0.185	0.184	0.002	0.190	0.183	0.005***
	(0.388)	(0.388)	(0.001)	(0.392)	(0.387)	(0.002)
Year 12 or equivalent	0.529	0.525	0.001	0.520	0.529	-0.007***
	(0.499)	(0.499)	(0.002)	(0.500)	(0.499)	(as discussed with Chris)
Father's non-school education						
No non-school qualification	0.210	0.211	0.001	0.212	0.210	0.001
	(0.408)	(0.408)	(0.001)	(0.409)	(0.407)	(0.002)
Certificate I to IV (incl. trade certificate)	0.246	0.248	-0.001	0.250	0.246	0.002
	(0.431)	(0.432)	(0.002)	(0.433)	(0.431)	(0.002)
Advanced diploma/Diploma	0.085	0.083	0.001	0.081	0.085	-0.003***
	(0.278)	(0.277)	(0.001)	(0.274)	(0.279)	(0.001)
Bachelor degree or above	0.183	0.176	0.003***	0.175	0.181	-0.003**
	(0.387)	(0.381)	(0.001)	(0.380)	(0.385)	(0.001)
Mother's non-school education						
No non-school qualification	0.313	0.319	-0.003**	0.320	0.315	0.003*
	(0.464)	(0.466)	(0.002)	(0.467)	(0.464)	(0.002)
Certificate I to IV (incl. trade certificate)	0.195	0.195	0.001	0.195	0.195	-0.001
	(0.396)	(0.396)	(0.001)	(0.396)	(0.396)	(0.002)
Advanced diploma/Diploma	0.123	0.121	0.002**	0.121	0.122	-0.001
	(0.329)	(0.326)	(0.001)	(0.326)	(0.328)	(0.001)
Bachelor degree or above	0.214	0.211	0.000	0.209	0.213	-0.002
	(0.410)	(0.408)	(0.001)	(0.407)	(0.410)	(0.002)
Unisex school	0.000	0.000	0.000	0.000	0.000	0.000*
	(0.021)	(0.018)	(0.000)	(0.021)	(0.019)	(0.000)
Size of peer group	58.484	58.450	0.004	58.376	58.498	-0.033
	(30.338)	(30.303)	(0.030)	(30.207)	(30.359)	(0.034)
N	163581	157619	321200	80799	240401	321200

Notes: Columns 1 and 2 give means separately by (own) gender. Column 3 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 4 and 5 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 6 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A3: Regression-based Differences in Observed Individual Characteristics by Own Gender and Age, Grades 3 & 5, Spelling

	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	403.912	421.146	-17.487***	417.619	410.637	7.315***
	(79.280)	(73.422)	(0.420)	(77.222)	(76.764)	(0.447)
Average Score in Y5	481.894	497.941	-16.508***	493.042	488.640	4.728***
	(71.668)	(65.556)	(0.375)	(69.500)	(69.076)	(0.375)
Average Score	443.124	459.586	-16.996***	455.553	449.740	6.014***
	(85.016)	(79.485)	(0.281)	(82.556)	(82.776)	(0.293)
Proportion in 2011	0.249	0.249	0.000	0.249	0.249	0.001
	(0.432)	(0.432)	(0.002)	(0.433)	(0.432)	(0.002)
Proportion in 2010	0.244	0.248	-0.003**	0.244	0.246	-0.002
	(0.430)	(0.432)	(0.002)	(0.430)	(0.431)	(0.002)
Proportion in 2009	0.252	0.249	0.002	0.250	0.251	0.000
	(0.434)	(0.433)	(0.002)	(0.433)	(0.433)	(0.002)
Proportion in 2008	0.255	0.254	0.001	0.256	0.254	0.002
	(0.436)	(0.435)	(0.002)	(0.436)	(0.435)	(0.002)
Proportion in Y3	0.497	0.499		0.497	0.499	
	(0.500)	(0.500)		(0.500)	(0.500)	
Proportion in Y5	0.503	0.501		0.503	0.501	
	(0.500)	(0.500)		(0.500)	(0.500)	
Indigenous	0.014	0.016	-0.002***	0.016	0.015	0.001
	(0.118)	(0.124)	(0.000)	(0.124)	(0.120)	(0.001)
Non-English language background	0.247	0.244	0.003**	0.238	0.248	-0.007***
	(0.431)	(0.430)	(0.001)	(0.426)	(0.432)	(0.001)
Father's occupation						
senior management & qual. professionals	0.138	0.134	0.001	0.134	0.137	-0.002
	(0.345)	(0.341)	(0.001)	(0.340)	(0.344)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.190	0.185	0.003**	0.186	0.188	-0.001
	(0.392)	(0.389)	(0.001)	(0.389)	(0.391)	(0.002)
Tradesmen/women, clerks and skilled office, sales and service staff	0.200	0.199	0.001	0.200	0.200	0.000
	(0.400)	(0.400)	(0.001)	(0.400)	(0.400)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.192	0.195	-0.001	0.193	0.194	-0.002
	(0.394)	(0.396)	(0.001)	(0.395)	(0.396)	(0.002)
Not in paid work in last 12 months	0.110	0.110	0.001	0.109	0.111	-0.001
	(0.313)	(0.313)	(0.001)	(0.312)	(0.314)	(0.001)
Mother's occupation						
senior management & qual. professionals	0.113	0.109	0.002*	0.111	0.111	0.001
	(0.317)	(0.312)	(0.001)	(0.314)	(0.314)	(0.001)
Other business managers, arts/media/sportspersons and associate professionals	0.145	0.142	0.002	0.143	0.143	0.000
	(0.352)	(0.349)	(0.001)	(0.350)	(0.350)	(0.001)
Tradesmen/women, clerks and skilled office, sales and service staff	0.163	0.162	0.001	0.162	0.163	-0.002
	(0.369)	(0.369)	(0.001)	(0.368)	(0.369)	(0.002)
Machine operators, hospitality staff, assistants, labourers and related workers	0.166	0.166	0.002	0.168	0.166	0.002
	(0.372)	(0.372)	(0.001)	(0.374)	(0.372)	(0.002)
Not in paid work in last 12 months	0.386	0.395	-0.007***	0.389	0.391	-0.003
	(0.487)	(0.489)	(0.002)	(0.487)	(0.488)	(0.002)

Father's school education						
Year 9 or equivalent or below	0.068	0.069	-0.001	0.069	0.068	0.001
	(0.252)	(0.254)	(0.001)	(0.254)	(0.252)	(0.001)
Year 10 or equivalent	0.150	0.152	-0.001	0.154	0.150	0.002
	(0.357)	(0.359)	(0.001)	(0.361)	(0.357)	(0.002)
Year 11 or equivalent	0.157	0.156	0.002	0.158	0.156	0.001
	(0.364)	(0.363)	(0.001)	(0.365)	(0.363)	(0.001)
Year 12 or equivalent	0.405	0.396	0.005***	0.392	0.404	-0.008***
	(0.491)	(0.489)	(0.002)	(0.488)	(0.491)	(0.002)
Mother's school education						
Year 9 or equivalent or below	0.068	0.070	-0.001	0.069	0.069	0.000
	(0.252)	(0.255)	(0.001)	(0.254)	(0.253)	(0.001)
Year 10 or equivalent	0.142	0.145	-0.003**	0.146	0.143	0.002
	(0.349)	(0.353)	(0.001)	(0.353)	(0.350)	(0.001)
Year 11 or equivalent	0.185	0.184	0.002	0.190	0.183	0.005***
	(0.388)	(0.388)	(0.001)	(0.392)	(0.387)	(0.002)
Year 12 or equivalent	0.529	0.525	0.001	0.520	0.529	-0.007***
	(0.499)	(0.499)	(0.002)	(0.500)	(0.499)	(0.002)
Father's non-school education						
No non-school qualification	0.210	0.211	0.001	0.212	0.210	0.001
	(0.408)	(0.408)	(0.001)	(0.409)	(0.407)	(0.002)
Certificate I to IV (incl. trade	0.246	0.248	-0.001	0.250	0.246	0.002
certificate)	(0.431)	(0.432)	(0.002)	(0.433)	(0.431)	(0.002)
Advanced diploma/Diploma	0.085	0.083	0.001	0.081	0.085	-0.003***
	(0.278)	(0.277)	(0.001)	(0.274)	(0.279)	(0.001)
Bachelor degree or above	0.183	0.176	0.003***	0.175	0.181	-0.003**
	(0.387)	(0.381)	(0.001)	(0.380)	(0.385)	(0.001)
Mother's non-school education						
No non-school qualification	0.313	0.319	-0.003**	0.320	0.315	0.003*
	(0.464)	(0.466)	(0.002)	(0.467)	(0.464)	(0.002)
Certificate I to IV (incl. trade	0.195	0.195	0.001	0.195	0.195	-0.001
certificate)	(0.396)	(0.396)	(0.001)	(0.396)	(0.396)	(0.002)
Advanced diploma/Diploma	0.123	0.121	0.002**	0.121	0.122	-0.001
	(0.329)	(0.326)	(0.001)	(0.326)	(0.328)	(0.001)
Bachelor degree or above	0.214	0.211	0.000	0.209	0.213	-0.002
	(0.410)	(0.408)	(0.001)	(0.407)	(0.410)	(0.002)
Unisex school	0.000	0.000	0.000	0.000	0.000	0.000*
	(0.021)	(0.018)	(0.000)	(0.021)	(0.019)	(0.000)
Size of peer group	58.484	58.450	0.004	58.376	58.498	-0.033
	(30.338)	(30.303)	(0.030)	(30.207)	(30.359)	(0.034)
N	163581	157619	321200	80799	240401	321200

Notes: Columns 1 and 2 give means separately by (own) gender. Column 3 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 4 and 5 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 6 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A4: Regression-based Differences in Observed Individual Characteristics by Peers' Gender Balance and Age, Grades 3 & 5, Numeracy

	Peers' sex	Peers' DOB
Average Score	-5.475**	0.417
	(2.188)	(2.480)
Proportion in 2011	0.002	0.032
	(0.058)	(0.065)
Proportion in 2010	-0.102*	-0.096
	(0.059)	(0.067)
Proportion in 2009	0.070	-0.003
	(0.058)	(0.069)
Proportion in 2008	0.030	0.066
	(0.059)	(0.067)
Indigenous	0.003	-0.003
	(0.003)	(0.003)
Non-English language background	-0.002	-0.003
	(0.008)	(0.009)
Father's occupation		
senior management & qual. professionals	-0.004	0.023***
	(0.007)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	0.002	0.008
	(0.009)	(0.010)
Tradesmen/women, clerks and skilled office, sales and service staff	0.013	-0.005
	(0.008)	(0.010)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.010	-0.008
	(0.008)	(0.010)
Not in paid work in last 12 months	0.000	-0.021***
	(0.007)	(0.008)
Mother's occupation		
senior management & qual. professionals	-0.002	0.007
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	-0.010	0.031***
	(0.008)	(0.009)
Tradesmen/women, clerks and skilled office, sales and service staff	0.015**	-0.003
	(0.007)	(0.009)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.007	-0.018*
	(0.008)	(0.010)
Not in paid work in last 12 months	0.000	-0.013
	(0.011)	(0.012)
Father's school education		
Year 9 or equivalent or below	-0.007	-0.013*
	(0.006)	(0.006)
Year 10 or equivalent	0.000	0.006
	(0.008)	(0.009)
Year 11 or equivalent	0.002	0.008
	(0.008)	(0.009)
Year 12 or equivalent	-0.011	-0.005
	(0.011)	(0.012)
Mother's school education		
Year 9 or equivalent or below	0.000	0.006
	(0.005)	(0.006)
Year 10 or equivalent	0.002	-0.011
	(0.008)	(0.009)
Year 11 or equivalent	-0.011	-0.008
	(0.008)	(0.010)
Year 12 or equivalent	-0.011	0.011
	(0.012)	(0.014)

Father's non-school education		
No non-school qualification	-0.016*	-0.002
	(0.009)	(0.010)
Certificate I to IV (incl. trade certificate)	0.006	-0.003
	(0.009)	(0.011)
Advanced diploma/Diploma	0.004	0.002
	(0.006)	(0.006)
Bachelor degree or above	-0.003	0.000
	(0.007)	(0.008)
Mother's non-school education		
No non-school qualification	-0.010	-0.009
	(0.010)	(0.012)
Certificate I to IV (incl. trade certificate)	0.003	0.013
	(0.009)	(0.010)
Advanced diploma/Diploma	0.009	0.003
	(0.007)	(0.008)
Bachelor degree or above	-0.012	-0.003
	(0.009)	(0.010)
Unisex school	0.004	0.004*
	(0.003)	(0.002)
Size of peer group	0.147	-0.681
	(1.061)	(1.249)
N	320,301	320,301

Notes: Columns 2 and 3 give the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and the relevant instrument (gender balance of peer group, proportion of peer group born in the first trimester). Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A5: Regression-based Differences in Observed Individual Characteristics by Peers' Gender Balance and Age, Grades 3 & 5, Grammar

	Peers' sex	Peers' DOB
Average Score	1.101	4.877*
	(2.396)	(2.741)
Proportion in 2011	-0.005	0.035
	(0.059)	(0.065)
Proportion in 2010	-0.119**	-0.091
	(0.059)	(0.068)
Proportion in 2009	0.073	-0.003
	(0.058)	(0.069)
Proportion in 2008	0.051	0.059
	(0.059)	(0.067)
Indigenous	0.004	-0.003
	(0.003)	(0.003)
Non-English language background	0.000	-0.002
	(0.008)	(0.009)
Father's occupation		
senior management & qual. professionals	-0.006	0.021***
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	0.006	0.008
	(0.009)	(0.010)
Tradesmen/women, clerks and skilled office, sales and service staff	0.010	-0.003
	(0.008)	(0.010)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.008	-0.007
	(0.008)	(0.010)
Not in paid work in last 12 months	0.001	-0.019**
	(0.007)	(0.008)
Mother's occupation		
senior management & qual. professionals	-0.002	0.004
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	-0.007	0.029***
	(0.008)	(0.009)
Tradesmen/women, clerks and skilled office, sales and service staff	0.015**	-0.001
	(0.007)	(0.009)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.007	-0.015
	(0.008)	(0.010)
Not in paid work in last 12 months	-0.005	-0.011
	(0.011)	(0.012)
Father's school education		
Year 9 or equivalent or below	-0.007	-0.015**
	(0.006)	(0.006)
Year 10 or equivalent	0.003	0.013
	(0.008)	(0.009)
Year 11 or equivalent	0.004	0.007
	(0.008)	(0.009)
Year 12 or equivalent	-0.013	-0.005
	(0.011)	(0.012)
Mother's school education		
Year 9 or equivalent or below	0.000	0.007
	(0.005)	(0.006)
Year 10 or equivalent	0.002	-0.011
	(0.008)	(0.009)
Year 11 or equivalent	-0.011	-0.008
	(0.008)	(0.010)
Year 12 or equivalent	-0.013	0.012
	(0.012)	(0.014)

Father's non-school education		
No non-school qualification	-0.015*	-0.003
	(0.009)	(0.010)
Certificate I to IV (incl. trade certificate)	0.005	0.001
	(0.009)	(0.011)
Advanced diploma/Diploma	0.005	0.000
	(0.006)	(0.006)
Bachelor degree or above	-0.004	-0.002
	(0.007)	(0.008)
Mother's non-school education		
No non-school qualification	-0.012	-0.005
	(0.010)	(0.012)
Certificate I to IV (incl. trade certificate)	-0.003	0.009
	(0.009)	(0.010)
Advanced diploma/Diploma	0.009	0.003
	(0.007)	(0.008)
Bachelor degree or above	-0.012	-0.007
	(0.009)	(0.010)
Unisex school	0.003	0.004*
	(0.003)	(0.002)
Size of peer group	0.137	-1.234
	(1.077)	(1.256)
N	321,200	321,200

Notes: Columns 2 and 3 give the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and the relevant instrument (gender balance of peer group, proportion of peer group born in the first trimester). Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A6: Regression-based Differences in Observed Individual Characteristics by Peers' Gender Balance and Age, Grades 3 & 5, Spelling

	Peers' sex	Peers' DOB
Average Score	1.683	3.919*
	(1.979)	(2.252)
Proportion in 2011	-0.005	0.035
	(0.059)	(0.065)
Proportion in 2010	-0.119**	-0.091
	(0.059)	(0.068)
Proportion in 2009	0.073	-0.003
	(0.058)	(0.069)
Proportion in 2008	0.051	0.059
	(0.059)	(0.067)
Indigenous	0.004	-0.003
	(0.003)	(0.003)
Non-English language background	0.000	-0.002
	(0.008)	(0.009)
Father's occupation		
senior management & qual. professionals	-0.006	0.021***
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	0.006	0.008
	(0.009)	(0.010)
Tradesmen/women, clerks and skilled office, sales and service staff	0.010	-0.003
	(0.008)	(0.010)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.008	-0.007
	(0.008)	(0.010)
Not in paid work in last 12 months	0.001	-0.019**
	(0.007)	(0.008)
Mother's occupation		
senior management & qual. professionals	-0.002	0.004
	(0.006)	(0.007)
Other business managers, arts/media/sportspersons and associate professionals	-0.007	0.029***
	(0.008)	(0.009)
Tradesmen/women, clerks and skilled office, sales and service staff	0.015**	-0.001
	(0.007)	(0.009)
Machine operators, hospitality staff, assistants, labourers and related workers	-0.007	-0.015
	(0.008)	(0.010)
Not in paid work in last 12 months	-0.005	-0.011
	(0.011)	(0.012)
Father's school education		
Year 9 or equivalent or below	-0.007	-0.015**
	(0.006)	(0.006)
Year 10 or equivalent	0.003	0.013
	(0.008)	(0.009)
Year 11 or equivalent	0.004	0.007
	(0.008)	(0.009)
Year 12 or equivalent	-0.013	-0.005
	(0.011)	(0.012)
Mother's school education		
Year 9 or equivalent or below	0.000	0.007
	(0.005)	(0.006)
Year 10 or equivalent	0.002	-0.011
	(0.008)	(0.009)
Year 11 or equivalent	-0.011	-0.008
	(0.008)	(0.010)
Year 12 or equivalent	-0.013	0.012
	(0.012)	(0.014)

Father's non-school education		
No non-school qualification	-0.015*	-0.003
	(0.009)	(0.010)
Certificate I to IV (incl. trade certificate)	0.005	0.001
	(0.009)	(0.011)
Advanced diploma/Diploma	0.005	0.000
	(0.006)	(0.006)
Bachelor degree or above	-0.004	-0.002
	(0.007)	(0.008)
Mother's non-school education		
No non-school qualification	-0.012	-0.005
	(0.010)	(0.012)
Certificate I to IV (incl. trade certificate)	-0.003	0.009
	(0.009)	(0.010)
Advanced diploma/Diploma	0.009	0.003
	(0.007)	(0.008)
Bachelor degree or above	-0.012	-0.007
	(0.009)	(0.010)
Unisex school	0.003	0.004*
	(0.003)	(0.002)
Size of peer group	0.137	-1.234
	(1.077)	(1.256)
N	321,200	321,200

Notes: Columns 2 and 3 give the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and the relevant instrument (gender balance of peer group, proportion of peer group born in the first trimester). Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A7: Regression-based Differences in Observed Peer Characteristics by Own Gender and Age, Grades 3 & 5, Numeracy

	All	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	412.978	413.109	412.842	-0.197**	412.793	413.040	-0.016
	(32.002)	(31.966)	(32.039)	(0.091)	(32.047)	(31.987)	(0.098)
Average Score in Y5	495.745	496.049	495.426	-0.108	495.572	495.803	0.038
	(31.496)	(31.566)	(31.420)	(0.081)	(31.337)	(31.549)	(0.090)
Average Score	454.491	454.792	454.178	-0.152**	454.380	454.529	0.011
	(52.159)	(52.238)	(52.076)	(0.061)	(52.130)	(52.170)	(0.066)
Proportion of boys	0.510	0.512	0.508	-0.006***	0.510	0.510	0.000
	(0.096)	(0.096)	(0.096)	(0.000)	(0.096)	(0.096)	(0.000)
Proportion of old	0.252	0.252	0.251	0.000	0.253	0.251	-0.007***
	(0.081)	(0.082)	(0.081)	(0.000)	(0.082)	(0.081)	(0.000)
Average number of days	182.737	182.799	182.673	0.054	182.936	182.670	-1.329***
since birthday (at 30th April)	(19.797)	(19.886)	(19.704)	(0.064)	(19.982)	(19.735)	(0.093)
Indigenous	0.015	0.015	0.015	0.000	0.015	0.015	0.000
	(0.036)	(0.036)	(0.035)	(0.000)	(0.035)	(0.036)	(0.000)
Non-English language	0.246	0.246	0.246	0.000	0.243	0.247	0.000
background	(0.236)	(0.235)	(0.236)	(0.000)	(0.235)	(0.236)	(0.000)
Father not in paid work in	0.110	0.110	0.111	0.000	0.110	0.110	-0.001***
last 12 months	(0.106)	(0.106)	(0.107)	(0.000)	(0.105)	(0.107)	(0.000)
Mother not in paid work in	0.390	0.389	0.391	0.000	0.390	0.390	0.000
last 12 months	(0.159)	(0.159)	(0.159)	(0.000)	(0.158)	(0.159)	(0.000)
Father graduated Y12 or	0.401	0.402	0.399	0.000	0.398	0.402	0.000
equivalent	(0.203)	(0.204)	(0.202)	(0.000)	(0.203)	(0.204)	(0.000)
Mother graduated Y12 or	0.527	0.529	0.526	0.000	0.526	0.528	0.000
equivalent	(0.196)	(0.197)	(0.196)	(0.000)	(0.195)	(0.197)	(0.000)
N	320301	163456	156845	320301	80561	239740	320301

Notes: Column 1 gives means and standard deviations for each observed variable across the whole sample. Columns 2 and 3 give means separately by (own) gender. Column 4 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 5 and 6 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 7 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A8: Regression-based Differences in Observed Peer Characteristics by Own Gender and Age, Grades 3 & 5, Grammar

	All	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	428.970	429.067	428.869	0.006	428.818	429.021	0.134
	(36.230)	(36.324)	(36.133)	(0.097)	(36.217)	(36.235)	(0.110)
Average Score in Y5	505.954	506.152	505.747	0.055	505.775	506.014	0.160
	(31.904)	(32.023)	(31.778)	(0.090)	(31.941)	(31.891)	(0.097)
Average Score	467.594	467.828	467.350	0.030	467.524	467.617	0.130*
	(51.443)	(51.547)	(51.333)	(0.066)	(51.436)	(51.445)	(0.073)
Proportion of boys	0.509	0.511	0.507	-0.006***	0.509	0.509	0.000
	(0.096)	(0.096)	(0.096)	(0.000)	(0.096)	(0.096)	(0.000)
Proportion of old	0.252	0.252	0.251	0.000	0.253	0.251	-0.007***
	(0.081)	(0.082)	(0.081)	(0.000)	(0.082)	(0.081)	(0.000)
Average number of days	182.750	182.813	182.686	0.050	182.942	182.686	-1.319***
since birthday (at 30th April)	(19.726)	(19.814)	(19.634)	(0.064)	(19.932)	(19.656)	(0.091)
Indigenous	0.015	0.015	0.015	0.000	0.015	0.015	0.000
	(0.036)	(0.037)	(0.035)	(0.000)	(0.035)	(0.037)	(0.000)
Non-English language	0.246	0.245	0.246	0.000	0.243	0.246	0.000
background	(0.235)	(0.235)	(0.236)	(0.000)	(0.235)	(0.235)	(0.000)
Father not in paid work in	0.110	0.110	0.111	0.000	0.110	0.110	-0.001**
last 12 months	(0.106)	(0.106)	(0.106)	(0.000)	(0.105)	(0.107)	(0.000)
Mother not in paid work in	0.390	0.389	0.392	0.000	0.391	0.390	0.000
last 12 months	(0.159)	(0.159)	(0.159)	(0.000)	(0.159)	(0.159)	(0.000)
Father graduated Y12 or	0.401	0.402	0.399	0.000	0.398	0.402	0.000
equivalent	(0.204)	(0.205)	(0.203)	(0.000)	(0.203)	(0.204)	(0.000)
Mother graduated Y12 or	0.527	0.528	0.525	0.000	0.525	0.527	0.000
equivalent	(0.196)	(0.197)	(0.196)	(0.000)	(0.195)	(0.197)	(0.000)
N	321200	163581	157619	321200	80799	240401	321200

Notes: Column 1 gives means and standard deviations for each observed variable across the whole sample. Columns 2 and 3 give means separately by (own) gender. Column 4 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 5 and 6 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 7 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.

A9: Regression-based Differences in Observed Peer Characteristics by Own Gender and Age, Grades 3 & 5, Spelling

	All	Boys	Girls	Difference	Old	Young	Difference
Average Score in Y3	412.389	412.443	412.333	0.014	412.292	412.421	0.141
	(29.427)	(29.570)	(29.279)	(0.083)	(29.466)	(29.414)	(0.092)
Average Score in Y5	489.750	489.936	489.557	0.079	489.587	489.805	0.069
	(25.174)	(25.308)	(25.032)	(0.072)	(25.250)	(25.147)	(0.077)
Average Score	451.202	451.409	450.987	0.047	451.168	451.214	0.105*
	(47.388)	(47.519)	(47.251)	(0.055)	(47.390)	(47.387)	(0.060)
Proportion of boys	0.509	0.511	0.507	-0.006***	0.509	0.509	0.000
	(0.096)	(0.096)	(0.096)	(0.000)	(0.096)	(0.096)	(0.000)
Proportion of old	0.252	0.252	0.251	0.000	0.253	0.251	-0.007***
	(0.081)	(0.082)	(0.081)	(0.000)	(0.082)	(0.081)	(0.000)
Average number of days since birthday (at 30th April)	182.750	182.813	182.686	0.050	182.942	182.686	-1.319***
	(19.726)	(19.814)	(19.634)	(0.064)	(19.932)	(19.656)	(0.091)
Indigenous	0.015	0.015	0.015	0.000	0.015	0.015	0.000
	(0.036)	(0.037)	(0.035)	(0.000)	(0.035)	(0.037)	(0.000)
Non-English language background	0.246	0.245	0.246	0.000	0.243	0.246	0.000
	(0.235)	(0.235)	(0.236)	(0.000)	(0.235)	(0.235)	(0.000)
Father not in paid work in last 12 months	0.110	0.110	0.111	0.000	0.110	0.110	-0.001**
	(0.106)	(0.106)	(0.106)	(0.000)	(0.105)	(0.107)	(0.000)
Mother not in paid work in last 12 months	0.390	0.389	0.392	0.000	0.391	0.390	0.000
	(0.159)	(0.159)	(0.159)	(0.000)	(0.159)	(0.159)	(0.000)
Father graduated Y12 or equivalent	0.401	0.402	0.399	0.000	0.398	0.402	0.000
	(0.204)	(0.205)	(0.203)	(0.000)	(0.203)	(0.204)	(0.000)
Mother graduated Y12 or equivalent	0.527	0.528	0.525	0.000	0.525	0.527	0.000
	(0.196)	(0.197)	(0.196)	(0.000)	(0.195)	(0.197)	(0.000)
N	321200	163581	157619	321200	80799	240401	321200

Notes: Column 1 gives means and standard deviations for each observed variable across the whole sample. Columns 2 and 3 give means separately by (own) gender. Column 4 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and a male dummy. Columns 5 and 6 give means separately by (own) position in the age distribution of the school year (old = born in the first trimester for standard entry into the school year, young = born during the rest of the year). Column 7 gives the regression coefficient and associated (clustered) standard error from regressions of each of the observed variables on school-grade fixed effects and an 'old' dummy. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 99%/95%/90% levels.