**QUEEN'S
UNIVERSITY
BELFAST**

Achievement Effects from New Peers: Who Matters to Whom?

McVicar, D., Moschion, J., & Ryan, C. (2018). Achievement Effects from New Peers: Who Matters to Whom? *Economics of Education Review*. <https://doi.org/10.1016/j.econedurev.2018.07.008>

Published in:
Economics of Education Review

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
Copyright 2018 Elsevier.
This manuscript is distributed under a Creative Commons Attribution-NonCommercial-NoDerivs License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Achievement Effects from New Peers: Who Matters to Whom?

Duncan McVicar , Julie Moschion , Chris Ryan

PII: S0272-7757(17)30596-4
DOI: <https://doi.org/10.1016/j.econedurev.2018.07.008>
Reference: ECOEDU 1823

To appear in: *Economics of Education Review*

Received date: 28 September 2017
Revised date: 18 July 2018
Accepted date: 31 July 2018

Please cite this article as: Duncan McVicar , Julie Moschion , Chris Ryan , Achievement Effects from New Peers: Who Matters to Whom?, *Economics of Education Review* (2018), doi: <https://doi.org/10.1016/j.econedurev.2018.07.008>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Achievement Effects from New Peers: Who Matters to Whom?

Duncan McVicar^{*,***,#}, Julie Moschion^{*****} & Chris Ryan^{**}

*Queen's Management School
Queen's University Belfast
Riddel Hall, 185 Stranmillis Road, Belfast BT9 5EE, United Kingdom
d.mcvicar@qub.ac.uk

**Melbourne Institute of Applied Economic and Social Research
University of Melbourne
Vic 3010, Australia
moschion@unimelb.edu.au
ryan.c@unimelb.edu.au
Tel: +61(0)3 8344 2100

*** IZA, Bonn

July 2018

JEL: I21, I24, J24

Keywords: peer effects, school achievement, education

Acknowledgements

This research was commissioned by the Victorian Department of Education and Training (DET). This paper uses unit record data from the NAPLAN administrative data collection, provided by DET. Thanks also to seminar participants at University Erlangen-Nuremberg, Melbourne Institute, University of Tilburg, University of Sydney, ESAM, and University of Zurich for comments and suggestions. The findings and views reported in this paper are those of the authors and should not be attributed to DET or any other branch of the Victorian or Australian Commonwealth government. Conflicts of interest: none.

Corresponding author.

Abstract

This paper presents estimates of achievement-related peer effects on school students' literacy using data from national test scores, across multiple literacy measures and student cohorts, for the population of public secondary school students in Years 7 and 9 (aged 12/13 and 14/15 years) in the Australian state of Victoria. Identification is achieved via individual fixed effects and by distinguishing between secondary school peers who attended the same primary school as the individual and those who did not. Estimates of peer effects are based on the new peers, whose primary school achievement could not have been affected by the individual. The results provide strong evidence for the existence of peer effects, with small but positive and statistically significant effects from having higher-achieving peers on average and from having a higher proportion of very high-achieving peers. Further, it is individuals in the middle of the ability distribution who benefit most from having high achieving peers.

Highlights

- Educational achievement peer effects are estimated for Australian school students
- The transition from primary to secondary school is exploited for identification
- Peer effects estimates are based only on peers from different primary schools
- Small but positive and statistically significant average peer effects are found
- There is also a shining light effect which is strongest for middle achieving students

1. Introduction

Peer effects refer to externalities in which the actions or characteristics of a reference group affect an individual's behaviour or outcomes. They have been studied in numerous contexts. This paper examines a specific form of peer effects, related to the effect on a student's achievement of the achievement of his or her peers, which if the effects are of sufficient magnitude, have critical implications for students, parents, schools and policy makers. Effects from peer averages imply that parents can improve their child's expected educational achievement by selecting a school with a higher-ability intake, and reforms introducing greater school choice could widen educational inequalities. Non-linear peer effects have additional efficiency implications, such that schools may be able to improve the average achievement of their students, and policy makers the efficiency of the schooling system as a whole, by manipulating the allocation of students across classes or schools. As a result, peer effects in school achievement have attracted an enormous amount of attention in the literature. Establishing the existence and magnitude of peer effects, however, is beset by practical difficulties (Manski, 1993, 2000; Moffitt, 2001, Angrist 2014), a fact that has probably also contributed to the longevity and ubiquity of the literature. Despite the vast number of studies, disagreement as to the nature, magnitude and even the existence of peer effects in school achievement remains.

The two key difficulties in estimating achievement peer effects arise from endogenous sorting into or within schools and the fact that the individual is a peer of their peers and may therefore influence their peers – the reflection problem. One branch of the school achievement peer effects literature has attempted to address the second challenge by utilising measures of peer quality that pre-date any potential social interactions between the individual and their peers. A promising new strand of this literature exploits the transition between primary and secondary school (so far only in England), and the fact that most of the secondary school peers of any individual attended a different primary school (Lavy et al., 2012; Gibbons and Telhaj, 2016; Zhang, 2016; Mendolia et al., 2018). Specifically, these papers use the prior performance of peers when they were in different schools – a measure that is immune from reflection problems – to measure the impact of peers' earlier achievement on current student performance. Lavy et al. (2012) also exploits multiple measures of individual student achievement in various learning domains to remove any fixed student-specific effect on achievement, which mitigates confounding effects from sorting into schools, making for a very strong identification strategy. They find no evidence of an average

peer effect on individual achievement. Gibbons and Telhaj (2016), Zhang (2016) and Mendolia et al. (2018), however, all find some evidence of small, positive average peer effects. All four studies use data drawn from the same underlying data set; the English National Pupil Database.

The research design adopted in this paper also exploits this transition, when students encounter new classmates at the start of secondary school who did not attend the same primary school. The secondary school achievement of individuals is regressed on the primary school achievement of their secondary school new peers, for whom there is no reflection problem. The approach also makes use of multiple measures of individual achievement, specifically a set of four language or literacy-related measures of achievement. This enables use of an individual fixed effects approach to account for common individual effects on achievement across the four language-related subjects of reading, writing, grammar and spelling. Because the data cover multiple cohorts of students, school-subject fixed effects can also be included to wash out time-invariant correlated effects that might differ across subjects. In other words, peer effects are estimated by exploiting within-individual variation in the performance of new peers across the literacy domain when they were in primary school.

In contrast to the other papers in this strand of the literature, however, this paper uses data from Australian schools rather than English schools and focuses on outcomes at age 12/13 rather than outcomes at age 13/14 or later. In doing so the paper makes a standalone contribution to the knowledge base supporting education policy makers by shedding light on how well the conclusions of these earlier English studies generalise to a similar but not identical educational context. Its single most important contribution, however, is to present evidence of peer effects, including small but non-trivial average peer effects, using a strong identification approach exploiting within-individual variation in test scores. A new placebo test is also added.

Specifically, this paper estimates peer effects using administrative data on the test scores of students attending public schools in the Australian state of Victoria. The data come from the National Assessment Program – Literacy and Numeracy (NAPLAN) conducted across Australia, which provides test scores for five subjects – numeracy, reading, spelling, grammar, and writing – for all students 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, and the paper exploits data from 2008 to 2013. This provides

four cohorts of Year 5 (primary school) students subsequently observed in Year 7 (in secondary school), and two cohorts with students also observed in Year 9.

In addition to estimating average peer effects, this paper also assesses the impact of having a high proportion of peers who were in either the top or bottom 10% of the achievement distribution, while also allowing these effects to differ between genders and across the distribution of individual ability (that is, allowing for non-linear and heterogeneous peer effects). This allows testing of empirical support in Australian secondary schools for a wide range of models including the ‘bad apple’ (disruptive students harm everyone), the ‘shining light’ (excellent students provide a great example for all), the ‘invidious comparison’ (outcomes are harmed by the presence of better achieving peers), and ultimately to assess the possible benefits of tracking, at least at the school grade level (see Sacerdote, 2011).

The remainder of this paper is set out as follows. The next section briefly reviews those studies in the school achievement peer effects literature to which this paper most closely speaks. The data are described in more detail in section 3, along with further discussion of the approach to estimation. The main results are discussed in section 4, with extensions in section 5 and concluding discussion in section 6.

2. Related Literature

Recent studies on peer effects in school achievement have looked for both average peer effects and peer effects from and to different points in the ability distribution, using a variety of strategies to overcome the identification problems associated with reflection and endogenous sorting. One strand of this literature exploits random or quasi-random assignment to new peer groups stemming from programs like Metco in the US (Angrist and Lang, 2004) or the Extra Teacher Program in Kenya (Duflo et al., 2011). These studies make credible claims to estimating causal peer effects (or more precisely the absence of causal peer effects) that are internally valid. But a dearth of suitable similar experiments in other contexts means we cannot yet draw general conclusions from this strand of literature alone. A second (non-experimental) group of studies regresses own achievement on lagged peer achievement under the assumption that an individual’s current achievement cannot impact her peers’ prior achievement (e.g. Hanushek et al., 2003; Lefgren, 2004; Vigdor and Nechyba, 2007; Atkinson et al., 2008). Manski (1993) questions the extent to which this approach truly solves the reflection problem on the grounds that peer scores are likely to be persistent. Additional identification steps – typically school fixed effects and individual controls – are also required

to try to wash out sorting effects, arguably with mixed success. Angrist (2014) also critiques this approach for being susceptible to ‘mechanical’ biases due, among other things, to measurement error and negative intra-group correlation between own test score and the leave-out-mean test score of peers. A third group of studies exploits demographic variation in peer group composition to instrument for peer achievement (e.g. Goux and Maurin, 2007), with the extent to which the resulting estimates can be interpreted as causal peer effects hanging on the plausibility of the exclusion restrictions. Sacerdote (2011) provides a review, summing up the body of evidence from this literature as “mixed”.

The current paper contributes to a fourth group of studies – Lavy et al. (2012), Gibbons and Telhaj (2016), Zhang (2016) and Mendolia et al. (2018) – which arguably falls somewhere between the first and second literature strands. Like the second group of studies described above, these studies also regress own achievement on lagged (school grade) peers’ achievement, with a number of additional identification steps to address sorting and other correlated effects. Like the first set of studies described above, however, they focus on the impact of being grouped with *new* peers, in this case as a result of the transition between primary and secondary school in England.¹ These studies are immune to reflection problems because the prior scores (in primary school) of new peers (in secondary school) cannot have been impacted by the individual. This research design also mitigates problems associated with “Angrist mechanics” such as the negative intra-group correlation.² Further, Lavy et al. (2012) makes a convincing case that correlation due to endogenous sorting is washed out by basing conclusions only on within-individual variation in educational achievement across different learning domains (English, mathematics and science).³ Although Lavy et al. (2012) finds no evidence of average peer effects – in contrast to the other studies in this strand of the literature – it does find that high achieving peers impact positively on low achieving girls (but not boys), and that both genders, across the ability distribution, suffer from having low achieving peers (bad apples). These effects, although statistically significant, are small in magnitude. Mendolia et al. (2018), but not Gibbons and Telhaj (2016), also finds some evidence of bad apple effects. Zhang (2016) does not examine non-linear peer effects.

¹ Mendolia et al. (2018) takes a slightly different approach from the others in using primary school test scores of peers of peers to instrument for the secondary school test scores of peers.

² The negative intra-group correlation arises because of the use of the leave-out mean used in this literature, which is unnecessary in the construction of the peer effects variable from new peers only.

³ This within-individual cross-domain approach has also been used in various studies of other education-related questions (see, for example, Dee, 2005; Lavy 2010; Schwerdt and Wuppermann, 2011; Altinok and Kingdon, 2012).

3. Data and Identification

The Australian testing system, NAPLAN, was introduced in 2008, with tests taking place each year in grades 3, 5, 7 and 9. Test score data from 2008 to 2013 are available for the population of students in the relevant grades in public schools in the state of Victoria. This provides four cohorts of Year 5 (primary school) students who are subsequently observed in Year 7 (in secondary school) from 2010 to 2013. In addition, there are two cohorts with students observed in Year 9 in 2012 or 2013. For both the Year 7 and Year 9 cohorts, secondary school peer groups are defined using the Year 7 data, and Year 7 peers who attended the same primary school (old peers) are differentiated from those who attended a different primary school (new peers). For both groups, Year 5 achievement outcomes are used as measures of new and old peer group quality.⁴ The focus here is primarily on the results from the analysis of Year 7 literacy, though results for Year 9 are also reported.

NAPLAN scores across all grades are measured on a scale from 0 to 1000, and are designed to measure absolute (rather than relative) competence in the subject concerned.⁵ These scores are converted within individual subjects, by year and grade, but across all schools, to percentile ranks. In the analysis, all individuals with the same subject score are given the same rank. This approach avoids the problem that scores across different subjects in NAPLAN are not intended to be directly comparable. The inclusion of separate subject-level effects in the regression equation is also intended to deal with this issue.

The process by which peer groups are constructed is as follows. Scores for all identified students in all four literacy-related subjects are required in the current year, as well as in the past. Since many students miss one test in a particular year, limiting the sample to only those with complete responses and allowing the peer effects to reflect only the achievement

⁴ Some schools, particularly in remote areas, encompass both primary and secondary school stages, so most secondary students in such schools only have old peers. These schools are excluded from the analysis. Only 1.2% of students are dropped from the sample because they have no new peers.

⁵ Average NAPLAN scores in a particular subject area are therefore higher in higher school grades. Note that just under 7% and 9% of students at Years 7 and 9 respectively, were not entered for, or otherwise missed, NAPLAN tests in 2010. Less than 5% of students did not sit the tests in Years 3 and 5 in 2010.

percentiles of those with complete data reduces the sample. Therefore, two data sets are constructed – one where only individuals with complete information were included (the *restricted* sample) and another where those with partial information were also included (the *unrestricted* sample), which increases the number of observations by about 10,000 for both the Year 7 and Year 9 estimates. The results discussed in the main text are those for the *unrestricted* sample, but we also report results for the *restricted* sample in the sensitivity analysis.

In addition to test score data, each student's school and grade (but not class) is identified and used to construct peer test scores. There is also 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 as controls at the individual level in OLS models but not in individual fixed effect models. This information is also used to examine evidence for heterogeneous peer effects.

Table 1 gives summary statistics for Year 7 students, for all students, 'regular' students (those whose scores lie between the bottom 10 percent and top 10 percent), as well as for those with at least one subject where they were in the top 10 percent of students and for those with at least one subject where they were in the bottom 10 percent of students. Unsurprisingly, the average rank in any Year 7 subject is always close to 50. More surprisingly, the average rank of students with at least one score in the top 10 percent is 10 percentile points lower than 90, suggesting the average rank of other scores is often a rank within the 70s, well down on their maximum score. Importantly, this demonstrates that there is significant variation in ranks within individuals across literacy subjects (to be exploited in our identification strategy). As expected, higher performing students have parents with higher levels of education and who work in management and professional or associate professional occupations. Males and those from Indigenous backgrounds are more likely to have achievement in at least one subject in the bottom 10 percent of students.

The size of the average Year 7 peer group is around 174 students. From Table 2, about 83 percent of those Year 7 peers were new peers. Indeed, there are 1287 primary schools for just 330 secondary schools in our sample and there is a lot of reshuffling between primary and secondary school in Australia. As a result, on average in each secondary school, students come from 36 different primary schools. The average percentile rank of the new peers was a

little below 50, their share being in the top 10 percent a little lower than 10 percent, and their share in the bottom 10 percent a little higher.⁶ Table 3 contains an analysis of the variance in the student's and his new peers' percentile rank variables. Although most of the variance arises from variation between students, there is also significant variation across subjects within students, representing between 35% and 65% of the between variation.

One feature of these data is that they include only students in public schools. Yet the transition from primary to secondary school is precisely the time when many students in Australia exit the public system to attend private schools. Although these lost peers do not have any internal validity implications for the analysis, they may limit generalisability beyond the specific set of individuals observed in public secondary schools in the relevant years. These private schools remain a key concern of policy makers in Australia, in part because they typically receive substantial public subsidies, but also because they draw disproportionately from the upper end of the ability distribution, which may have both equity and efficiency implications in the presence of peer effects.

The approach here is to examine whether students who have new peers with higher ability in one subject (e.g. reading) than in the other subjects (writing, spelling, grammar) have better test scores in that subject compared to the other subjects. The estimation allows for individual, subject and school-subject fixed effects, denoted in equation (1) below by α_i , β_q and η_{qs} respectively. Estimation proceeds separately for grades 7 and 9. The subject effects are allowed to vary by gender, though note that only subjects in the broad language domain are considered, at least initially. In addition, it is anticipated that individual i 's achievement A in subject q in grade g will be influenced by peer average effects measured by performance in an earlier grade ($g = c$), denoted by $\bar{A}_{qcst}$, as well as by the proportion of high (top 10%) and low (bottom 10%) achieving peers they have, denoted by $\bar{P}_{qcst}^h$ and $\bar{P}_{qcst}^l$, respectively. Rather than entering these peer effects aggregated to the whole grade level, however, the effects of 'new' and 'old' peers are distinguished via $\bar{A}_{qcst}^n$ and $\bar{A}_{qcst}^o$, where 'new' peers are those in grade g at school s with individual i who did not attend the same school as individual i in grade c , while the 'old' peers attended the same school as individual i in grades c and g . The same distinction is made for the variables denoting high achieving and low achieving peers. Further, an individual i 's achievement in subject q (say, reading) in grade g will also be

⁶ The departures from 10% reflect the movement of higher achieving students from public to private schools in the transition from primary to secondary education common in Australia. Hence, more than 10% of the remaining students come from the original bottom decile, fewer than 10% from the original top decile.

influenced by her prior achievement in grade $g = c$ in that subject, A_{iqcst} , along with the cross-subjects ($q = d, e, f$, i.e. writing, spelling and grammar) A_{idcst} , A_{iecst} and A_{ifcst} (with subject specific coefficients). This gives equation (1) as:

$$(1) \quad A_{iqgst} = \alpha_i + \beta_q + \beta_q \times \text{gender} + \eta_{qs} + \delta_1 \bar{A}_{qcst}^n + \delta_2 \bar{A}_{qcst}^o + \theta_1 \bar{P}_{qcst}^{nh} + \theta_2 \bar{P}_{qcst}^{nl} + \theta_3 \bar{P}_{qcst}^{oh} + \theta_4 \bar{P}_{qcst}^{ol} + \lambda_{1q} A_{iqcst} + \lambda_{2q} A_{idcst} + \lambda_{3q} A_{iecst} + \lambda_{4q} A_{ifcst} + \varepsilon_{iqgst}$$

where δ_1 , δ_2 , θ_1 , θ_2 , θ_3 , θ_4 , λ_1 , λ_2 , λ_3 and λ_4 are parameters to be estimated and ε_{iqgst} is a random error term.

The main coefficients of interest are δ_1 , θ_1 , and θ_2 , the parameters on the prior test scores of new peers and on the proportion of new peers who are high or low ability. Following the earlier papers in this strand of the literature, estimated associations with old peer test scores are not interpreted as peer effects, although they may be partly driven by peer effects. Rather, old peers are included as controls for shared primary school*cohort*subject effects that might persist into secondary school. The inclusion of school*subject dummies in equation (1) is to control for possible time-invariant subject-specific sorting into schools, although such effects seem unlikely here given that all four subjects are literacy-related.

While the data are different, the specification of equation (1) is similar to the preferred specification of Lavy et al. (2012), although here the high/low achieving peer variables are defined at 10% thresholds rather than 5% thresholds, and school by subject fixed effects are included instead of school by year fixed effects.⁷ The various fixed effects wash out potentially confounding effects from sorting and other correlated effects, and the estimates are immune from reflection problems and common shocks (e.g. teacher effects) because only the impact of prior test scores of new peers and in a single learning domain is considered. Given the data, there is no perfect placebo test to examine whether any biases still remain in the most complete specification. But something close to such a test can be obtained by replacing the individual's current (Year 7) test score with her Year 5 test score, predating the transition to secondary school. Peer effects should not be possible in this scenario because the individual is yet to meet her new peers. Any sorting effects, however, would still show up as

⁷ The 10% threshold is the one where the estimated top/bottom peer effects are largest. Sensitivity to the particular choice of threshold is discussed in Section 4.2. School by year fixed effects are not included because they did not change the estimates that did not include school by subject fixed effects.

correlation between own and new peer scores in Year 5. The one weakness of this test is that own prior score cannot be included as a control, so it is a placebo test on not quite the preferred specification.

This research design mitigates the threat to placing a causal interpretation on the resulting estimates from the “Angrist mechanics”. The leave-out mean is not used, so one source of mechanical bias – from the negative correlation between own score and peer score within group – is avoided. Systematic rather than random variation in group composition means weak instrument bias is also unlikely to be an issue.⁸ Peer groups are also not symmetric in the sense that they vary within the secondary school grade for those from different feeder schools. To support these claims a further placebo test is conducted which keeps the real peers of individuals but randomly mixes average peer scores across the different subjects for each individual.

What are estimated here are ‘net’ peer effects, i.e. net of ability spillovers between subjects. If such spillovers were strong enough they could mask subject-specific spillovers so that this approach would find zero peer effects. In other words we are setting an extremely high bar for peer effects.

4. Main Results and Discussion

4.1 Homogenous Peer Effects in Year 7

Table 4 contains parameter estimates on the new peers variables included in equation (1) for Year 7 students (OLS equivalents – without individual student fixed effects – are given in Appendix Table 1). The first three columns show the parameters on the average, bottom 10% and top 10% new peer effects when each of the peer variables are entered singly in equation (1), and without controlling for old peers or any individual prior achievement scores. Each subsequent column in Table 4 corresponds to a version of (1) with an increasing degree of control, with the full version of equation (1) as presented in the previous section corresponding to specification 6 in what follows. As we will see, estimated peer effects tend to fall in absolute value as tighter controls are progressively included.

⁸ The within-individual design and the suite of other included fixed effects mean there is not an endogenous sorting price to pay for this non-random group composition.

The first three columns (1a-1c) of Table 4 point to substantial and positive within-individual association between own score and the average prior score of new peers, as well as significant associations with the proportion of new peers in the bottom 10% of the ability distribution (negative) or with the proportion of new peers in the top 10% of the ability distribution (positive). The effects are of the order that a one-point increase in the average percentile rank of new peers is associated with a 0.4 unit rank increase in the subject's percentile rank, and a one percentage point increase in the proportion of top and bottom students is associated with a 0.3 and -0.2 change in the percentile rank of the individual, respectively.

The second set of results (specification 2) show the parameter estimates when the three new peers variables are entered simultaneously in equation (1), but still without controlling for old peers or individual prior achievement. Inclusion of the proportions at the top and bottom of the achievement distribution has only a small effect on the average effect parameter, but including the average substantially reduces the parameter on the effect of the proportion of top and bottom achievers. From this point on, none of the estimates point to there being any detrimental impact from having a higher share of peers in the bottom 10% of the achievement distribution over and above the effect that works through the peer average effect. This contrasts with the findings of Lavy et al. (2012) and, for 18-year olds if not for 16-year olds, with Mendolia et al. (2018) for England.

The third set of results show the parameter estimates when the three peer variables are entered in equation (1), this time controlling for the old-peer equivalents, but still without any individual prior achievement scores. The old peer variables are intended to control for shared primary school*cohort*subject effects that might persist into secondary school, and the estimated parameters on the new peers variables change little with the inclusion of the old peer variables.⁹

The fourth set of results show the parameter estimates when controls for same-subject own prior score interacted with subject dummy variables are included. This is intended to address sorting into schools that might be subject specific, in addition to the individual fixed effect which washes out non-subject specific sorting. The interaction with subject dummies allows own prior score to impact differently in different subjects. Again, both remaining associations

⁹ Note that, while not reported, the estimates on the old peers variables are always smaller in absolute values than the new peer values reported in Table 4. See Zhang (2016) for a discussion of a potential downwards bias in the estimated old peer effect.

between own (Year 7) score and the scores of new peers fall in absolute value, but remain statistically significant.

The fifth set of results builds on the fourth by including cross-subject own prior scores interacted with subject dummy variables. This allows for cross-subject spillovers from own prior scores that differ by subject. Inclusion of these additional controls lowers the estimated magnitudes of the effects, which are of the order that an increase in the average percentile rank of new peers by one-unit rank is associated with an individual's percentile rank that is around 0.1 of a unit rank higher. A one percentage point increase in the proportion of top students is associated with a 0.03 higher percentile rank in the subject, substantially smaller effects than those reported in the simplest specification. Nevertheless, these estimated effects remain non-trivial in magnitude and statistically significant. The cross-subject within-individual approach sets a high bar for finding evidence of peer effects, but the peer effects still show through.

The sixth and final set of results in Table 4 builds on the fifth by including school by subject dummies to wash out any remaining time-invariant subject-school unobservables, e.g. subject-specific sorting reflecting school specialism in a particular subject. This is the preferred specification here, since it offers a further degree of control over and above specification 5 above, and these results lay the strongest claim in the paper to a causal interpretation. Again, there is no impact from low-achieving peers in the bottom 10% of the achievement distribution (the coefficient is negative but very small and nowhere near statistical significance). There is also little change in the estimated impact from high-achieving peers. The effect from the peer average decreases by a third. A one percentage point increase in the proportion of students in the top 10% of the achievement distribution increases own score by 0.04 of a percentile rank. A one-point increase in the average percentile rank of new peers leads to a 0.08 increase in the percentile rank of the individual in Year 7.¹⁰

4.2 Robustness

The outcomes of a number of robustness tests, in each case of the sixth specification of the model with the highest degree of control, are presented in Table 5. Overall, these tests

¹⁰ Sample size falls as we add more controls, so that the sample for specification 6 in table 4 is around 15% smaller than that for specification 1. Appendix Table 2 re-estimates all 6 specifications with the common (specification 6) sample. The results are almost identical to those in Table 4.

confirm the presence of peer effects from the average and the top but not from the bottom, although the magnitudes vary. First, using the restricted sample (only those individuals for whom we observe test scores in all subjects at both Year 5 and Year 7) in place of the unrestricted sample leads to almost identical estimates. Second, restricting the sample to the first two cohorts only (whose Year 9 achievement we examine in Section 5.2) suggests a similar pattern of peer effects, albeit with a smaller effect from the average and a higher effect from the top 10%.

Third, a version of (1) where numeracy is added to the set of student achievement measures is estimated. The addition of numeracy does not change the estimated average peer effect or the zero effect from low-achieving peers, but does lead to a larger estimated impact of the top peer effect (it increases to 0.057 from 0.038). This may in part reflect the arguments put forward by Bettinger (2012) for differences in behavioural responsiveness between reading and mathematics and is consistent with the larger estimated peer effects for maths scores than for English scores reported by Gibbons and Telhaj (2016). Note that the difference between the 5th and 6th specifications – with and without school*subject dummies – is a little larger with numeracy than in the literacy-only case, consistent with there being more scope for school specialization when considering the wider set of subjects spanning two learning domains.

Fourth, raw test scores are used in place of percentiles. So far raw test scores have been converted to percentiles, which are more comparable across grades and subjects than the raw scores. But the inclusion of subject dummies in (1) and the grade-level focus should mean that the estimates are not qualitatively sensitive to this conversion. This is indeed the case: estimating (1) using raw scores suggests a small but positive and statistically significant average peer effect, a positive and statistically significant impact from high-achieving peers in the grade, and zero impact from low-achieving peers.

Finally, we explore the robustness of the estimated impacts of high/low achieving peers to varying the thresholds at either end of the distribution. Figure 1 shows how the estimated low and high-achieving peer effects vary as the threshold is varied from top/bottom 5% through to top/bottom 25%. There is no threshold for which the bottom peer effect is negative and statistically significant, and the biggest negative impact is where the threshold is set at 10%. In contrast, the statistical significance (but not the sign) of the high-achieving peer effect is sensitive to threshold, again with the biggest positive impact corresponding to the 10% threshold.

4.3. Placebo Tests

Confidence in the interpretation of the sixth set of estimates in Table 4 as causal would be increased were there no ‘evidence’ of peer effects in scenarios where peer effects are not possible, by design. The biggest remaining concern in terms of how confidently these estimates can be interpreted as peer effects are whether there might be any remaining subject-specific sorting that is not washed out by own prior scores and school*subject dummies. Two placebo tests are conducted to examine these threats to the causal interpretation. Results are presented in Table 6.

The first test replaces own current score with own prior score on the left-hand side of (1) (own prior score therefore obviously gets dropped from the right-hand side of (1)). Peer effects should be impossible in this scenario because the individual is yet to meet the new peers. But any remaining sorting effects would still show up as correlation between own and new peer scores in Year 5. Both the average peer effect and the high-achieving peer effect disappear in this case but the low-achieving peer effect becomes statistically significant at the 10% level. So the test suggests broad support for the interpretation of the relevant Table 4 estimates as capturing causal peer effects. Were the Table 4 coefficients to be adjusted by subtracting the relevant Table 5 coefficients, only the estimated impact of high-achieving peers would be qualitatively sensitive. Both the zero peer effect from low-achieving peers and the positive average peer effect would remain unchanged.

The second test randomly mixes the achievement scores of an individual’s new peers across the different subjects. Since within-individual variation in average peer scores across subjects is the main source of variation exploited in estimation of peer effects in this paper, these should disappear if the peer scores are randomised, i.e. peer effects should be ruled out in this scenario, even if there are cross-subject spillovers, by this randomisation. The results show that the estimated effects are all small in magnitude and not statistically different from zero.

5. Extensions

5.1 Heterogeneity in the Year 7 Peer Effects

Studies have examined evidence for heterogeneous peer effects by characteristics like gender, race and ability in addition to homogenous peer effects. Table 7 contains parameter estimates on the new peers variables for males and females separately and across the distribution of prior school achievement for males and females jointly and separately (using specification 6 of table 4). Table 7 panel A shows that the average new peer effect varies little across the distribution of prior achievement, broadly consistent with Gibbons and Telhaj (2016). In contrast, the estimates suggest that the high-achieving peer effect is much smaller for those at the top of the prior achievement distribution, and largest for those in the middle of the distribution, although confidence intervals for the estimates across the range overlap. This is consistent with ‘shining-light’ and inconsistent with ‘invidious-comparison’ peer effects. Note that the impact from low-achieving peers is everywhere small and statistically insignificant.

Panels B and C of Table 7 show that the average peer effects are common to boys and girls, though the magnitude of the average effect is marginally larger for females. The positive effect from the proportion of high-achieving students is predominantly a female phenomenon; the estimated parameter is smaller and not significantly different from zero for males. The zero impact from low-achieving peers is common to both genders. The gender-specific estimates across the prior achievement distribution reflect those in Panel A, i.e. no variation in the average peer effect across the distribution, stronger shining light effects in the middle of the distribution, and no convincing evidence of bad apple effects at any point in the distribution, with the partial exception of a small and marginally statistically significant negative effect from bottom 10% peers at the bottom of the ability distribution for girls.

Both Lavy et al. (2012) and Zhang (2016) restrict their sample to smaller schools, in the former case to try to reduce the extent to which ability grouping within grades takes place in the sampled schools. Table 8 reports estimated new peer effects by school size (such that half of the students attend small schools and half of the students attend large schools). Limiting the sample to students from small schools does have an impact on the estimated average peer effects, with the relevant coefficient falling from 0.081 in Table 4 to 0.039 in Table 8 (significant only at the 10% level) when only students in the smallest 50% of schools are used, and to 0.030 (and not significant even at the 10% level) when only students in the smallest 33% of schools are used. Conversely, we find larger estimated average peer effects for larger schools.

Were this capturing the effect of less tracking within grades in smaller schools, the estimated peer effects might be expected to be larger in small schools than in large schools.

Indeed, more tracking implies grade-level measures of peer quality that involve greater measurement error and therefore smaller peer effects in large schools. More likely is that restricting to small schools in the Australian context – where the population is highly concentrated in the big cities and sparse elsewhere – means restricting to rural schools, some with multiple grades in a single class, which likely absorb the full population from feeder primary schools and with relatively few new peers compared to old peers.¹¹ With more old peers and fewer new peers, smaller schools may offer less scope for interacting with new peers leading to smaller peer effects given our research design. Interestingly, while the percentage of new peers is always very high in larger schools, it is more heterogeneous in smaller schools, offering an opportunity to test whether peer effects increase with the proportion of new peers (in small schools).

To that end Table 8 also contains estimates where the proportion of new peers in the grade is split into quintiles. If differences between small and large schools in the proportions of new peers explain the difference in average peer effects by school size, then we would also expect increasing average peer effects with the proportion of new peers within small schools. And indeed, for small schools the average effect of new peers tends to increase with the proportion of new peers, although with a statistically insignificant tailing off for the top quintile. There is no clear pattern by proportion of new peers in the estimated impacts of high-achieving peers, and higher standard errors make these estimates statistically insignificant with the exception of the larger and marginally significant estimate for the top quintile. The estimated impact of low-achieving peers in small schools increases in absolute magnitude with the proportion of new peers, becoming statistically significant and negative for the highest quintile. This suggests that a higher proportion of new peers may be conducive to larger estimated peer effects, given our research design, at least up to a certain level.

5.2 Peer effects for Year 9

Table 9 contains parameter estimates on the new peers variables for Year 9 students. Own and peer prior scores are measured in Year 5, as before. The motivation for estimating Year 9 peer effects in addition to the Year 7 peer effects is twofold. First, comparing both sets of estimates may be informative about whether achievement peer effects from new peers build

¹¹ The percentage of new peers is significantly lower in small schools than in large schools with an average of 77% compared to 88%, and many more students in small schools that are in grades with fewer than 50% new peers (which is the case of 12.7% of students in small schools but only 0.3% of those in large schools).

up over time spent together, or fade as Year 5 achievement becomes a more distant proxy for current peer quality. Second, NAPLAN tests in Year 9 take place closer to the age at which the equivalent tests in England, studied by the earlier cited papers, take place. Note that just two cohorts of Year 9 students are available for the analysis. (Table 5 presents the corresponding estimates of Year 7 peer effects restricted to these two cohorts.)

The progression of the results in Table 9 from substantial to small, as controls for old peers, individual prior achievement and school*subject dummies are added, is similar to that for the Year 7 students in Table 4. In the preferred (sixth) specification, the magnitude of the statistically significant and positive average peer effect is slightly smaller than that for Year 7 students (but in line with the Year 7 estimates restricted to the same two cohorts), and again there is no evidence for an additional peer effect specifically from low-achieving peers. The positive impact from high-achieving peers reported in Table 4 (and for the same two cohorts in Table 5), however, is not there in Year 9; this effect actually becomes negative, although non-significant and sensitive to precise model specification. One interpretation could be that the average new peer effect neither substantially accumulates nor substantially declines over time; it is immediate or almost immediate, long-lasting, but not cumulative. In support of this, estimating (1) where Year 5 prior scores for individuals and peers are replaced by Year 7 prior scores yields zero peer effects for all three new peer variables. In other words, the impact of new peers is already reflected in the Year 7 own prior score. The disappearing impact of the top 10% of peers is consistent with attenuation bias reflecting a greater degree of ability tracking in Year 9. As for the Year 7 results, these estimates are broadly robust to a number of different sample restrictions and other variations. Estimates vary by gender (only girls are affected by their peers and only by the average new peer variable) and by school size (smaller and statistically insignificant for small schools).

6. Discussion

This paper provides evidence about the existence of achievement-related peer effects in the first year of secondary schooling. It exploits measurements of peer quality prior to any social interaction and uses individual and school-subject fixed-effects for identification. This approach controls for common individual factors that influence achievement across subjects and domains, such as ability, motivation and background factors, along with school by

subject effects common across individuals. Together, this provides a strong basis for interpreting any resulting peer effects as causal.

The estimates presented here do indeed suggest that there are positive and significant average peer effects and that the proportion of top peers has an additional positive impact on individual achievement. These effects are small. Imagine two otherwise similar students – one with new peers who were 10 percentile rank points higher in their average peer score for a particular subject (just over a one standard deviation increase) than another student. This would translate into a 0.8 percentile rank gain for the student with better peers in Year 7, slightly smaller for a Year 9 student. This is equivalent to an improvement of about 0.05 of the within-student standard deviation of Year 7 test scores. The impact of the proportion of good peers an individual has is smaller again. But bear in mind the approach to identification here is very conservative.

The results presented here are also partly consistent with existing evidence that peer effects may be heterogeneous along various dimensions. The average effect is common to boys and girls, though its magnitude is marginally larger for females, and varies little across the prior achievement distribution. The positive effect from the proportion of high achievement students is statistically significant for girls but not for boys, and is strongest for those in the middle of the prior achievement distribution.

In presenting evidence of an average achievement peer effect, this paper contrasts with recent zero findings in the (quasi-)experimental literature and in Lavy et al. (2012) – this paper's closest predecessor. Further, the sensitivity analysis presented here suggests the contrast with Lavy et al. (2012) is not easily explained, e.g. by differences in the subjects studied or differences in the thresholds adopted for low/high achieving peers. Rather, the explanation may lie in institutional differences between the English schools and Australian schools studied, although note that the different studies using English data in this particular strand of the literature themselves draw contrasting conclusions. One potential culprit is differences in the extent of ability grouping within schools between England and Australia, between subject areas and by age-group. Such differences are relevant in part because grade-level measures of peer ability are a poorer proxy for class-level peer ability where such tracking is more extensive. Tracking is less common in language-related subjects than in mathematics and science, less common in the first year of secondary school than in later years, and less prevalent more generally in Australian public schools than in English public

schools (see Thomson et al., 2003; Forgasz, 2010; Lavy et al., 2012).¹² The results presented here might therefore be comparatively free from this source of attenuation bias. Evidence that tracking also reduces peer effects within classes, particularly those by which low-ability students can benefit from being paired with high-ability students, is presented by Kimbrough et al. (2017). Another potential culprit is differences in the extent to which ‘high-stakes’ national testing has been embedded in the school systems of England and Australia. National testing for the grades studied here was only introduced in Australia in 2008 but has been in place in England since 1988. Further, league tables for English secondary schools, based on these tests, have been reported since 1992, whereas such tables were never officially released for Australian schools (although the media generated their own league tables using NAPLAN data from 2010 onwards). The implication is that the English schools have had longer to develop behavioural responses to national testing – e.g. teaching to the test – and may have more incentive to do so than the Australian schools studied here. The result may be more direct teaching methods (e.g. Au, 2007), where bad apple effects are presumably more likely, and less collaboration in the classroom (e.g. Thompson, 2013), where average peer effects may similarly be less likely.

A key policy implication of the lack of a bad apple effect and the presence of a shining light effect is that further ability tracking in Australian schools would have no beneficial effect but would reduce the potential for the positive impact of high achievers on those further down the ability distribution observed in the data. Having said that, the magnitude of the estimated peer effects presented here suggest that the achievement peer-effects-related costs or benefits of more or less tracking in countries such as Australia are likely to be small.

We conclude with one final thought experiment relating to the roughly 12% of each cohort in the final year of public primary school who transit not into public secondary schools but private secondary schools, including elite private schools. The loss of these students induces concerns about the “residualisation” of low-achievement and low socio-economic status students in public schools. Imagine that some intervention induced these students to stay predominantly in the public system and that any students who transferred were no longer disproportionately drawn from the top of the achievement distribution. In that case, both the

¹² For example, Lavy et al. (2012) state that ability tracking takes place within 46%, 59% and 80% of schools in their sample at Year 9 for English, science and mathematics, respectively. Thomson et al. (2003) report comparable Australian figures of 58% for maths and <20% for science in Year 8, which are likely to be upper bounds on the extent of ability tracking in Year 7 given tracking increases with age and tracking is more common in non-government schools within Australia. Forgasz (2010) reports survey-based proportions of 37% (70%) of public schools in Victoria that ability track in mathematics in Year 7 (Year 9).

proportion of new peers from the top and bottom deciles in public secondary schools would average 10%, while the average peer rank would be 50%. Would ability-related peer effects lead to a substantial achievement gain in the public school sector under this scenario? The mechanisms for such a gain are the positive impact of high achieving peers on others, particularly those in the middle of the ability distribution, and the average peer effect. The answer is ‘not really’; this scenario would have only a small impact on student achievement in public schools because the estimated peer effects are small. A 1.5 percentage point increase in the top group, multiplied by the 0.038 parameter from Table 4, and an increase in average peer quality of 2.5 percentage points combine to provide the equivalent of a 0.26 percentile improvement in average achievement across all public school students. This is less than 0.02 of the within-student standard deviation of Year 7 test scores and constitutes an estimated effect just one-tenth of the magnitude Hattie (2009) views as “small”.

References

- Altinok, N. and Kingdon, G. (2012). 'New evidence on class size effects: a pupil fixed effects approach.' *Oxford Bulletin of Economics and Statistics* 74(2): 207-234.
- Angrist, J.D. (2014). 'The Perils of Peer Effects.' *Labour Economics* 30: 98-108.
- 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.
- Au, W. (2007). 'High-stakes testing and curricular control: A meta-synthesis.' *Educational Researcher* 36(5): 258-267.
- Bettinger, E.P. (2012). 'Paying to learn: the effect of financial incentives on elementary school test scores.' *The Review of Economics and Statistics* 94(3): 686-698.
- Dee, T.S. (2005). 'A teacher like me: Does race, ethnicity, or gender matter?' *American Economic Review* 95(2): 158-165.
- 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.
- Forgasz, H. (2010). 'Streaming for Mathematics in Years 7-10 in Victoria: An Issue of Equity?' *Mathematics Education Research Journal* 22(1): 57-90.
- Gibbons, S. and Telhaj, S. (2016). 'Peer effects: evidence from secondary school transition in England.' *Oxford Bulletin of Economics and Statistics* 78(4): 548-575.

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.

Hattie, J.A.C (2009), *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*, Routledge, Abingdon, UK.

Kimbrough, E.O., McGee, A.D. and Higeoka, H. (2017). 'How do peers impact learning? An experimental investigation of peer-to-peer teaching and ability tracking.' IZA Discussion Paper No. 10783, Institute of Labor Economics, Bonn.

Lavy, V. (2010). 'Do differences in school's instruction time explain international achievement gaps in math, science, and reading? Evidence from developed and developing countries.' NBER Working Paper 16227, National Bureau of Economic Research.

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.

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

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.

Mendolia, S., Paloyo, AR., Walker, I. (2018). 'Heterogeneous effects of high-school peers on educational outcomes.' *Oxford Economic Papers* (forthcoming, doi: 10.1093/oep/gpy008).

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

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 Woessman, L. (eds.) *Handbook of the Economics of Education Volume 3*, Amsterdam: Elsevier.

Schwerdt, G. and Wuppermann, A.C. (2011). 'Is traditional teaching really all that bad? A within-student between-subject approach.' *Economics of Education Review* 30(2): 365-379.

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

Thompson, G. (2013). 'NAPLAN, MySchool and Accountability: Teacher perceptions of the effects of testing.' *The International Education Journal: Comparative Perspectives* 12(2): 62-84.

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.

Zhang, H. (2016). 'The role of testing noise in the estimation of achievement-based peer effects.' *Economics of Education Review* 54: 113-123.

Figure 1: Sensitivity of estimated new peer effects to top and bottom peer group thresholds – Year 7 results

Table 1: Sample Means & Standard Deviations, Year 7 students in public schools

	All	"Regular students"	At least 1 subject in the top 10%	At least 1 subject in the bottom 10%
A. Student's outcomes				
Y5 percentile rank, grammar	48.458 (28.577)	47.036 (24.163)	75.117 (19.700)	21.594 (19.912)
Y5 percentile rank, reading	48.251 (28.529)	46.367 (23.964)	75.368 (19.792)	21.804 (20.071)
Y5 percentile rank, spelling	48.919 (28.715)	47.244 (23.993)	75.437 (20.914)	22.908 (21.259)
Y5 percentile rank, writing	48.599 (28.518)	47.959 (25.293)	70.879 (23.203)	24.548 (21.215)
Y7 percentile rank, grammar	50.568 (28.621)	49.689 (21.708)	81.222 (16.909)	20.398 (18.782)
Y7 percentile rank, reading	50.710 (28.631)	49.905 (21.599)	80.960 (17.424)	20.358 (19.395)
Y7 percentile rank, spelling	50.455 (28.611)	49.610 (22.003)	79.797 (18.791)	21.586 (20.167)
Y7 percentile rank, writing	50.489 (28.659)	50.484 (22.378)	78.025 (20.521)	21.333 (20.002)
B. Student's characteristics				
Male	0.517 (0.500)	0.491 (0.500)	0.468 (0.499)	0.635 (0.482)
Speaks another language than English at home	0.231 (0.422)	0.209 (0.407)	0.298 (0.457)	0.219 (0.414)
Indigenous	0.018 (0.133)	0.016 (0.124)	0.005 (0.073)	0.036 (0.187)
<i>Mother's occupation</i>				
senior managmnt & qual. professionals	0.084 (0.277)	0.071 (0.257)	0.165 (0.371)	0.032 (0.176)
Other business managers, arts/media/sportspersons and associate professionals	0.144 (0.351)	0.145 (0.352)	0.195 (0.396)	0.090 (0.286)
Tradesmen/women, clerks and skilled office, sales and service staff	0.214 (0.410)	0.229 (0.420)	0.212 (0.409)	0.177 (0.382)
Machine operators, hospitality staff, assistants, labourers and related workers	0.237 (0.425)	0.248 (0.432)	0.182 (0.386)	0.265 (0.441)
Not in paid work in last 12 months	0.306 (0.461)	0.292 (0.454)	0.234 (0.424)	0.418 (0.493)

<i>(continued)</i>	All	"Regular students"	At least 1 subject in the top 10%	At least 1 subject in the bottom 10%
<i>Father's occupation</i>				
senior managmnt & qual. professionals	0.081 (0.273)	0.068 (0.252)	0.166 (0.372)	0.026 (0.159)
Other business managers, arts/media/sportspersons and associate professionals	0.149 (0.357)	0.150 (0.357)	0.211 (0.408)	0.085 (0.278)
Tradesmen/women, clerks and skilled office, sales and service staff	0.204 (0.403)	0.219 (0.414)	0.197 (0.398)	0.173 (0.379)
Machine operators, hospitality staff, assistants, labourers and related workers	0.213 (0.410)	0.219 (0.414)	0.160 (0.367)	0.252 (0.434)
Not in paid work in last 12 months	0.114 (0.318)	0.107 (0.308)	0.094 (0.292)	0.155 (0.362)
<i>Mother's school education</i>				
Year 9 or equivalent or below	0.094 (0.293)	0.088 (0.283)	0.045 (0.207)	0.163 (0.369)
Year 10 or equivalent	0.187 (0.390)	0.193 (0.395)	0.106 (0.308)	0.254 (0.436)
Year 11 or equivalent	0.216 (0.411)	0.235 (0.424)	0.160 (0.366)	0.226 (0.418)
Year 12 or equivalent	0.478 (0.500)	0.460 (0.498)	0.669 (0.470)	0.326 (0.469)
<i>Father's school education</i>				
Year 9 or equivalent or below	0.086 (0.280)	0.083 (0.276)	0.047 (0.212)	0.133 (0.340)
Year 10 or equivalent	0.166 (0.372)	0.178 (0.383)	0.111 (0.314)	0.194 (0.395)
Year 11 or equivalent	0.160 (0.367)	0.173 (0.378)	0.138 (0.345)	0.151 (0.358)
Year 12 or equivalent	0.339 (0.473)	0.318 (0.466)	0.525 (0.499)	0.200 (0.400)
<i>Mother's non school education</i>				
No non school qualification	0.366 (0.482)	0.373 (0.484)	0.248 (0.432)	0.471 (0.499)
Certificate I to IV (incl. trade certificate)	0.264 (0.441)	0.281 (0.449)	0.207 (0.405)	0.280 (0.449)
Advanced diploma/Diploma	0.129 (0.335)	0.133 (0.339)	0.152 (0.359)	0.093 (0.291)
Bachelor degree or above	0.168 (0.374)	0.139 (0.345)	0.340 (0.474)	0.065 (0.246)

<i>(continued)</i>	All	"Regular students"	At least 1 subject in the top 10%	At least 1 subject in the bottom 10%
<i>Father's non school education</i>				
No non school qualification	0.246 (0.431)	0.253 (0.435)	0.179 (0.384)	0.297 (0.457)
Certificate I to IV (incl. trade certificate)	0.258 (0.438)	0.278 (0.448)	0.220 (0.414)	0.248 (0.432)
Advanced diploma/Diploma	0.083 (0.276)	0.084 (0.277)	0.112 (0.316)	0.052 (0.221)
Bachelor degree or above	0.129 (0.335)	0.101 (0.302)	0.283 (0.451)	0.039 (0.194)
Size of peer group (in Y7)	174 (78)	173 (78)	191 (78)	158 (76)
Number of students	118,356	65,525	26,883	25,948

Table 2: Average Peer rankings

	Mean	Std	Min	Max
Percentage of new peers in Y7	82.67	17.71	3.33	99.77
A. Average of new peers' Y5 percentile rank				
in grammar	47.72	9.65	1.07	98.41
in reading	47.52	9.83	1.02	93.41
in spelling	48.19	9.31	1.45	92.10
in writing	47.79	9.47	1.05	88.84
B. Percentage of new peers' Y5 percentile rank in top 10%				
in grammar	8.67	6.10	0	100
in reading	8.50	6.37	0	100
in spelling	9.03	6.47	0	74.06
in writing	8.57	6.31	0	50
C. Percentage of new peers' Y5 percentile rank in bottom 10%				
in grammar	11.28	7.66	0	100
in reading	11.33	7.63	0	100
in spelling	11.16	7.19	0	100
in writing	11.19	7.45	0	100

Table 3: Analysis of variance of peer rankings

	Mean	Overall Std	Between Std	Within Std
Y7 percentile rank	50.56	28.63	24.55	14.92
Y5 percentile rank	48.56	28.59	24.15	15.52
Average of new peers' Y5 percentile rank	47.81	9.57	9.02	3.17
Percentage of new peers' Y5 percentile rank in top 10%	8.69	6.32	5.66	2.78
Percentage of new peers' Y5 percentile rank in bottom 10%	11.24	7.49	6.69	3.39

Table 4: Student fixed effects regression results – New peers and Year 7 achievement

	1a	1b	1c	Specification				
				2	3	4	5	6
Average of new peers' rank								
Coefficient	0.363***	-	-	0.315***	0.267***	0.179***	0.120***	0.081***
se	(0.022)	-	-	(0.026)	(0.025)	(0.021)	(0.019)	(0.021)
% of new peers' rank in the bottom 10%								
Coefficient	-	-0.168***	-	-0.018	-0.019	-0.014	-0.013	-0.005
se	-	(0.019)	-	(0.016)	(0.015)	(0.014)	(0.014)	(0.015)
% of new peers' rank in the top 10%								
Coefficient	-	-	0.280***	0.080***	0.067***	0.042**	0.034*	0.038**
se	-	-	(0.023)	(0.021)	(0.021)	(0.019)	(0.018)	(0.019)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for old peers	No	No	No	No	Yes	Yes	Yes	Yes
Same-subject own prior score*subject dummies	No	No	No	No	No	Yes	Yes	Yes
Cross-subject own prior scores*subject dummies	No	No	No	No	No	No	Yes	Yes
School*subject dummies	No	No	No	No	No	No	No	Yes
Number of observations	461,861	461,861	461,861	461,861	418,821	397,896	393,677	393,677
Number of students	118,356	118,356	118,356	118,356	107,197	102,866	101,663	101,663
Number of schools	330	330	330	330	327	326	326	326

Controls: Dummies for grammar, reading, spelling, and grammar*sex, reading*sex and spelling*sex.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Table 5: Robustness checks – Student fixed effects (Year 7)

	Robustness checks			
	1	2	3	4
Average of new peers' rank				
Coefficient	0.079***	0.058*	0.080***	0.050**
se	(0.021)	(0.035)	(0.018)	(0.024)
% of new peers' rank in the bottom 10%				
Coefficient	-0.008	-0.037	-0.001	-0.049
se	(0.015)	(0.027)	(0.013)	(0.046)
% of new peers' rank in the top 10%				
Coefficient	0.037*	0.091***	0.057***	0.158***
se	(0.019)	(0.034)	(0.016)	(0.060)
Number of observations	383,444	199,130	486,097	397,896

Controls: As in specification 6 in Table 4.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Robustness 1: sample restricted to individuals for whom we observe test scores in all subjects.

Robustness 2: Year 7 in 2010-2011.

Robustness 3: adding numeracy.

Robustness 4: raw test scores.

Table 6: Placebo Tests, Student fixed effects (Year 7):**(1) replacing own Year 7 score with own Year 5 score****(2) replacing subject peer scores with those from random subjects**

	Placebo 1	Placebo 2
Average of new peers' rank		
Coefficient	-0.009	-0.010
se	(0.024)	(0.009)
% of new peers' rank in the bottom 10%		
Coefficient	-0.026*	0.001
se	(0.013)	(0.006)
% of new peers' rank in the top 10%		
Coefficient	0.030	0.000
se	(0.021)	(0.008)
Number of observations	397,896	397,896

Controls: Dummies for grammar, reading, spelling, and grammar*sex, reading*sex and spelling*sex.

Placebo 1: Yr5 outcome, peer variables entered together, controls for old peers but not for own prior score + school*subject dummies

Placebo 2: Random peer values entered for subjects, peer variables entered together, controls for old peers and for cross-subject own prior scores interacted with subject dummies + school*subject dummies

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Table 7: Student fixed effects (Year 7) – by gender and ability

	Overall	By ability (wrt YR5 percentile rank)					
		Rank < 20	Rank 20-35	Rank 35-50	Rank 50-65	Rank 65-80	Rank>=80
<u>All</u>							
Average of new peers' rank							
Coefficient	0.081***	0.082***	0.089***	0.083***	0.083***	0.077***	0.068***
se	(0.020)	(0.020)	(0.021)	(0.020)	(0.021)	(0.022)	(0.021)
% of new peers' rank in the bottom 10%							
Coefficient	-0.005	-0.015	-0.008	0.001	0.004	-0.011	0.002
se	(0.014)	(0.015)	(0.016)	(0.017)	(0.019)	(0.017)	(0.017)
% of new peers' rank in the top 10%							
Coefficient	0.038**	0.041	0.056**	0.054**	0.050**	0.026	0.015
se	(0.018)	(0.025)	(0.027)	(0.027)	(0.025)	(0.025)	(0.023)
Number of observations	393,677				393,677		
<u>Boys</u>							
Average of new peers' rank							
Coefficient	0.074***	0.070***	0.078***	0.075***	0.082***	0.070***	0.057**
se	(0.023)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.025)
% of new peers' rank in the bottom 10%							
Coefficient	-0.007	-0.003	-0.014	-0.005	-0.012	-0.024	0.006
se	(0.017)	(0.019)	(0.021)	(0.023)	(0.025)	(0.023)	(0.025)
% of new peers' rank in the top 10%							
Coefficient	0.028	0.039	0.051	0.034	0.010	0.016	0.015
se	(0.022)	(0.032)	(0.034)	(0.041)	(0.036)	(0.031)	(0.032)
Number of observations	201,566				201,566		

Girls

Average of new peers' rank

Coefficient	0.087***	0.093***	0.095***	0.087***	0.084***	0.084***	0.077***
se	(0.025)	(0.027)	(0.027)	(0.025)	(0.026)	(0.027)	(0.027)
% of new peers' rank in the bottom 10%							
Coefficient	-0.006	-0.034*	-0.008	0.000	0.015	-0.004	0.000
se	(0.017)	(0.020)	(0.022)	(0.023)	(0.023)	(0.022)	(0.020)
% of new peers' rank in the top 10%							
Coefficient	0.051**	0.050	0.072*	0.076**	0.082***	0.033	0.017
se	(0.021)	(0.038)	(0.037)	(0.033)	(0.030)	(0.032)	(0.025)
Number of observations	192,111			192,111			

Controls: As in specification 6 in Table 4.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Table 8: Student fixed effects (Year 7) –by school size and proportion of new peers

	Overall	By quintiles of the % of new peers				
		<Q1	>=Q1 & <Q2	>=Q2 & <Q3	>=Q3 & <Q4	>=Q4
<u>Small schools</u>						
Average of new peers' rank						
Coefficient	0.039*	0.019	0.051*	0.069**	0.069**	0.041
se	(0.023)	(0.024)	(0.029)	(0.030)	(0.033)	(0.039)
% of new peers' rank in the bottom 10%						
Coefficient	-0.008	0.005	-0.028	-0.024	-0.057*	-0.094**
se	(0.014)	(0.015)	(0.025)	(0.029)	(0.032)	(0.043)
% of new peers' rank in the top 10%						
Coefficient	0.036*	0.034	0.054	0.003	0.023	0.103*
se	(0.020)	(0.023)	(0.047)	(0.046)	(0.043)	(0.059)
Number of observations	193,395			193,395		
<u>Large schools</u>						
Average of new peers' rank						
Coefficient	0.282***	0.290***	0.269***	0.316***	0.282***	0.271***
se	(0.062)	(0.060)	(0.065)	(0.065)	(0.068)	(0.071)
% of new peers' rank in the bottom 10%						
Coefficient	-0.019	-0.046	0.012	-0.010	0.012	0.006
se	(0.064)	(0.069)	(0.072)	(0.073)	(0.090)	(0.081)
% of new peers' rank in the top 10%						
Coefficient	0.018	0.019	0.046	0.008	-0.008	0.032
se	(0.052)	(0.062)	(0.058)	(0.064)	(0.063)	(0.065)
Number of observations	200,282			200,282		

Controls: As in specification 6 in Table 4.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Table 9: Student fixed effects – Year 9 achievement

	Specification							
	1a	1b	1c	2	3	4	5	6
Average of new peers' rank								
Coefficient	0.336***	-	-	0.278***	0.233***	0.136***	0.078***	0.068**
se	(0.026)	-	-	(0.032)	(0.031)	(0.028)	(0.026)	(0.031)
% of new peers' rank in the bottom 10%								
Coefficient	-	-0.194***	-	-0.050**	-0.051**	-0.033	-0.028	0.015
se	-	(0.025)	-	(0.024)	(0.024)	(0.022)	(0.021)	(0.021)
% of new peers' rank in the top 10%								
Coefficient	-	-	0.240***	0.068**	0.061**	0.043	0.030	-0.043
se	-	-	(0.033)	(0.029)	(0.029)	(0.029)	(0.027)	(0.031)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for old peers	No	No	No	No	Yes	Yes	Yes	Yes
Same-subject own prior score*subject dummies	No	No	No	No	No	Yes	Yes	Yes
Cross-subject own prior scores*subject dummies	No	No	No	No	No	No	Yes	Yes
School*subject dummies	No	No	No	No	No	No	No	Yes
Number of observations	204,320	204,320	204,320	204,320	186,326	178,123	176,371	176,371
Number of students	52,886	52,886	52,886	52,886	48,179	46,479	45,972	45,972
Number of schools	336	336	336	336	335	334	334	334

Controls: Dummies for grammar, reading, spelling, and grammar*sex, reading*sex and spelling*sex.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Appendix Table 1: OLS (Year 7)

	Specification							
	1a	1b	1c	2	3	4	5	6
Average of new peers' rank								
Coefficient	0.470***	-	-	0.404***	0.288***	0.153***	0.135***	0.133***
se	(0.015)	-	-	(0.024)	(0.022)	(0.018)	(0.017)	(0.019)
% of new peers' rank in the bottom 10%								
Coefficient	-	-0.425***	-	0.016	0.006	-0.001	0.000	0.000
se	-	(0.030)	-	(0.024)	(0.021)	(0.016)	(0.015)	(0.016)
% of new peers' rank in the top 10%								
Coefficient	-	-	0.627***	0.140***	0.085***	0.055***	0.037**	0.044**
se	-	-	(0.024)	(0.022)	(0.021)	(0.018)	(0.017)	(0.019)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for old peers	No	No	No	No	Yes	Yes	Yes	Yes
Same-subject own prior score*subject dummies	No	No	No	No	No	Yes	Yes	Yes
Cross-subject own prior scores*subject dummies	No	No	No	No	No	No	Yes	Yes
School*subject dummies	No	No	No	No	No	No	No	Yes
Number of observations	461,861	461,861	461,861	461,861	418,821	397,896	393,677	393,677
Number of students	118,356	118,356	118,356	118,356	107,197	102,866	101,663	101,663
Number of schools	330	330	330	330	327	326	326	326

Controls: Dummies for grammar, reading, spelling, and grammar*sex, reading*sex and spelling*sex. Additional controls included in the OLS specifications are: sex, language spoken at home, indigenous status, mother's and father's occupation, mother's and father's school & non-school education. Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.

Appendix Table 2: Student fixed effects (Year 7) – sample restricted to specification 6 sample from Table 4

	Specification							
	1a	1b	1c	2	3	4	5	6
Average of new peers' rank								
Coefficient	0.359***	-	-	0.309***	0.265***	0.179***	0.120***	0.081***
se	(0.027)	-	-	(0.030)	(0.027)	(0.022)	(0.019)	(0.021)
% of new peers' rank in the bottom 10%								
Coefficient	-	-0.171***	-	-0.025	-0.023	-0.013	-0.013	-0.005
se	-	(0.024)	-	(0.018)	(0.017)	(0.015)	(0.014)	(0.015)
% of new peers' rank in the top 10%								
Coefficient	-	-	0.273***	0.077***	0.069***	0.043**	0.034*	0.038**
se	-	-	(0.029)	(0.024)	(0.023)	(0.021)	(0.018)	(0.019)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for old peers	No	No	No	No	Yes	Yes	Yes	Yes
Same-subject own prior score*subject dummies	No	No	No	No	No	Yes	Yes	Yes
Cross-subject own prior scores*subject dummies	No	No	No	No	No	No	Yes	Yes
School*subject dummies	No	No	No	No	No	No	No	Yes
Number of observations	393,677	393,677	393,677	393,677	393,677	393,677	393,677	393,677
Number of students	101,663	101,663	101,663	101,663	101,663	101,663	101,663	101,663
Number of schools	326	326	326	326	326	326	326	326

Controls: Dummies for grammar, reading, spelling, and grammar*sex, reading*sex and spelling*sex.

Robust standard errors clustered at the school level. ***/**/* denotes statistical significance at the 1%/5%/10% levels.