THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 11/17

Secondary School Teacher Effects on Student Achievement in Australian Schools

Chris Ryan

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Secondary School Teacher Effects on Student Achievement in Australian Schools*

Chris Ryan
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 11/17
May 2017

* The research in this paper was commissioned by the Victorian Department of Education and Training (DET). The findings and views in this paper are those of the author and should not be attributed to the Department of Education and Training. All errors are responsibility of the author. Contact: <ryan.c@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* +61 3 8344 2100**
***Fax* +61 3 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This study finds that approaching 10% of the variation in high school student achievement is explained by teacher effects in Australia. It uses data from the 2011 Trends in Mathematics and Science Study (TIMSS) sample of Australian Year 8 students to estimate achievement in mathematics and science with student fixed effects, calculating teacher effects as part of this estimation. Like results in other studies, these teacher effects do not appear to be strongly related to observed teacher characteristics, despite attempts to account for the composition of the classes teachers face. Nor are the teacher effects related to self-assessments of how well prepared teachers view themselves as being able to teach the content of the TIMSS tests.

JEL classification: I21

Keywords: Teacher effects, teacher characteristics, class composition

1. Introduction

International studies show that differences in teacher performance can have a substantial impact on student achievement. Estimates suggest that a one standard deviation improvement in teacher performance is associated with an increase in student achievement by 0.1 of a standard deviation in the current year (Leigh 2010; Rockoff 2004; Rivkin, Hanushek and Kain 2005; Slater *et al*. 2012 find larger effects for England). So, teacher performance matters. However, these estimated teacher effects seem to have little apparent relationship with the characteristics of teachers typically observed by researchers in international studies (Hanushek and Rivkin 2006; Slater, Davies & Burgess 2012).

There is, nevertheless, uncertainty about the applicability of these results for Australia and uncertainty surrounds at least some of the effects associated with teacher characteristics, most notably the role of experience. While the results may generalise for many of the readily observable characteristics of teachers, the impacts of recent experience of subject-relevant professional development and of subject-specific self-confidence about expertise, for example, are less clear.

There are many complexities in estimating teacher effects. These include accounting for timing issues, including the duration with the current teacher, difficulties in disentangling the contributions of teachers where classes or student careers in individual subjects are split between teachers and the role of measurement error. Further confounding any assessment of teacher performance from student achievement is that teachers are not randomly assigned students – some selection process assigns specific teachers to groups of students. Schools may choose to assign particularly effective teachers to specific groups of students, depending on what they are trying to achieve - for example, if they aim to have as many high achieving students as possible (with a consequent large variance in student achievement) compared with as few low achieving students as possible (with a smaller variance in achievement), their staff allocation decisions will differ.

The formal analysis in the paper involves the estimation of regression equations where there are two estimates of achievement per student, one in mathematics and one in science. Hence identification of effects will be based on factors that differ between students' science and mathematics classes that are associated with differences in their achievement between the subjects (an approach followed in Dee 2005, 2007, Ammermüller and Dolton 2006, Clotfelter, Ladd & Vigdor 2007, Slater *et al*. 2012, among others). Much of the descriptive analysis will also involve comparisons of individual or class average differences in science and mathematics achievement and their relationship with differences in their experiences with the subjects. Not all students in the same mathematics class have the same science teachers, and *vice versa*. Consequently, it is possible to estimate equations with both student and teacher fixed effects, providing the number of students from the classes is large enough (Meltzer and Woessmann 2010). Student-level differences in prior achievement or in their experience of quality teaching between mathematics and science will be reflected in idiosyncratic, individual-level residuals that will not be picked up in the estimated teacher effect, other than through class composition effects, discussed below.

Because the data also include teacher surveys which elicit information on teacher background, qualifications, approaches to teaching, professional development and matters such as whether they were specifically trained for the mathematics or science class they teach, it is possible to regress the estimated teacher effects on these characteristics to assess how much they affect teacher performance (an approach followed in many studies, for example Slater *et al*. 2012).

A final important issue is how to deal with the "composition" of the class which may make it easier or harder to teach, an important part of the allocation of teachers to classes (Clotfelter, *et al.* 2007). The estimated teacher effect is likely to consist of two elements – the "true" contribution from the teacher and an element that reflects class composition. Ideally, measures of prior achievement could assist in dealing with this issue (Kalogrides and Loeb 2013), but they are not available in the data used here, from the Trends in Mathematics and Science Study (TIMSS). There are some sources of information about class composition available in TIMSS. These include an indication from students about their intended level of educational attainment and responses from teachers about the number of students from non-English speaking backgrounds in their classes, and whether they face difficulties in teaching the class because of the poor grasp of pre-requisite knowledge or skills, disruptive and/or uninterested students and the presence of students with special learning needs. It will also be possible to regress these measures reflecting sorting against the teacher characteristics to provide a picture of student sorting in early secondary school within Australian schools.

The paper uses the 2011 Trends in Mathematics and Science Study (TIMSS) data sample of Australian Year 8 students. This data contains student achievement scales in mathematics and science, along with questionnaire responses from the mathematics and science teachers of those students that can be linked to their achievement.

The paper begins with a review of three relevant literatures about teachers: the magnitude of estimated teacher effects; their impact on achievement and what factors this depends on; and how they are assigned to groups of students across and within schools. The methodology section discusses the problems associated with estimation of teacher effects, particular issues with the data and our approach to this study. We present our analysis in the following section and finish with the conclusion.

2. The literature on teacher effects

There are three related elements of the teacher literature of relevance to this study – the size of any teacher effects on student achievement, the relationship between teacher effects and their characteristics and the allocation of teachers to classes of students.

2.1. The size of teacher effects

As already noted, there is a general consensus that a one standard deviation improvement in teacher performance is associated with an increase in student achievement of 0.1 of a standard deviation in the current year (Leigh 2010; Rockoff 2004; Rivkin, Hanushek and Kain 2005; Meltzer and Woessmann 2010) However, Slater *et al*. 2012 find effects for England that are more than twice this magnitude. Nye, Konstantonopoulos and Hedges (2004), cited in Hattie (2009), summarise the education literature as indicating that somewhere between 7 and 21 per cent of the variance in student achievement gains was associated with variation in teacher effectiveness.

2.2. Teacher effects and teacher characteristics

In what follows, we will distinguish between readily-observed teacher characteristics, such as age and gender, and harder to observe characteristics such as teachers' own achievement, their educational practices and their attitudes. Hanushek and Rivkin (2006) summarise the literature on the measured impact of the readily-observed teacher characteristics on outcomes of various kinds. While mostly performance in achievement tests has been analysed in studies, other outcomes studied include attainment and future earnings. They report that few reliable studies find any effect for higher educational qualification levels, notably a masters-level degree on student outcomes, while only a minority of studies find a positive experience effect, citing summaries of the literature in Hanushek (1997, 2003). However, the experience effect is highly non-linear, with positive effects with the only the first few years of experience (Rivkin, Hanushek & Kain 2005, Hanushek, Kain, O'Brien & Rivkin 2005). Leigh (2010) found that student achievement increased with teacher experience for up to 20 years, though these are not based on a panel that follows the same teachers over time. Winters *et al.* (2012) found student reading declined for students with teachers beyond 4-6 years experience, but was constant for students in numeracy with teachers beyond 4-6 years experience.

With data for England, Slater *et al*. (2012) found no teacher characteristics included in their regression equation (among age, experience, gender, qualification level, field of qualification and salary level) helped explain student achievement.

Among other factors that might be readily-observed about teachers, studies typically find no role for remuneration packages in explaining student achievement (Hanushek and Rivkin 2006; Slater *et al*. 2012). The impact of certification procedures, where teachers must reach some set of minimum certification standards, also remains contentious (Wayne and Youngs 2003). There are studies that find subject-specific certification is related to achievement in relevant subjects. For example, Goldhaber and Brewer (2000) and Jepsen and Rivkin (2002) found positive certification effects on student mathematics achievement and/or its growth.

By contrast, Hattie (2009) concludes teacher subject matter knowledge has only a small impact on student achievement. These results are related to the issue of teachers teaching in fields in which they are qualified to teach. Another factor that may contribute to subject matter knowledge is subject-related experience of professional development. Hattie (2009) summarises the literature on professional development as indicating that it can change teacher learning or knowledge, but has less effect on actual teacher behavior and still less effect on student learning,

Teachers own achievement in specific subjects will also reflect their subject matter knowledge, but is harder to measure. Studies of its impact tend to point to positive effects of teacher achievement on student achievement, though the results are much stronger for subject-specific achievement that relates directly to the test being undertaken by the students (Wayne and Youngs 2003). Meltzer and Woessmann (2010) studied the effect of teacher subject knowledge on student achievement using within-teacher within-student variation, exploiting data where both students and their teachers were tested in two subjects. They found that a one standard deviation in subject-specific teacher achievement increased student achievement by about 10 percent of a standard deviation.

Hattie (2009) suggests that other aspects of teachers' attitudes, behaviours and philosophies might be related to better student achievement outcomes. These include teacher expectations of students, their relationships with students, their adoption of specific learning approaches and student assessments of the teacher's performance. Hattie concludes that it is teachers skilled in specific teaching methods, with high expectations for their students and who create positive relationships with students who are likely to have above average effects on student achievement.

This view involves a shift in emphasis from studying who teachers are to what they do as the factors likely to make the main contribution to their impact on student achievement. Unfortunately, these sets of attitudes and behaviours are harder and more complex to observe than simple demographic characteristics. Bietenbeck (2014) used an earlier version of the TIMSS study to look at how different teaching practices affected achievement. Bietenbeck (2014) distinguished 'traditional' and 'modern' teaching practices. He found that traditional teaching practices fostered factual knowledge and competency in solving routine problems, skills that have traditionally been emphasized in schools (and achievement tests). In contrast, modern teaching practices promote reasoning, but that only a small fraction of the questions in standardized tests, both in TIMSS and elsewhere, measured students' reasoning skills.

One likely important attitudinal characteristic of teachers is their belief that they can do the job, or their self-efficacy about their skills and capabilities. Self-efficacy refers to people's appraisal of their capabilities to execute courses of action that are context specific (Bandura 1977, 1997). Self-efficacy has to do with self-perceptions of competence, rather than an individual's actual level of competence. For teaching, this refers to teachers' confidence in their ability to promote students' learning, even those who might be difficult to teach or be otherwise unmotivated (Guskey & Passaro 1994, Tschannen-Moran, Woolfolk Hoy & Hoy 1998).

Studies in the education literature have looked at the measurement, determinants and impact of teacher self-efficacy, albeit measured in different ways from the approach used here. Two main approaches have been used, but both involve analysing answers to direct questions of teachers in surveys. The first approach involves the use of two questions, about the teacher's sense of being able to influence the outcomes or learning of their students, despite the motivation level of students. The second approach distinguishes an "efficacy" expectation (that an individual can execute the actions necessary to accomplish a specific task at a specific level) from an "outcomes" expectation (the consequences of performing the task at the specified level). Scales based on this approach are longer and distinguish these two elements. Bandura (1997) developed a 30-item teacher self-efficacy scale, one that goes beyond classroom learning into their capacity to influence school-level decisions, parent and other external interactions and other factors.

Bandura's critique of earlier attempts to capture teacher self-efficacy argued that measures need to capture a teacher's assessment of their competence across the wide range of activities and tasks they are asked to perform. The level of specificity of the measure should reflect the purposes of the research. In this paper, where the purpose of the research is to analyse student achievement, the measure of self-efficacy should be based on a teacher's confidence about teaching the content of the set of assessable items actually contained in the TIMSS test. Such an approach is described in the data section below.

2.3. The Allocation of teachers to classes of students

Since much of the literature that looks at how teachers are allocated to students comes from the United States, it has tended to focus on issues of interest specific to that country, such as the racial segregation of minority-background and White students and the match of such classes to teachers (Mickelson 2002). Sorting is found to occur both *across* or *between* schools and *within* schools, being greater *across* schools than *within* them, but the extent of *within* school sorting increases at high school compared to lower levels of schooling (Clotfelter, Ladd & Vigdor 2002, Conger 2005). African-American students are more segregated from White students than are Hispanic students. Both explicit and *de facto* ability tracking within schools is one of the primary mechanisms that also drives *within* school racial and social-background segregation (Oakes & Guiton 1995, Lucas & Berends 2002). Schools with tracking systems in place tend to group students on the basis of prior school achievement (Conger, Long & Iterola 2009), though other factors may impinge on this allocation, including parental representations (Oakes & Guiton 1995).

Parents also intervene to ensure their children are taught by specific teachers, a practice more common among those from more advantaged social backgrounds (Lareau 1987, 2000). If acted on, this could act to match more capable teachers (if parents are well informed) to students from more advantaged social backgrounds. Alternatively, teacher retention practices within schools may allow the "best" teachers to choose the classes they wish to be assigned to, which could result in such teachers being assigned classes somewhat easier to teach.

Kalogrides & Loeb (2013) found evidence of substantial sorting by social background and race in a study of student sorting *within* schools in three school districts in the United States, predominantly driven by sorting based on student achievement. Further, Kalogrides & Loeb (2013) found that classes with the most low-achieving, minority and poor students tended to be taught by "novice" teachers, which they equate to such classes being taught by "low quality" teachers.

3. Methodology and data

The principal focus in this paper is on the relationship between estimated teacher effects on student achievement and the observed characteristics of the teachers. This relationship is estimated via a two stage approach, as in Slater *et al*. (2012), for example. In the first stage, student achievement in the form of ranked science and mathematics achievement is regressed on student characteristics and their experience of the two subjects to derive measures of the teacher effects that contribute to achievement across relevant groups of students. The approach incorporates fixed student effects to account for factors like family background, individual ability, motivation and educational and occupational ambitions that might influence achievement in both science and mathematics. In the second stage, the teacher effects are regressed against teacher characteristics observed in the data. At either the first or the second stage, we can incorporate measures of the composition of classes to attempt to remove the impact of these factors from the estimated teacher effects. In that way, we hope to remove from the estimated teacher effect that part of class achievement that reflects that the class was particularly easy or challenging to teach. In the process of estimating these two stages, there are a number of supplementary phenomena we are interested in to round out the picture of how teachers might influence achievement. These include how teachers view their own performance and how that might be related to their characteristics and whether the characteristics of teachers are related to how they are assigned to classes.

3.1. Relationships estimated in this paper

We are interested in exploring a number of relationships in this research. The first just relates student achievement in different subjects to their characteristics and effort levels in the subject (the first stage estimation). The grouping of students into separate classes for mathematics and science allows estimation of a class effect that equates to a teacher effect in cases where teachers teach only one class in the data used here. The second relationship is between average class achievement and factors that reflect the composition of the class, either its social composition or its planned educational attainment or a teacher assessment that it is a difficult class to teach. This shows how student achievement is affected by the class composition variables. The third type of relationship is between these class composition variables and the characteristics of teachers. These show how teacher are matched to classes of students. The fourth relationship of interest relates how confident teachers are about their skills to teach the specific topics covered in the TIMSS tests to their background characteristics. The fifth relationship looks at whether the teacher's skills self-efficacy is related to average class achievement. The sixth relationship (the second stage regression estimation) looks at how the teacher effect is related to the teacher's background characteristics, taking account of the difficulty of teaching their particular class and their own skills self-efficacy.[1]

The exact regression equations use to estimate these relationships are set out more formally in Appendix I of this paper. The results for each of these relationships are presented in order below in the results section of the paper.

3.2. Data

The study uses Australian data from the Trends in International Mathematics and Science Study (TIMSS) for 2011 (described in more detail in Thomson *et al*. 2012). TIMSS involves an assessment of the mathematical and scientific achievement of students in Years 4 and 8, though the focus here is on the Year 8 sample. For the 2011 survey, a stratified random sample of 277 secondary Australian schools (43 in Victoria) participated in the data collection for the TIMSS 2011 Year 8 study. The sample was stratified by state, school sector, SEIFA (socio-economic status) index and rural geography. The initial national sample sizes of the Australian school and student data used here are set out in Table 1. Some 7556 students were included in the TIMSS sample.

Two classes of mathematics students per school were sampled in the TIMSS survey in the Australian Capital Territory and the Northern Territory, along with all of the indigenous students found in Year 8 in the sampled schools. In other states, the sample was not based on intact classes. In general, students were distributed across more science than mathematics classes in schools. These students were linked to 802 mathematics classes and 740 teachers, and 1049 science classes and 902 teachers.

---

[1] Following the literature, the estimated teacher effects are "shrunk" to minimise the role of estimates of teacher effects based on few student observations, as described in Koedel, Mihaly and Rockoff (2015).

The samples used in estimation depart from the entire sample because of item non-response by students and questionnaire and item non-response by teachers. Further, for the analysis of classes and/or teachers, a minimum class size of five students is imposed to provide some comfort about data quality. The observations lost are set out in Table 2. Rather than the 1028 teachers that teach the students in the TIMSS data in mathematics and science, just 303 observations are used for the analysis of teacher characteristics. Some check on the potential impact on the data quality of the lost observations can be provided by assessing whether the teachers used in the characteristics analysis typically have different teacher effects on student achievement from those who are not used. The two distributions appear in Figure 1 and are practically identical, so in outcomes at least, the teachers used in the main analysis appear representative of the teachers of this sample.

The data contain three types of variables used in this study: individual student characteristics and attitudes; teacher characteristics, attitudes and behaviours; and variables that reflect the match of students to teachers, some of which are aggregations of the student characteristics, while others are reports by teachers about their classes.

3.2.1 Student characteristics

The principal student level variables used in this analysis are the mathematics and science achievement of students, based on the TIMSS tests. The outcomes of the achievement tests in mathematics and science appear in Table 3. Australian students performed relatively better in science than they did in mathematics compared to the international average of 500 in both tests in TIMSS 2011. Since the values of the achievement tests have no real cardinal interpretation, comparisons of scores across the domains has little meaning. For example, a higher score in science compared to mathematics does not necessarily convey that the student is better in or knows more science than mathematics. In this paper, ranks of students within the two subjects are used, normalised so the lowest ranked student is assigned a value of 0 in the two subjects, the highest ranked student 1. With ranks, a student who has a higher rank in science than mathematics clearly is better than more students in science than mathematics, so a difference in ranks does convey information about the student’s relative performance in the subjects. Like other international studies, the TIMSS data contains five “plausible” values for achievement in both mathematics and science. The five alterative values for achievement reflect that there is a degree of uncertainty surrounding any individual's true achievement based on the specific test they undertook. In this study, we use individuals' ranks within mathematics and science of the average of their five plausible values to measure their achievement.[2]

Students in TIMSS report their attitudes towards mathematics and science (the *likes*, *confident*, *values* mathematics and science scales used in TIMSS, along with the *engaged* by mathematics and science scales). These scales are used in some regression specifications in this paper and their average values are presented in Table 3. For the *confident* and *values* scales, Year 8 students have more positive attitudes towards mathematics than science, but are more engaged by their science classes.

Another measure that differs between science and mathematics is the weekly hours of homework the student completes in both subjects, calculated from the number of times they complete homework each week and the hours they typically undertake each time they do it. Students report doing more hours of mathematics than science homework. In the regression estimation that follows, this variable is converted to capture something more like a relative homework effort concept, by dividing the student hours in mathematics (or science) by the average number of hours of homework by other students at their school. The idea is to capture the time students spend doing homework in the specific subjects relative to their peers.

A number of other student demographic characteristics are also utilised in this paper, though their values do not differ between science and mathematics. These include scales intended to capture the social background of students and their conscientiousness. The first is based on responses about the presence of (potentially) education-related household items (a computer, a study desk, and an internet connection, the presences of the student's own books, among other items), while the second was calculated as the student's completion rate of the last fifty-four items asked of them in their questionnaire, mostly about their attitudes towards mathematics and science (the *likes*, *confident*, *values* and *engaged* in mathematics and science scales used in TIMSS).

[2] Using the rank of the average score does not induce the same statistical problems as using the average of the plausible values, since the distribution of the rank of the average is the same as the rank of any of the individual plausible values.

Summary statistics for these variables and other individual characteristics that do not differ between science and mathematics appear in Table 3 (and their definitions in Appendix B). The other variables include revealed planned student attainment that is either "low" (incomplete secondary school) or "high" (a university qualification), gender and the books in the household. These are presented entirely for descriptive purposes and the student-level variables are not analysed further in this paper, since they do not differ across mathematics and science and therefore cannot explain any of the difference in achievement between them. Their average values, at the class level, are however used and described in more detail in the next sub-section

3.2.3 Teacher characteristics

The descriptive statistics for the set of teacher characteristics, behaviours and attitudes are set out in Table 4, for teachers in mathematics and science. These reflect the teacher's gender, age, experience, highest qualifications, mathematics and science-related qualifications, the homework they set, their recent exposure to professional development opportunities and their confidence about their preparation to teach the material in the TIMSS test, referred to as teacher skills self-efficacy. In general, the various measures differ very little between responding mathematics and science teachers. The age and experience profiles of teachers are very similar. The proportion of teachers with subject-specific qualifications is substantially higher in science than in mathematics, though again this partly reflects the limits on the sample arising from class size restrictions that work remove disproportionately science teachers without relevant qualifications.

The teacher skills self-efficacy measure was constructed from two sets of information. First, in their surveys, teachers indicated that a specific topic covered in the relevant TIMSS exam was in taught in Year 8 at their school. Second, teachers indicated whether they personally were "very well" (=3), "somewhat" (=2) or "not well" (=1) prepared to teach the topic, with their aggregate skills self-efficacy measured as the sum of these scores for all topics taught in Year 8 divided by three. Hence, the value of one in this scale means that overwhelmingly a teacher views themselves as being very well prepared to teach all topics in the TIMSS test that are relevant to their grade level in their school. Around two thirds of teachers had a score of one for this measure (see Table 5). The balance indicate lesser levels of skills self-efficacy. In general, the pattern suggests that typically few people rate themselves as “not well prepared” for many topics.

The expectation is that teachers with higher skills self-efficacy are likely to do a better job at conveying material than others, if no reason other than that their effort and motivation will be higher if they believe themselves capable of doing a good job. The distribution of responses is examined, along with teacher characteristics associated with their skills self-efficacy and the simple relationship between it and achievement in Figure 3. The distribution of teacher skills self-efficacy is presented in the top left panel of Figure 3 and shows the distribution skewed towards the maximum value of one.

The right panel at the top of Figure 2 summarises the apparent relationship between skills self-efficacy and average class achievement in mathematics and science. Over the range to 0.9, where there are relatively few observations, average class achievement in both mathematics and science increases with teacher skills self-efficacy. However, average class achievement in mathematics and science is lower where teacher’s skills self-efficacy measure is equal to one. This is also evident in Table 5 with shows the class average achievement ranks for specified ranges of the skills self-efficacy measure.

The lower panel of Figure 2 relates the difference in the self-efficacy by students’ science and mathematics teachers to the difference in their achievement between the subjects, which mirrors the regression estimation approach used in the paper. The dashed line shows the distribution of values the difference in teacher self-efficacy takes, while the solid line shows how this difference is associated with the difference in achievement ranks between science and mathematics. The lines indicate that, over the range where the overwhelming majority of the differences in teacher self-efficacy occur (from -0.1 to 0.4), the relationship with achievement might be positive, but any positive relationship is not a strong one. Teacher self-ratings of their capability may not mirror their impact on the achievement of students very well, though it remains possible that class composition and matching effects obscure the true relationship, something which the regression equations discussed later can take into account.

3.2.2 Class characteristics

Aggregations of these variables across classes can, however, be used to explain difference in science and mathematics achievement, since the peers in a student's mathematics and science classes differ, so a student can be in a mathematics class with more peers who hope to attend university than in their science class. Summary statistics for variables that reflect the match of students to teachers in mathematics and science are presented in Table 6. In addition to the class averages based on expected attainment, class averages of the home possessions and student conscientiousness variables are used in this paper, as well as the gender compositions of classes (leaving aside single sex schools, in the peer effects literature, classes with a higher the proportion of boys in them are viewed as being prone to more disruption). Further, a number of teacher responses about the classes are used to indicate aspects of the match of students to teachers. These responses reflect class size and the extent to which the teacher indicates that either some factor limits generally their ability to teach the class (student characteristics, disabilities, nutrition, attitudes and behaviours) or specifically that disruptive or uninterested students limit their ability to teach the class.

In general, the science classes appear slightly more favourable for achievement than the mathematics classes. This reflects the pattern of teacher responses and limits on the sample arising from class size restrictions, which exclude more science classes than mathematics classes, making the former a more select group.

The set of class composition and match variables are explored in more detail in Figures 3 though 9. These follow a similar format to Figure 2, in that the top left panel in each Figure shows the distribution of the composition variable, the right panel the simple relationship between the variable and the average class achievement rank in science and mathematics and the lower panel the relationship between the difference in the composition variable between science and mathematics and the difference in student achievement ranks in science and mathematics.

The left-hand panel of each Figure should inform interpretation of the right-hand panel, since it shows where the observations may be plentiful or scarce. In order, Figures 3 through 9 show the panels for class average SES, the proportion who plan to attend to university, the class proportion who plan to leave education before completing Year 12, class size, the student conscientiousness scale, the teacher identified learning limitation and the proportion of the grade cohort who are male.

In all cases, the top right panels of the Figures conform to expectations. Achievement increases with average class SES; with the proportion of the class who plan to attend to university; falls with the proportion who plan to leave school early; mostly rises with class size up to about classes with 35 students, but there are very few classes with more than 35 students; rises with the class average of valid responses to the attitudes questions, the conscientiousness scale; falls as the extent of limitations to teaching the class rises; increases with the familiarity with the content of the TIMSS test; and is higher in all single-sex schools than in mixed gender schools, though single sex schools are more often private, higher SES schools.

The lower panels also tend to support the expected directions of relationships: the difference in achievement ranks between subjects tends to increase with differences in class SES over the relevant range, as it does with higher average educational attainment expectations and class conscientiousness, and falls with larger differences in identified student problems. However, the relationships are not obviously strong, with most of the apparent differences in achievement ranks over the relevant range of the composition difference variables being very small.

Descriptive statistics for one further variable is included in Table 6. It indicates how much of the content of the TIMSS test had been covered by Year 8 in schools, based on teacher reports (higher values correspond with the content having been taught before Year 8). The variable is considerably higher in mathematics than in science, suggesting that more of the mathematics content of the TIMSS test should be familiar to students than the science test. Figure 10 follows the same format as the preceding set of figures and shows the distribution of this content timing variable and its relationship with achievement. Once more, while the apparent relationship with achievement rank appears quite strong, the relationship between differences in achievement ranks between subjects and differences in the timing of the TIMSS content is not obviously strongly positive.

3.2.4 School-level factors

Various distinctions between mathematics and science were also drawn in the school questionnaire. These included remuneration bonuses for teachers in the subjects, recruitment difficulties, and evaluation practices for staff performance. Generally, these phenomena did not differ within schools between mathematics and science teachers in Australia in the TIMSS data. For example, if a school used financial incentives for mathematics teachers, they overwhelmingly also used them for science teachers too. Further, schools tended to use observed student achievement, teacher peer review and observations of performance by senior staff to evaluate the performance of teachers, but not the use of external persons or inspectors. However, they used the same practices for the evaluation of teachers in both mathematics and science, so that those practices cannot be used to explain any differences in performance between the subjects.

By contrast, thirty percent of schools had different experiences in the recruitment of teachers between mathematics and science. However, few (less than half) of those who found it difficult to recruit for vacancies in mathematics and science put it down to a shortage or inadequacy in the availability of teachers with a specialization in mathematics or science. Factor analysis of the shortage question returned factors that indicated that the availability of computers, library materials and the inadequacy of physical facilities mattered more for schools' capacity to provide instruction than teacher recruitment issues. Consequently, factors at the school level are not analysed specifically in the analysis that follows, though some of the regression specification include school fixed effects to remove their impact on other relationships of interest.

4. Results

The presentation of results follows the same order as the list of relationships of interest from section 3.1, beginning with the teacher skills self-efficacy measure.

4.1. Student achievement

In this section, the role of student-related factors that influence their achievement rank in science and mathematics is examined. The principal purpose of estimation of the equation is just to derive the teacher effects for subsequent analysis. Only factors that differ between the two subjects can be considered given the fixed student effects approach adopted here. The results of three different specifications are reported in Table 7. In the first specification (column 1), just the content timing and student homework variables (and an indicator for mathematics) are included as the explanatory variables. In the second specification (column 2), attitudinal variables towards mathematics and science are added to the set of explanatory variables. In the third specification (column 3), a number of the class composition variables are added to the second specification, to assess whether taking account of them at this stage influences the estimated relationships between the estimated teacher effects and teacher characteristics, discussed later.

Not surprisingly, content timing is related to student achievement rank, with students who should be more familiar with the TIMSS test content being ranked higher. Relative homework effort does not appear to influence relative achievement ranks for students. From the second specification, positive attitudes towards subjects in terms of the *like* and *confidence* scale are associated with higher achievement ranks, though we do not emphasise features of this specification too much. First, it is unclear which way the causation runs between attitudes and achievement, and second, it is possible that one of the ways teachers affect student achievement is through their attitudes towards their subjects. Hence, these variables may remove the role of an important channel from the estimated teacher effects.

The third specification suggests that the class composition variables contribute very little to individual student achievement in the two subjects. The home possessions (SES) variable and existence of any factors that limit the teachers' efforts to teach the class effects are both not significantly different from zero in the third column. Neither were other class composition variables significant when they were included in the regressions.

One key issue is how much of the variance in student achievement is explained by variation in the teacher effects. The variance in student achievement ranks is 0.08, while the variance in teacher effects is 0.005 in specification 1 and 0.006 in specification 3. This indicates that under10 per cent of the variation in student achievement is explained by the teacher effects in this case. A one standard deviation change in the teacher effect is associated with about a 0.15 of a standard deviation change in student achievement, a result between those of Slater *et al*. (2012) for England and the smaller estimated effects of around 0.1 of a standard deviation in student achievement reported for the United States.

4.2. Class composition and average achievement

The relationship between the class composition variables and student achievement has already been described non-parametrically, for each separate variable, in the series of Figures 3 through 9 in the top right and lower panels. The results when the equation is estimated jointly for all of the variables, focussing on the difference between the science and mathematics class composition on the difference in achievement rank appear in Table 8. Like the results from the third specification of the previous sub-section, the results suggest that the relationship between the class composition variables and achievement is fairly weak. If estimated over individual observations (column 2), the only variables that are significant are of the wrong sign. The first column contains estimates using differences in class-level means, which provide more encouraging results, but unfortunately this is not the level at which the teacher effects are estimated.

4.3. Class composition and the match of students to teachers

In this sub-section, the way students are matched to teachers is studied. The approach taken is to regress each of the match or class composition variables against teacher characteristics, including teacher skill self-efficacy, to see whether there are any systematic patterns in the way teachers are matched with students.

The regression relationships, which relate each of these match variables for classes with at least 5 teachers to teacher characteristics appear in Table 9. These show how students are allocated to teacher within schools, not how teachers are allocated across schools.[3] That is, they show whether it is young teachers in schools who are allocated the lowest SES class or the class with the low aspirations about attending university within schools, not whether young teachers are allocated to the lowest SES schools.

[3] The equations reflect the *within* school matching because they are estimated with school fixed effects.

The set of explanatory teacher characteristics include the gender, age, experience and qualifications of teachers, along with the subject, whether their qualifications relate to science of mathematics, their recent experience of professional development and their skills self-efficacy.

In general, the results show little regularity. While it appears the youngest group of teachers tend to be assigned classes with the lowest levels of expected educational attainment, in most equations just one of the variables is significant. Consequently, it seems unsafe to make too much of the results. Males appear to be assigned classes with more high SES students, as do teachers with higher levels of skills self-efficacy.

4.4. Teachers views of their skills and their characteristics

In this sub-section, the focus is on the reports by teachers about how well prepared they believe themselves to be to teach the content of the relevant TIMSS mathematics or science tests and their own characteristics.

The relationship between teachers skills self-efficacy and their personal characteristics is presented in the last column of Table 9, which contains the results of a regression equation of skills self-efficacy against a set of teacher characteristics such as gender, age, experience, qualifications and their relevance to the teacher's field, recent experience of professional development programs and whether they teach mathematics or science. Teacher skills self-efficacy appears to increase with age, experience, recent experience of professional development programs and is higher among mathematics teachers than science teachers, though having a relevant science qualification largely offsets the last effect. Teachers who tend to set more weekly hours of homework also exhibit higher skills self-efficacy.

4.5. Teachers skills self-efficacy and average class achievement

In this sub-section, the focus is on reports by teachers about how well prepared they believe themselves to be to teach the content of the relevant TIMSS mathematics or science tests and whether this is related to average class achievement. From the bottom panel of Figure 10, we might anticipate that any positive relationship might not be large, though that Figure did not take into account any class composition or match effects.

The relationship between teachers' skills self-efficacy and average class achievement is presented in the last column of Table 8. The estimated parameters indicate what the impact of differences in skills self-efficacy between science and mathematics teachers has any impact on differences in the achievement ranks of students between science and mathematics. The explanatory variables include the teacher skills self-efficacy variable and the set of class composition or match variables. The results indicate that skills self-efficacy does not help explain individual student achievement, since the parameter on this variable is not significantly different from zero.

4.6. The relationship between teacher effects and characteristics

We turn now to the second stage regression estimation that looks at how the teacher effects are related to the teacher's background characteristics, taking account of the difficulty of teaching their particular class and their own skills self-efficacy. We know from the student achievement equations that the teacher effects are important, explaining about a quarter of the variance in student achievement. Now, we wish to ascertain whether these effects are systematically related to observed teacher characteristics. The results, for the specifications associated with the three specifications in Table 7, appear in Table 11. The teacher characteristic variables included in the regression equations include a set of teacher characteristics such as gender, age, experience, qualifications and their relevance to the teacher's field, recent experience of professional development programs, whether they teach mathematics or science and the amount of homework they expect students in their classes to undertake each week.

In general, very few of the included characteristics are important in explaining the teacher effects. The do not appear to be any regularities to the significance of any variable, suggesting that the estimated teacher effects are not related to any of the teacher characteristics observed in this study.

5. Conclusions

This study has confirmed the importance of teacher effects on student achievement. In Australia, teachers indeed have a substantial impact on student achievement. Equally, the study has confirmed that not much about the teacher effects in Australia can be explained by the observed characteristics of teachers. It seems likely that what teachers *do* is much more important than who they *are*. It is also possible the potentially observable factors, such as aspects of teachers' personalities and their individual traits may be associated with the estimated teacher effects.

One of the implications of the study is that teacher self-assessments of how well-prepared they are to teach particular topics may not be reliable. Students sat tests that teachers, by and large, believed they were well-prepared to teach the content of, and those assessments bore no relation to the student outcomes. Clearly, only objective measures, or informed and reflective, peer assessments of teacher performance should form the basis of any appraisal of teacher performance.

Another implication of this study is that providing rewards on the basis of teacher characteristics may not involve actually rewarding performance, or incentivising the workforce. If performance is not strongly associated with characteristics, providing incentives to teachers on the basis of these characteristics will reward teachers independently of their performance. This concern is relevant to programs that aim to increase the proportion of teachers with relevant qualifications for the areas or disciplines in which they currently teach.

There are some uncertainties surrounding the results in this study. The class composition variables seemed promising in some specifications, yet were not significantly different from zero in the equations that mattered. That some of the teacher effects were subject to measurement error, since the number of students attached to teachers was so small, must mean that some of the class characteristics and teacher effects were measured with substantial error. Measurement error is known to bias towards zero parameter estimates in liner models. This problem does not affect estimation of the relationship between teacher effects and characteristics – here an estimated zero relationship really means there is no relationship. Finally, the results relate to Year 8 students and teachers. In senior school, it is possible that qualifications relevant to subject content and experience teaching the material may be more closely related to teacher performance and hence to teacher effects.

References

Ammermüller, A & Dolton P (2006), *Pupil-Teacher Gender Interaction Effects on Scholastic Outcomes in England and the USA*, ZEW Discussion Paper 06-060, Mannheim: Centre for European Economic Research.

Bandura, A (1977), 'Self-efficacy: Towards a Unifying Theory of Behavioral Change', *Psychological Review*, 84, 191 – 215.

Bandura, A (1997), *Self-efficacy: The Exercise of Control*, New York: WH Freeman.

Bietenbeck, J (2014), Teaching Practices and Cognitive Skills, *Labour Economics*, 30, 143-153.

Clotfelter, CT, Ladd, HF & Vigdor, JL. (2002), 'Segregation and resegregation in North Carolina's public school classrooms', *North Carolina Law Review*, *81*, 1463–1512.

Clotfelter, CT, Ladd, HF & Vigdor, JL. (2006), 'Teacher-student matching and the assessment of teacher effectiveness', *Journal of Human Resources*, Vol. 41, 778–820.

Clotfelter, CT, Ladd, HF & Vigdor, JL (2007), Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects. NBER Working Paper 13617. Cambridge, MA: National Bureau of Economic Research.

Conger, D (2005), 'Within-school segregation in an urban school district', *Educational Evaluation and Policy Analysis*, *27*, 225–244.

Conger, D, Long, MC, & Iatarola, P (2009), 'Explaining race, poverty, and gender disparities in advanced course-taking', *Journal of Policy Analysis and Management*, *28*, 555–576.

Dee, TS (2005), A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review* 95 (2), 158–165.

Dee, TS (2007), Teachers and the Gender Gaps in Student Achievement. *Journal of Human Resources* 42 (3), 528–554.

Goldhaber, D.D., Brewer, D.J. (2000), 'Does teacher certification matter? High school teacher certification status and student achievement', *Educational Evaluation and Policy Analysis*, 22 (2), 129–145.

Gusskey, TR & Passaro, PD (1994), 'Teacher efficacy: A study of Construct Dimensions', *American Educational Research Journal*, 31, 439 – 453.

Hanushek, EA (1997), Assessing the Effects of School Resources on Student Performance: An Update, *Educational Evaluation and Policy Analysis,* 19 (2), 141–164.

Hanushek, EA (2003), 'The failure of input-based schooling policies', *Economic Journal*, 113 (485), F64 – F98.

Hanushek, E.A., Kain, J.F., O'Brien, D.M., Rivkin, S.G. (2005), *The market for teacher quality*, Working Paper 11154, National Bureau of Economic Research, Cambridge.

Hanushek, EA & Rivkin, SG (2011), Teacher Quality, in Hanushek, EA, and Welch F (eds), *Handbook of the Economics of Education*, Vol II, North-Holland, Amsterdam.

Hattie, JAC (2009), *Visible Learning*, United Kingdom: Routledge.

Jepsen, C., Rivkin, S.G. (2002), *What is the trade-off between smaller classes and teacher quality*?, National Bureau of Economic Research.

Kalogrides, D & Loeb, S (2013), 'Different Teachers, Different Peers: The Magnitude of Student Sorting Within Schools', *Educational Researcher*, 42, 6, 304– 316.

Koedel, C Mihaly, K Rockoff, JE (2015), 'Value-added modeling: A review', *Economics of Education Review*, 47, 180–195.

Lareau, A (1987), 'Social class differences in family-school relationships: The importance of cultural capital', *Sociology of Education*, *60*, 73–85.

Lareau, A (2000), *Home advantage: Social class and parental intervention in elementary education*, Lanham, MD: Rowman & Littlefield.

Leigh, A (2010), 'Estimating teacher effectiveness from two-year changes in students' test scores', *Economics of Education Review*, 29, 480 – 88

Lucas, SR, & Berends, M (2002), 'Sociodemographic diversity, correlated achievement, and de facto tracking', *Sociology of Education*, *75*, 328–348.

Mickelson, RA (2002), *The academic consequences of desegregation and segregation: Evidence from the Charlotte-Mecklenburg Schools*. Available at http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/the-academic-consequences-of-desegregation-and-segregation-evidence-from-thecharlotte-mecklenburg-schools

Oakes, J & Guiton, G (1995), 'Matchmaking: The dynamics of high school tracking decisions', *American Educational Research Journal*, *32*, 3–33.

Rivkin, SG, Hanushek, EA, & Kain, JF (2005), 'Teachers, schools, and academic achievement', *Econometrica*, 73(2), 417–458.

Rockoff, JE (2004), 'The impact of individual teachers on student achievement: Evidence from panel data', *American Economic Review*, 94(2), 247–252.

Slater, H, Davies, NM, & Burgess, S, (2012), 'Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England', *Oxford Bulletin of Economics and Statistics*, 74, 5, 629 – 645.

Thomson S, Hillman, K & Wernert, N (2012), *Monitoring Australian Year 8 student achievement internationally: TIMSS 2011*, Australian Council for Educational Research, Camberwell.

Tschannen-Moran, M, Woolfolk Hoy, A & Hoy, WK (1998), 'Teacher Efficacy: Its Meaning and Measure', *Review of Educational Research*, 68, 202 – 248.

Wayne, AJ & Youngs, P (2003), 'Teacher Characteristics and Student Achievement Gains: A Review', *Review of Educational Research*, 73 (1), 89–122.

Winters, MA, Dixon, BL and Greene, JP, (2012), 'Observed characteristics and teacher quality: Impacts of sample selection on a value added model', *Economics of Education Review* 31, 19– 32

Table 1: 2011 TIMSS Year 8 data dimensions

Original TIMSS Year 8 sample	
Number of students	7556
Number of schools	277
Number of teacher – class pairings in the sample	
mathematics	802
science	1049
Number of linked teachers	
mathematics	740
science	902
Number of teacher respondents	
mathematics	486
science	579

Table 2: 2011 TIMSS Year 8 data – derivation of analysis sample

Original TIMSS Year 8 sample	7556	
Missing one of mathematics or science achievement		0
Missing one of the student-level explanatory variables		770
Missing teacher survey response		2225
Observations used in the student achievement equation	4561	
Missing one of the teacher-level explanatory variables		181
Excluded from teacher analysis because of small class size		339
Observations used in the teacher fixed effect equation	384	

Table 3 Student-level characteristics

	Mean	Std Dev		
Male (%)	50	50		
Books: 1 - 20 (%)	10	30		
Books: 11 - 25 (%)	16	37		
Books: 26 - 100 (%)	31	46		
Books: 101 - 200 (%)	20	40		
Books: 201 or more (%)	22	41		
Home possessions scale (0-100)	88	15		
Conscientiousness (%)	98	11		
Expected schooling <Year 12 (%)	23	42		
Expected education >=degree (%)	32	47		
	Mathematics		Science	
	Mean	Std Dev	Mean	Std Dev
Achievement (1st plausible value)	503	84	518	84
Hours of weekly homework	1.5	1.6	0.8	1.0
Likes	9.3	1.9	9.3	2.1
Confident	10.2	2.1	9.8	1.9
Values	9.9	1.8	9.1	2.1
Engaged by	9.3	1.8	9.5	1.9

Source: Estimated from TIMSS 2011, weighted data.

Table 4 Teacher characteristics

	Mathematics		Science	
	Mean	Std Dev	Mean	Std Dev
Male (%)	46	50	44	50
Age: <25 (%)	3	17	5	21
Age: 25 - 29 (%)	17	38	14	35
Age: 30 - 39 (%)	28	45	29	46
Age: 40 - 49 (%)	22	42	24	43
Age: 50 - 59 (%)	25	44	24	43
Age: >=60 (%)	4	19	4	19
Experience (years)	15	11	14	11
Post-graduate qualification (%)	62	49	79	41
Qualified in field (%)	65	48	89	31
Professional development last 2 years – any (%)	91	28	93	25
Professional development last 2 years – subject specific (%)	88	33	88	32
Professional development last 2 years – individual student needs (%)	72	45	69	46
Teacher self-efficacy	98	8	95	10
Class size	23.9	6.3	25.0	5.4
Weekly homework hours set	1.3	1.0	0.8	0.6

Source: Estimated from TIMSS 2011 data.

Table 5 Teacher self-efficacy and student achievement

	% Distribution	Average achievement rank
Self-efficacy range: <0.8	6.6	42
Self-efficacy range: 0.8 to <0.9	8.6	50
Self-efficacy range: 0.9 to <1	18.0	57
Self-efficacy =1	66.7	52

Source: Estimated from TIMSS 2011 data.

Table 6 Class composition or match variables

	Mathematics		Science	
	Mean	Std Dev	Mean	Std Dev
Expected schooling <Year 12 (%)	23	13	23	27
Expected education >=degree (%)	30	21	34	27
Conscientiousness (%)	98	4	98	5
Home possessions scale (0-100)	88	7	89	7
Male share (%)	50	24	52	30
Any limitation to teaching class	-0.04	0.84	-0.10	0.78
Class teaching limited by disruptive student behaviour	82	39	86	35
Class size	24	6	25	5
Teacher observations	247		278	
Content timing	68	9	59	8

Source: Estimated from TIMSS 2011 data.

Table 7 **Student achievement regression with individual fixed effects**

	FEs no attitudes /match	FEs include attitudes, no match	FEs include match
Hours of weekly homework	0.005 (0.007)	-0.007 (0.006)	-0.008 (0.007)
Mathematics	-0.006 (0.010)	-0.019** (0.009)	-0.019** (0.009)
Content timing	0.218*** (0.076)	0.243*** (0.067)	0.260*** (0.071)
Like scale		0.014*** (0.001)	0.015*** (0.001)
Confidence scale		0.018*** (0.001)	0.019*** (0.001)
Limitation to teaching class			-0.007 (0.010)
Home possessions			-0.343 (0.220)
R^2	0.18	0.40	0.40
N	9,121	8,899	8,683

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 8 Achievement differences and class composition

	Class average obs	Individual obs	Individual obs, with self-efficacy
Expected schooling <Year 12 (%)	-0.006 (0.031)	0.016 (0.026)	0.018 (0.027)
Expected education >=degree (%)	0.138*** (0.030)	-0.016 (0.025)	-0.021 (0.026)
Conscientiousness	0.301*** (0.111)	-0.029 (0.070)	-0.029 (0.069)
Home possessions	0.372*** (0.076)	-0.119** (0.059)	-0.094 (0.060)
Limitation to teaching class	-0.021*** (0.005)	0.010*** (0.003)	0.009*** (0.003)
Class size	0.001* (0.001)	-0.000 (0.001)	-0.000 (0.001)
Self-efficacy	0.068** (0.035)		0.012 (0.021)
Constant	-0.010** (0.005)	-0.007** (0.003)	-0.008*** (0.003)
R^2	0.15	0.01	0.01
N	637	3,969	3,763

Notes: * p<0.1; ** p<0.05; *** p<0.01; Columns (2) and (3) estimated with standard errors clustered at the teacher level.

Table 9 Class composition and match with teacher characteristics

	Expect low	Expect high	Conscientious ness	SES	Learning limited	Class size	Self-efficacy
Male	0.007	-0.009	0.003	0.010**	0.016	-0.168	-0.014*
	(0.014)	(0.016)	(0.003)	(0.005)	(0.096)	(0.564)	(0.008)
Aged 25-29	-0.090**	0.082*	-0.010	-0.011	0.235	-0.326	-0.014
	(0.041)	(0.045)	(0.009)	(0.014)	(0.273)	(1.583)	(0.022)
Aged 30-39	-0.019	0.004	-0.009	-0.003	-0.222	-1.359	0.008
	(0.039)	(0.042)	(0.008)	(0.014)	(0.258)	(1.502)	(0.020)
Aged 40-49	-0.031	0.024	-0.007	-0.006	-0.162	-1.276	0.001
	(0.042)	(0.046)	(0.009)	(0.015)	(0.282)	(1.635)	(0.022)
Aged 50-59	-0.083	0.060	-0.011	-0.001	-0.335	-2.932	0.013
	(0.050)	(0.054)	(0.011)	(0.018)	(0.338)	(1.938)	(0.025)
Aged >=60	-0.063	0.014	-0.011	0.013	-0.507	-3.638	0.037
	(0.065)	(0.071)	(0.014)	(0.023)	(0.439)	(2.505)	(0.033)
Experience	0.000	0.000	0.000	-0.000	0.008	0.028	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.009)	(0.050)	(0.001)
High quals	0.014	-0.019	0.001	-0.002	0.044	-0.049	0.020**
	(0.016)	(0.017)	(0.003)	(0.006)	(0.106)	(0.608)	(0.009)
Maths qual	0.012	-0.026	-0.004	-0.005	-0.138	-0.439	0.014
	(0.021)	(0.023)	(0.005)	(0.007)	(0.141)	(0.834)	(0.012)
Science qual	0.024	-0.023	-0.005	-0.007	-0.217*	-0.053	0.055***
	(0.019)	(0.020)	(0.004)	(0.007)	(0.125)	(0.713)	(0.016)
Homework hours	-0.001	0.017	0.001	0.003	-0.228***	-0.113	0.011**
	(0.010)	(0.011)	(0.002)	(0.004)	(0.068)	(0.403)	(0.005)
Prof Devl-ment	0.019	-0.004	0.015**	-0.009	0.062	-0.906	0.068***
	(0.029)	(0.032)	(0.006)	(0.010)	(0.198)	(1.181)	(0.014)
Self-efficacy	-0.056	0.110	-0.043***	0.056**	0.143	0.540	
	(0.076)	(0.082)	(0.016)	(0.027)	(0.507)	(2.917)	
Science							-0.067***
							(0.018)
R^2	0.76	0.89	0.87	0.85	0.64	0.73	0.16
N	482	482	482	482	471	475	482

Notes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$; Classes with at least five students observed. Specification for columns (1)-(6) included school fixed effects.

Table 10 Teacher effects regression explained by teacher characteristics

	FEs exclude match	FEs include attitudes	FEs include match
Male	-0.005	-0.022	-0.001
	(0.012)	(0.014)	(0.013)
Age: 30 - 39	-0.012	-0.023	0.017
	(0.018)	(0.021)	(0.021)
Age: 40 - 49	-0.007	-0.033	0.020
	(0.021)	(0.025)	(0.024)
Age: 50 - 59	-0.004	-0.021	0.003
	(0.027)	(0.032)	(0.031)
Age: >=60	0.009	0.009	-0.039
	(0.042)	(0.049)	(0.051)
Experience	0.000	0.000	0.002*
	(0.001)	(0.001)	(0.001)
Post-graduate qualification	0.006	-0.002	-0.012
	(0.012)	(0.015)	(0.015)
Qualified in field	-0.011	0.013	-0.015
	(0.014)	(0.018)	(0.016)
Professional development	0.022	0.012	0.041
	(0.028)	(0.029)	(0.027)
Self-efficacy	-0.169**	-0.112	-0.066
	(0.074)	(0.080)	(0.086)
Limitation to teaching class			0.011
			(0.008)
Home possessions			-0.032
			(0.065)
_cons	0.157**	0.109	0.042
	(0.076)	(0.080)	(0.102)
R^2	0.02	0.02	0.04
N	384	383	354

* p<0.1; ** p<0.05; *** p<0.01

Appendix A: Regression methodology used in the paper

The various relationships listed in the previous sub-section are specified, in order, in the regression equations that follow in this sub-section. In what follows, let *i* index students, *j* index subjects, *c* index classes, *t* index teachers and *s* index schools.

(1) The individual student achievement equation

Let the achievement rank of individual student *i* in subject *j,* who attends school *s,* r_{ijs} be explained by characteristics common across subjects, denoted by X_i , factors that differ between science and mathematics, denoted by Z_{ij}, the impact of their teachers in science and mathematics, denoted by f_{jt}, an individual effect common across both subjects α_i, a fixed school effect θ_s and a random error term ε^r_{ijs}. The estimation equation for outcome r_{ijs} can be written as:

$$r_{ijs} = \alpha^r_{0j} + Z'_{ij}\delta + X'_i\pi + f_{jt} + \alpha_i + \theta_s + \varepsilon^r_{ijs} \quad \text{or} \qquad (1)$$

$$r_{ij=Ss} - r_{ij=Ms} = \left(\alpha^r_{0j=S} - \alpha^r_{0j=M}\right) + (Z'_{ij=S} - Z'_{ij=M})\,\delta + f_{j=St} - f_{j=Mt} + \varepsilon^r_{ij=Ss} - \varepsilon^r_{ij=Ms}$$

where α^r_{0j}, δ and π are parameters or parameter vectors. The second specification just involves taking differences between subjects *S* and *M*. This has the effect of removing the impact of subject-invariant factors such as family background and general ability, any other fixed individual factors and school fixed effects. This equation is estimated over all observations of students with achievement measures in the two subjects and valid observations of Z_{ij} for each subject.

(2) The relationship between average class achievement and the class composition variables

Let the average achievement rank of students in subject *j,* who attend classes with teacher *t,* $\bar{r}_{jt}$ be explained by the characteristics of students who attend that class, denoted by $\bar{M}_{jt}$, the impact of their teachers in science or mathematics, denoted by ω_{jt}, and a random error term $\varepsilon_{jt}^{\bar{r}}$. The estimation equation for outcome $\bar{r}_{jt}$ can be written as:

$$\bar{r}_{jt} = \alpha_0^{\bar{r}} + \bar{M}_{jt}'\tau + \omega_{jt} + \varepsilon_{jt}^{\bar{r}} \qquad (3)$$

This equation is estimated for each class of students with valid observations of the class compositions variables $\bar{M}_{jt}$ for the relevant subject.

(3) The relationship between the match of students (or class composition variables) and teacher characteristics

Let be the average characteristics of students who attend classes with teacher *t* at school *s,* $\bar{M}_{ts}$ (otherwise known as the class composition measure), be related to the characteristics of teachers, denoted by W_t , a school effect, denoted by ω_s, and a random error term ε_{ijs}^{M}. The estimation equation for outcome $\bar{M}_{ts}$ can be written as:

$$\bar{M}_{ts} = \alpha_0^{M} + W_t'\,\tau + \omega_s + \varepsilon_{ijs}^{M} \qquad (3)$$

This equation is estimated for each class of students with valid observations of the class compositions variables $\bar{M}_{jt}$ for the relevant subject matched to individual teachers who have provided valid values for their own characteristics, denoted by W_t .

(4) The relationship between the teacher's skills self-efficacy and their own characteristics

Let be the skills self-efficacy of teacher *t,* P_t , be related to their characteristics, denoted by W_t , and a random error term ε_{ijs}^{P}. The estimation equation for outcome P_t can be written as:

$$P_t = \alpha_0^{P} + W_t'\,\tau + \varepsilon_{ijs}^{P} \qquad (3)$$

This equation is estimated for teacher who provided valid observations of their skills self-efficacy P_t for the relevant subject for their own characteristics, W_t .

(5) The relationship between average class achievement and class composition

In addition to the average characteristics of students who attend specific classes, let the average achievement rank of students in subject *j,* who attend classes with teacher *t,* $\bar{r}_{jt}$ be explained by the teachers skills self-efficacy of teacher *t,* P_t , a school fixed effect, denoted by ω_s, and a random error term $\varepsilon_{js}^{\bar{r}}$. The estimation equation for outcome $\bar{r}_{jt}$ can be written as:

$$\bar{r}_{jt} = \alpha_0^{\bar{r}} + \bar{M}_j'\tau + \mu P_t + \omega_t + \varepsilon_{jt}^{\bar{r}} \qquad (3)$$

This equation is estimated for each class of students with valid observations of the class compositions variables $\bar{M}_{jt}$ for the relevant subject where a teacher also provided their skills self-efficacy P_t .

(6) The relationship between the teacher's effect and their characteristics, taking account of the class composition variables and the teacher's skills self-efficacy

Let the teacher effect estimated in equation (1), f_{jt}, be explained by the characteristics of students who attend that class, denoted by $\bar{\mathrm{M}}_{\mathrm{jt}}$, the teachers won characteristics, denoted by W_t , the teachers skills self-efficacy of teachers', P_t , and a random error term $\varepsilon_{ijs}^{\hat{f}}$. The estimation equation for outcome $\hat{f}_{jt}$ can be written as

$$\hat{f}_{jt} = \alpha_0^M + W_j'\gamma + \bar{M}_j'\varphi + \vartheta P_t + \varepsilon_{ijs}^{\hat{f}} \qquad (3)$$

This equation is estimated for each class of students with valid observations of the class compositions variables $\bar{M}_{jt}$ for the relevant subject, where a teacher also provided their skills self-efficacy P_t and valid values for their own characteristics, W_t . Following the literature, the estimated teacher effects are "shrunk" to minimise the role of estimates of teacher effects based on few student observations, as described in Koedel, Mihaly and Rockoff (2015).

Appendix Table B.1: Data and variable description

Student characteristics	Variable Description	Mean	Std dev
Male	A dummy variable taking the value 1 if the individual was Male (=1)	50	50
Books at home: 1-10	Reported books in the students home: 1 – 10 (=1).	10	30
Books at home: 11-25	Reported books in the students home: 11 – 25 (=1).	16	37
Books at home: 26-100	Reported books in the students home: 26 – 100 (=1).	31	46
Books at home: 101-200	Reported books in the students home: 101 – 200 (=1).	20	40
Books at home: 200 or more	Reported books in the students home: 200 or more (=1).	22	41
Home possessions scale	Number of study resources (desk to study on, computer, internet connection, own books) student indicates are present in their household (scaled to range from 0 -100)	88	15
Conscientiousness	Proportion of the 50 individual items regarding attitudes towards mathematics and science the student answered (items lie behind the *likes*, *confidence*, *values* and *engaged* scales) (variable ranges from 0 -100)	98	11
Expected attainment - incomplete high school	Student's self-reported expected highest level of education: less than Year 12 (=1)	23	42
Expected attainment - university	Student's self-reported expected highest level of education: compete a university qualification 12 (=1)	32	47
Mathematics achievement	First plausible value for mathematics achievement - TIMSS variable: *bsmmat01*.	503	84
Science achievement	First plausible value for science achievement. - TIMSS variable: *bsssci01*	518	84

Likes mathematics	TIMSS variable: *bsbgslm*, based on 5 items about whether they enjoy mathematic or would prefer not to have to study it	9.3	1.9
Likes science	TIMSS variable: bsbgsl based on 5 items about whether they enjoy science or would prefer not to have to study it s	9.3	2.1
Confidence about mathematics	TIMSS variable: *bsbgscm*, based on 9 items about how well they do in mathematics, including relative to other subjects	10.2	2.1
Confidence about science	TIMSS variable: *bsbgscm*, based on 9 items about how well they do in science, including relative to other subjects s	9.8	1.9
Values mathematics	TIMSS variable: *bsbgsvm*, based on 6 items about its value for future study and employment	9.9	1.8
Values science	TIMSS variable: *bsbgsvs,* based on 6 items about its value for future study and employment	9.1	2.1
Engaged by mathematics classes	TIMSS variable: *bsbgseml*, based on 5 items about classes and their maths teacher's ability to engage them in the class content	9.3	1.8
Engaged by science classes	TIMSS variable: *bsbgsesl*, based on 5 items about classes and their science teacher's ability to engage them in the class content	9.5	1.9
Weekly hours of maths homework	Number of times undertook homework each week times typical hours usually undertaken	1.5	1.6
Weekly hours of science homework	Number of times undertook homework each week times typical hours usually undertaken	0.8	1.0

(continued . . .)

Appendix Table B.1: Data and variable description (continued)

Teacher characteristics	Variable Description	Mathematics		Science	
		Mean	Std dev	Mean	Std dev
Male	A variable taking the value 1 if the teacher was male (=1), 0 if female	46	50	44	50
Teacher aged ≤25	Teacher aged under 25 years (=1)	3	17	5	21
Teacher aged 25-29	Teacher aged 25 – 29 years (=1)	17	38	14	35
Teacher aged 30-39	Teacher aged 30 – 39 years (=1)	28	45	29	46
Teacher aged 40-49	Teacher aged 40 – 49 years (=1)	22	42	24	43
Teacher aged 50-59	Teacher aged 50 – 59 years (=1)	25	44	24	43
Teacher aged ≥60	Teacher aged 60 years or more (=1)	4	19	4	19
Teaching experience	Years spent teaching at the end of current school year	15	11	14	11
Teaching qualifications high	Completed more than one university level qualification (=1)	62	49	79	41
Teaching qualifications in mathematics	Completed a post-secondary qualification with the main area of study in Mathematics, or Education – Mathematics (=1)	65	48		
Teaching qualifications in science	Completed a post-secondary qualification with the main area of study in Biology, Physics, Chemistry, Earth Sciences, or Education – Science (=1)			89	31
Teaching in field	Has a qualification directly relevant to current teaching of mathematics or science	65	48	89	31
Weekly homework hours set	Number of hours of homework typically set	1.3	1.0	0.8	0.6
Professional development	Participated in professional development in past two years on science/mathematics content, pedagogy, curriculum, integration of	91	28	93	25

	technology, assessment and improving students skills or addressing their needs				
Self-efficacy	Scale based on sum of how many of 20 (19) items in mathematics (science) teacher indicated they felt *very well prepared* (=3), *somewhat prepared* (=2) or *not well prepared* (=1)	98	8	95	10
Test content timing	Scale based on sum of how many of 20 (19) items in TIMSS mathematics (science) test the teacher indicated were mostly taught *before this year* (=3), *this year* (=2) or *not yet taught or just introduced* (=1)	68	9	59	8

Appendix Table B.1: Data and variable description (continued)

Student-teacher match variables	Variable Description	Mathematics		Science	
		Mean	Std dev	Mean	Std dev
Proportion of class with low attainment expectations	Proportion of class whose expected highest level of education: less than Year 12	23	13	23	27
Proportion of class with university attainment expectations	Proportion of class whose expected highest level of education: compete a university qualification 12 (=1)	30	21	34	27
Class average conscientiousness	Class average of the individual Conscientiousness scale.	98	4	98	5
Class average home possessions	Class average of the individual home possessions scale.	88	7	89	7
Class average male	Proportion of class who were male	50	24	52	30
Teacher limited by class characteristics	First factor from 6 responses of factors that limited how teacher could teach class reflecting student skills, motivation, nutrition and behaviour	-0.04	0.84	-0.10	0.78
Teacher limited by class behaviour	Teacher indicated they were limited in how they could teach class by student motivation and behaviour	82	39	86	35
Class size	Report from teacher on how many students in class taught.	24	6	25	5

Figure 1: Teacher effects by survey response: science and mathematics

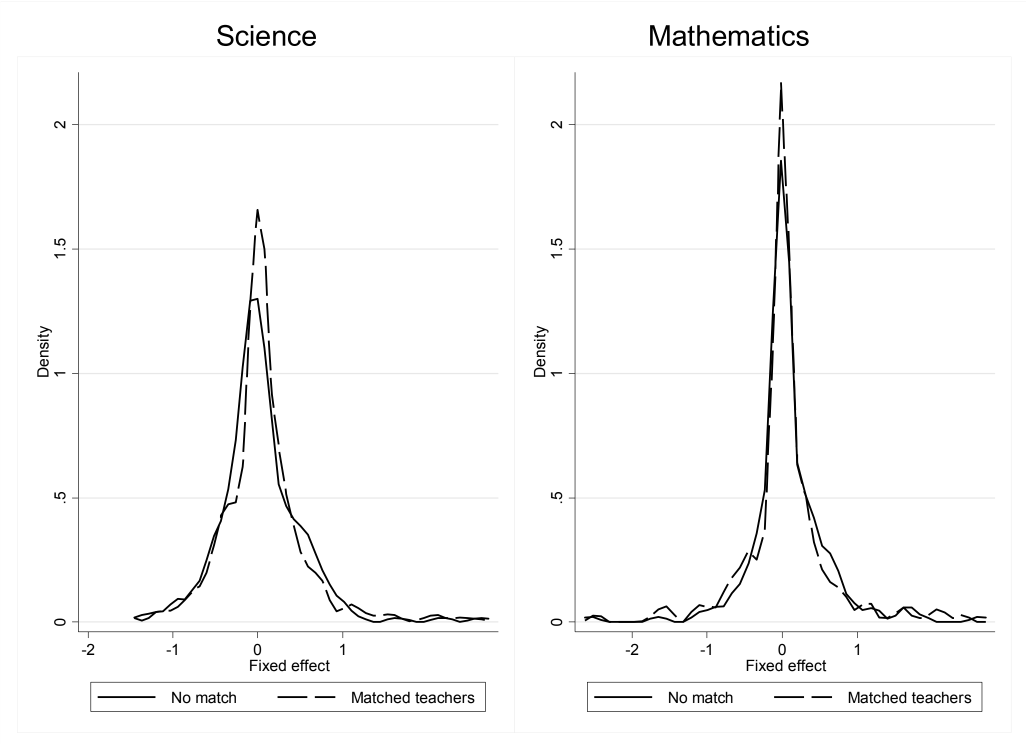
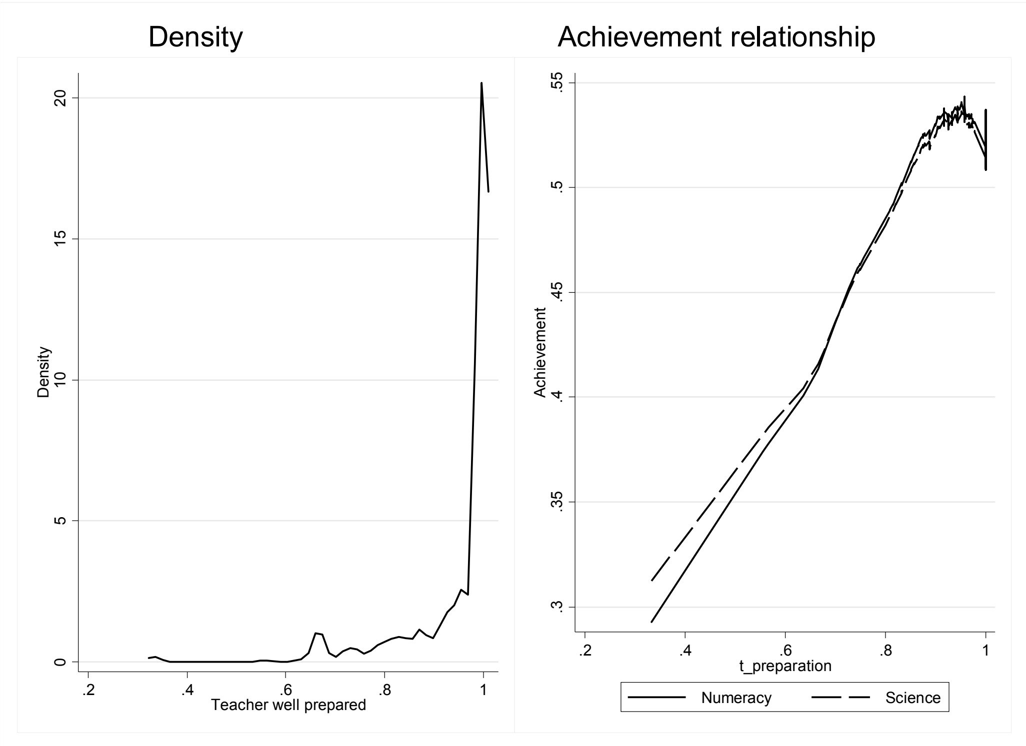


Figure 2: Teacher skills self-efficacy scale and relationship with achievement



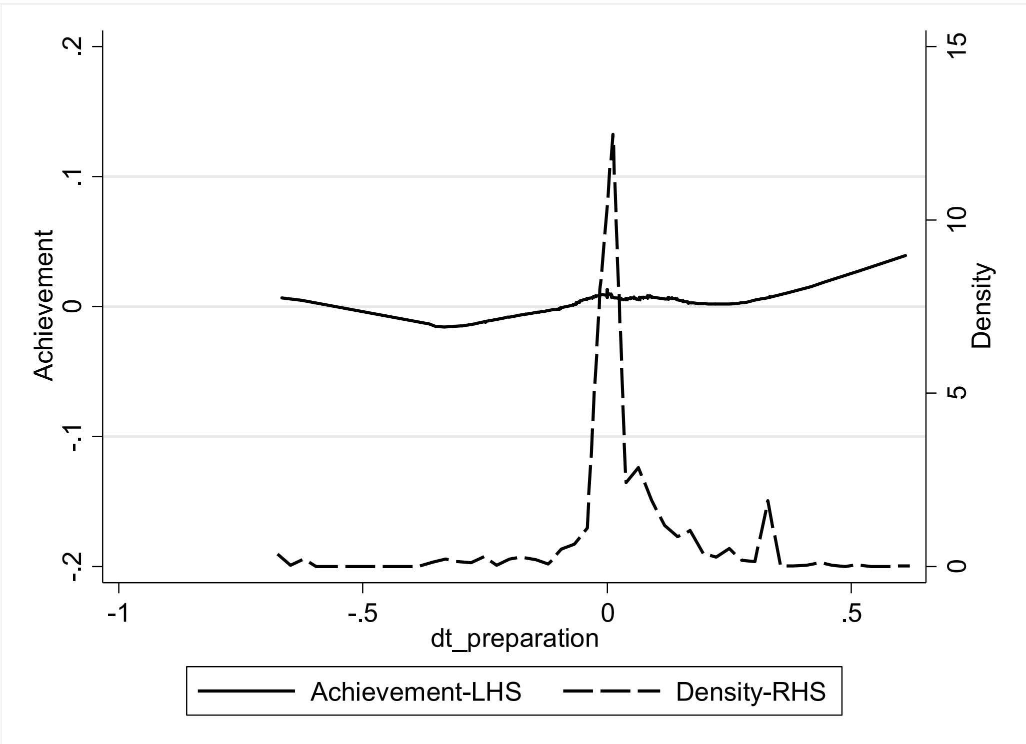
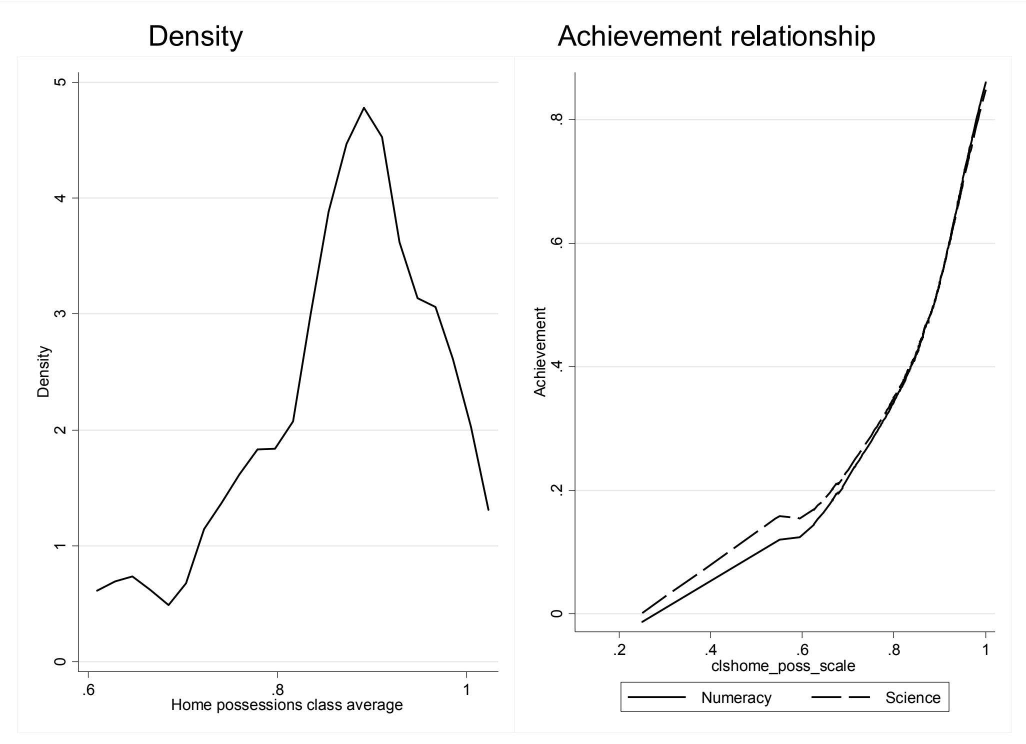


Figure 3: Average class SES and relationship with achievement



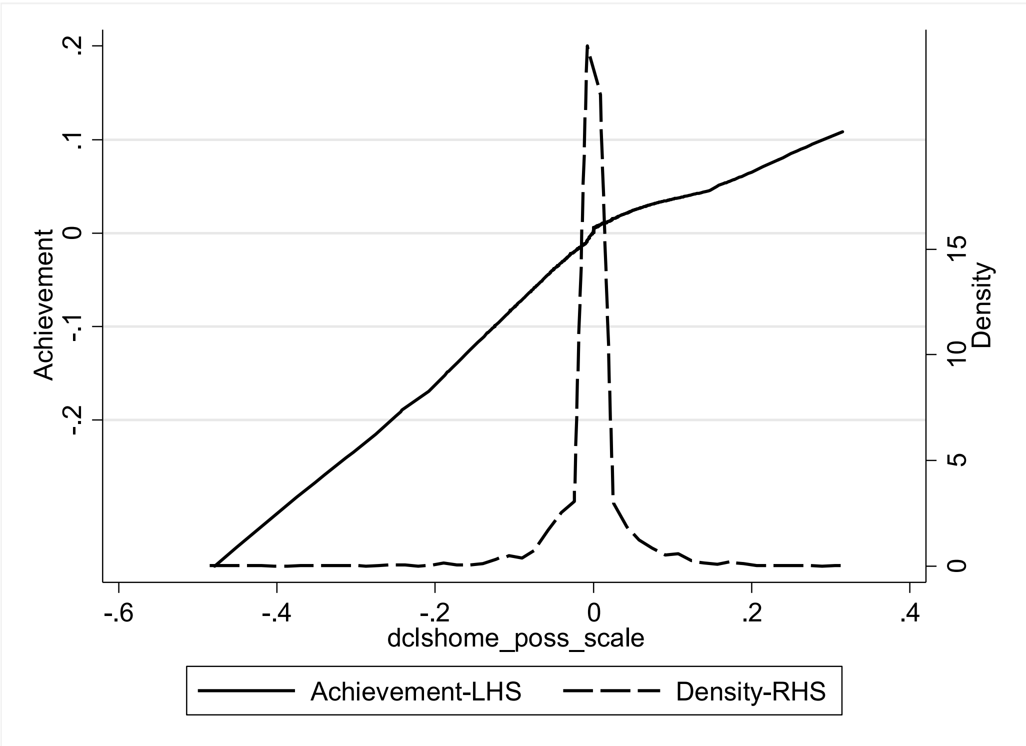


Figure 4: Class proportion with plans to go university and relationship with achievement

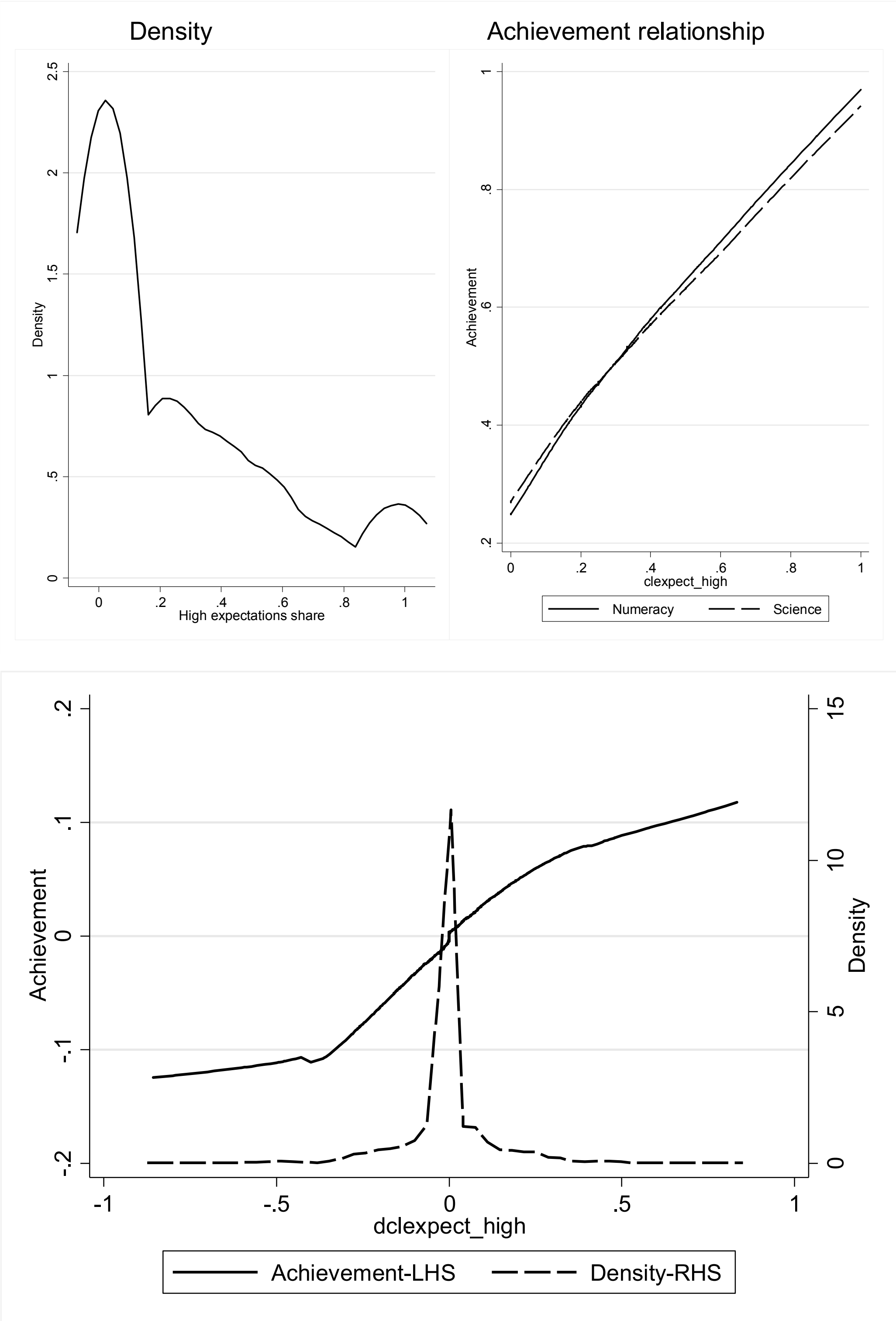


Figure 5: Class proportion with plans to leave before Year 12 and relationship with achievement

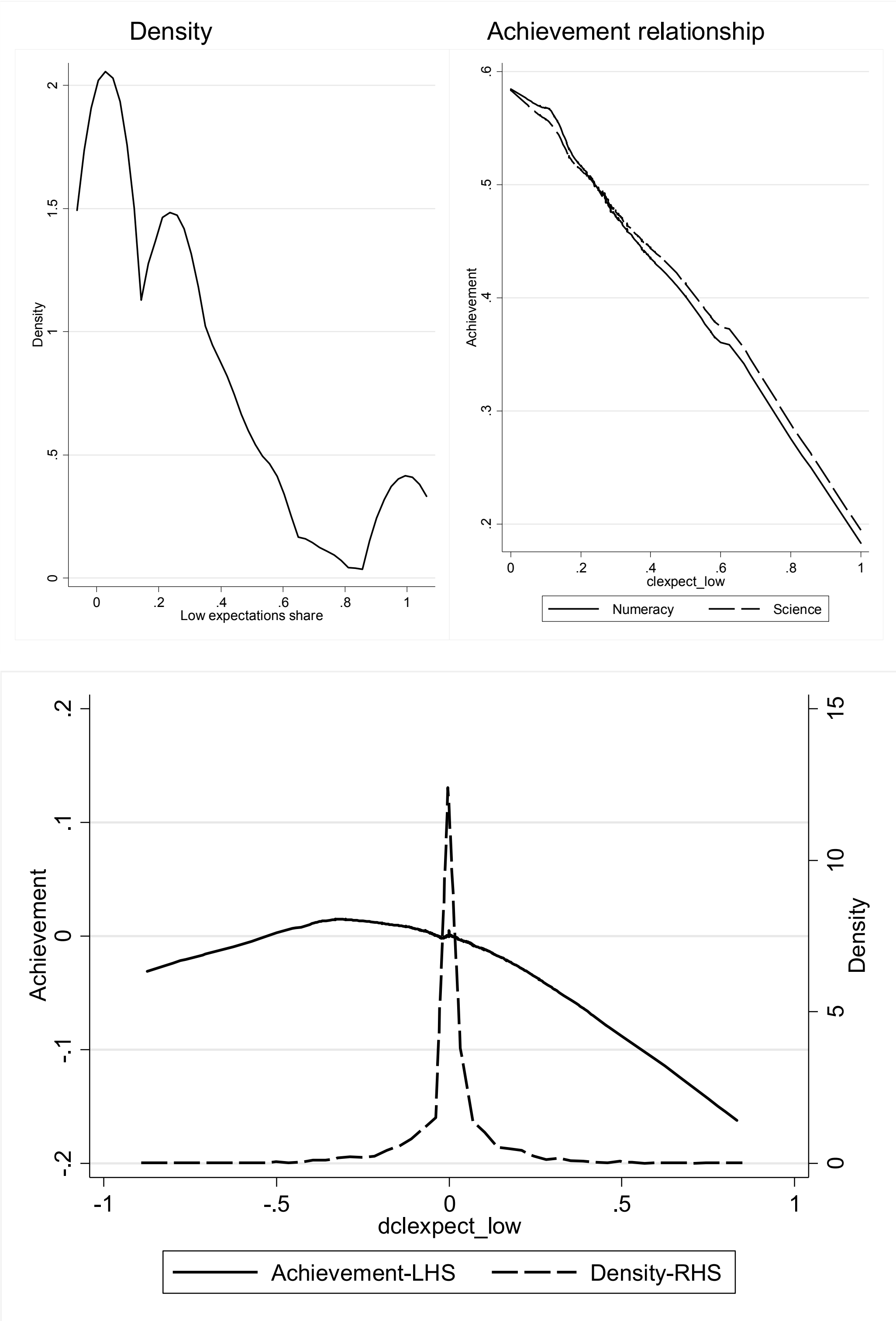
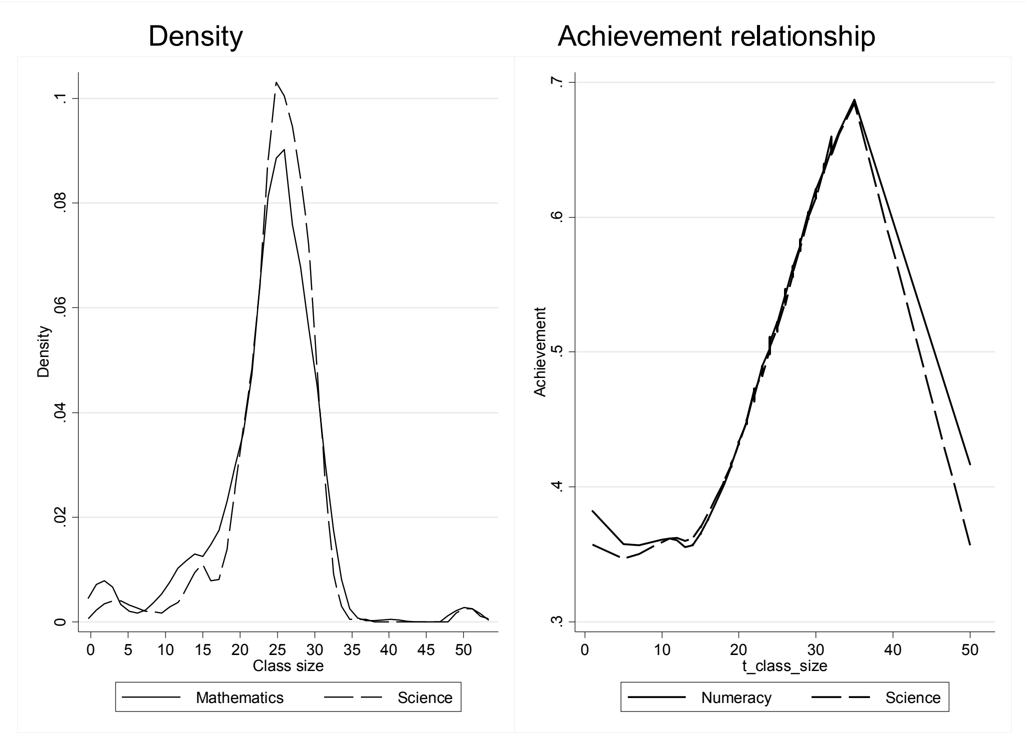


Figure 6: Class size and relationship with achievement



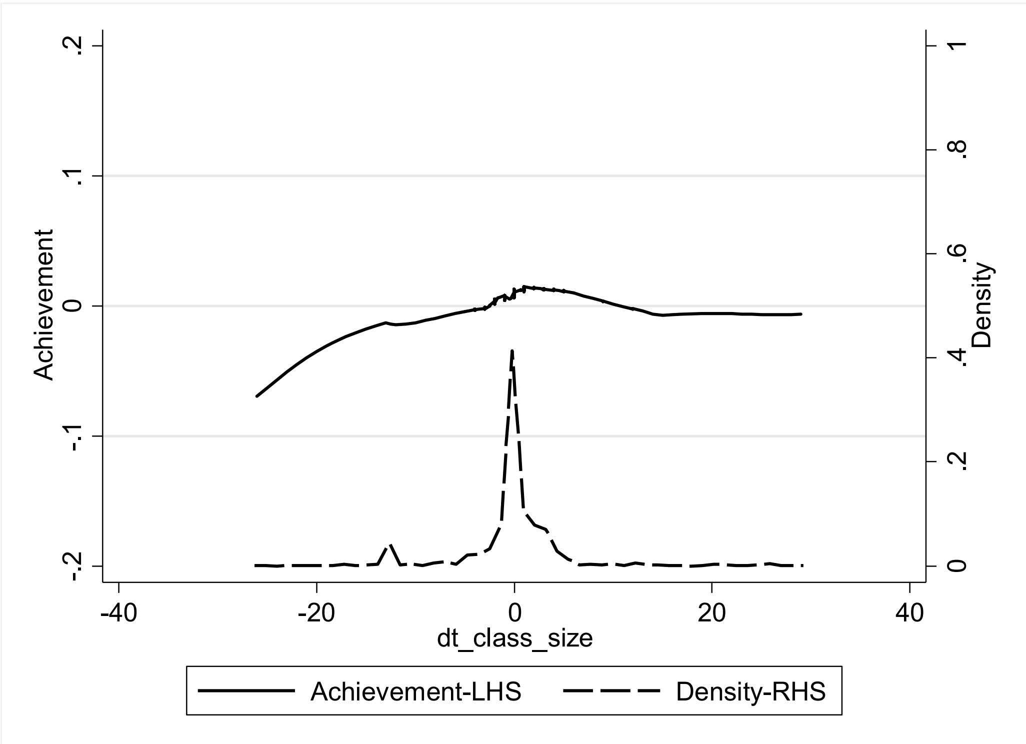


Figure 7: Average class conscientiousness scale and relationship with achievement

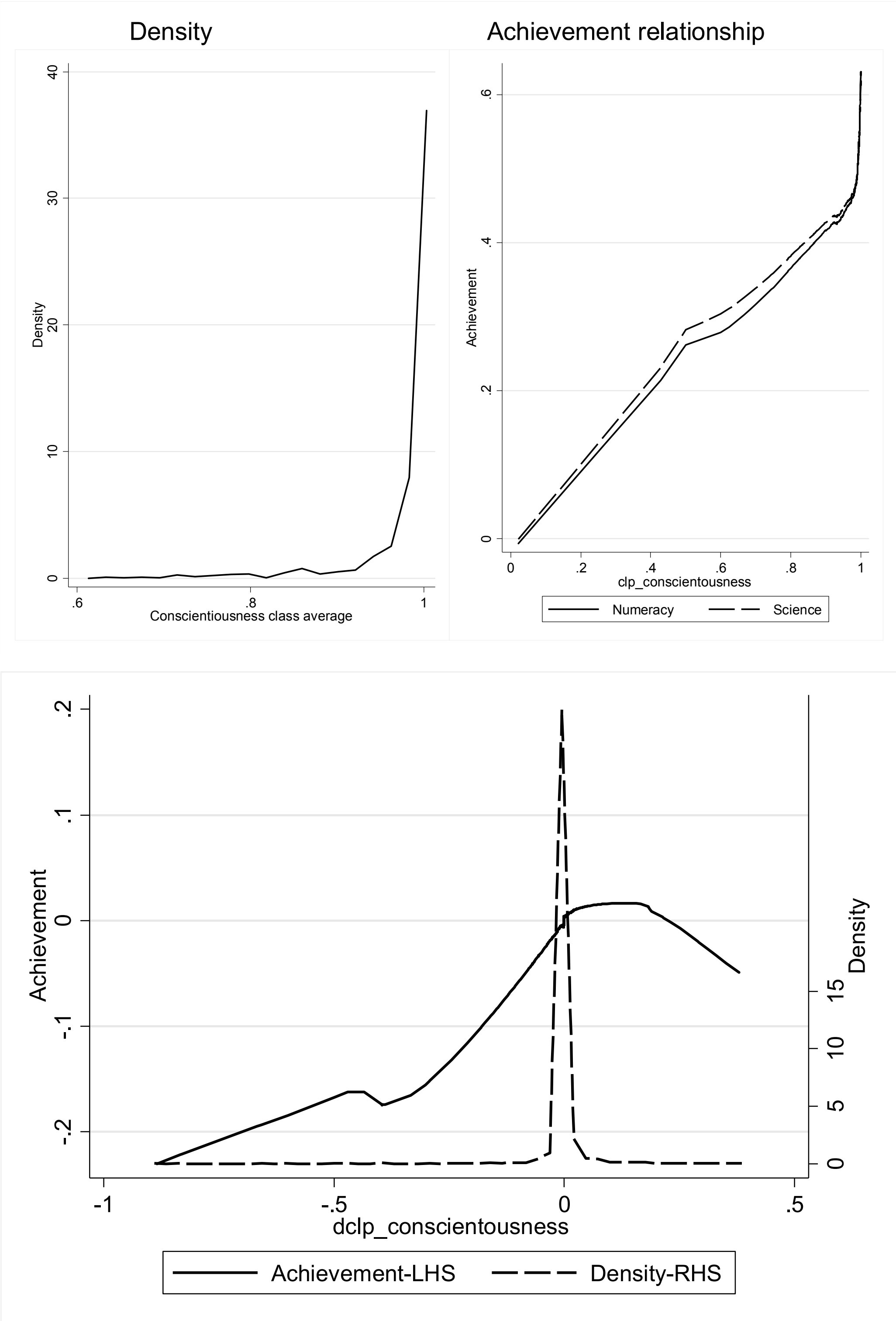


Figure 8: Class limits on learning scale and relationship with achievement

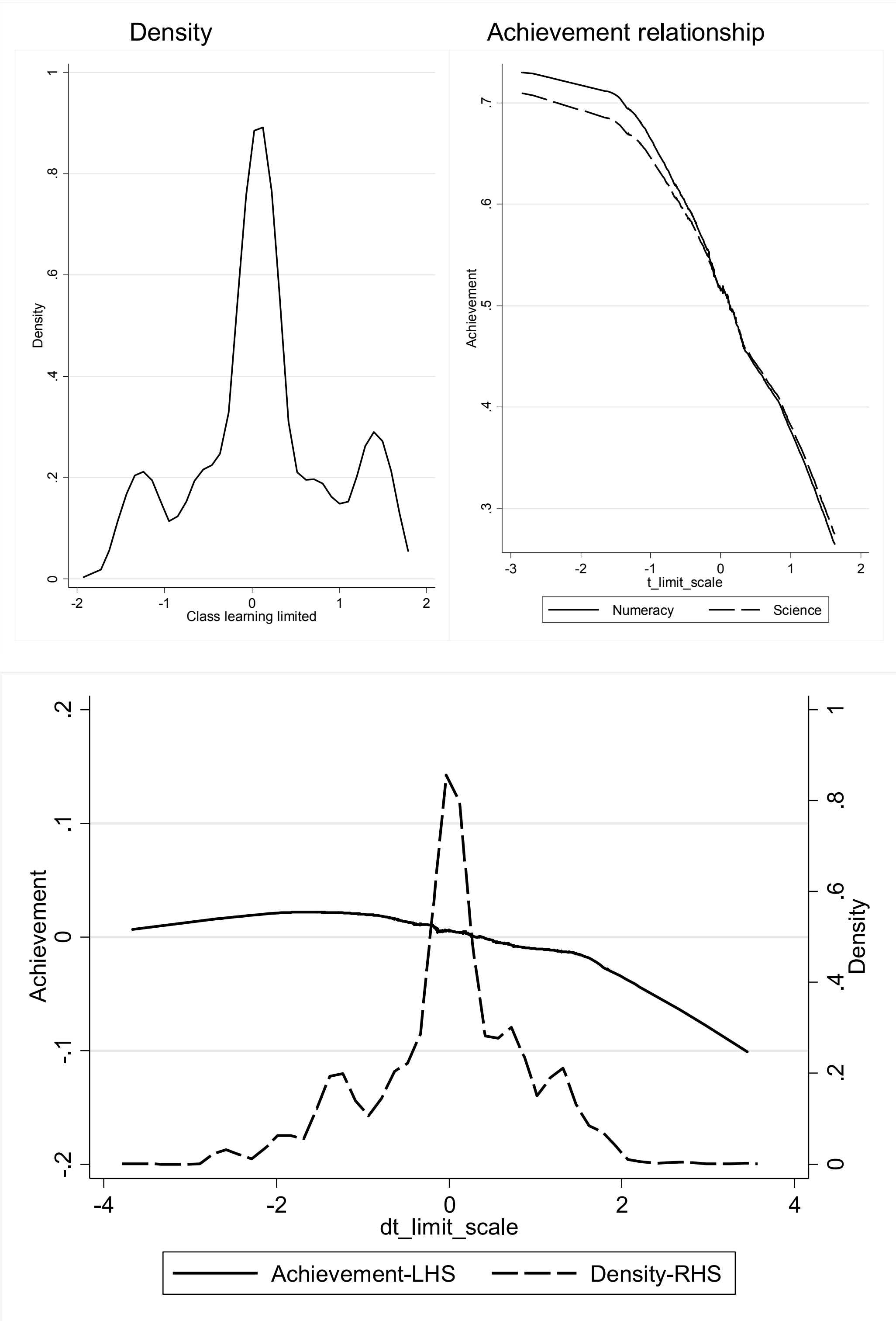


Figure 9: Proportion male and relationship with achievement

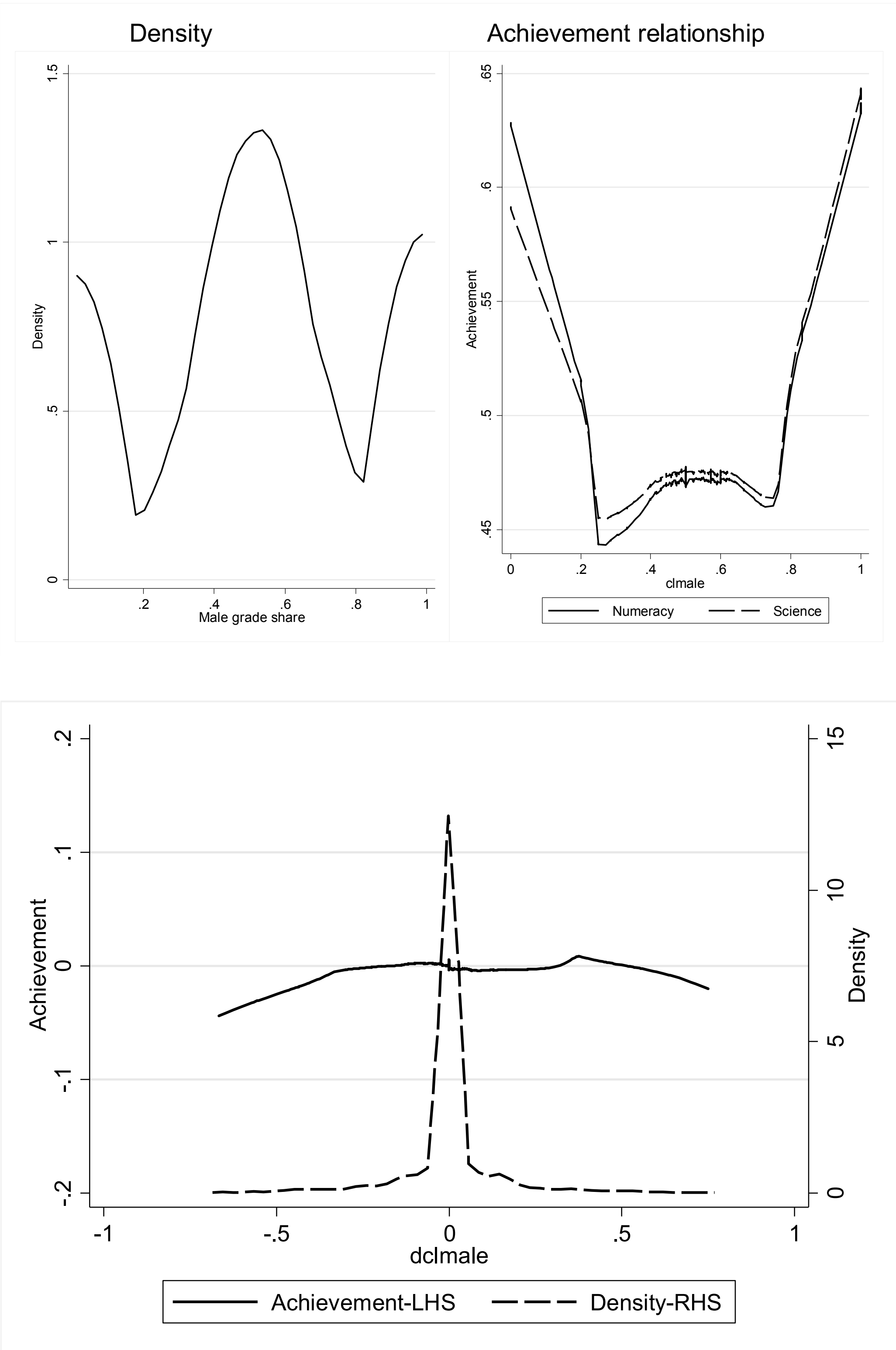
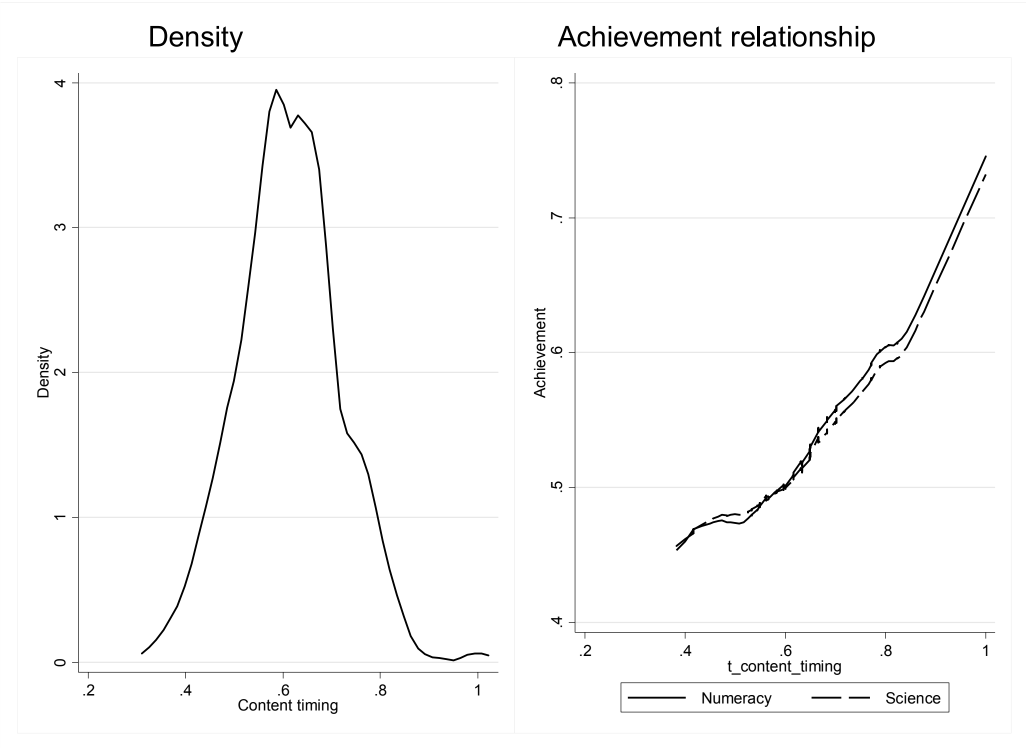
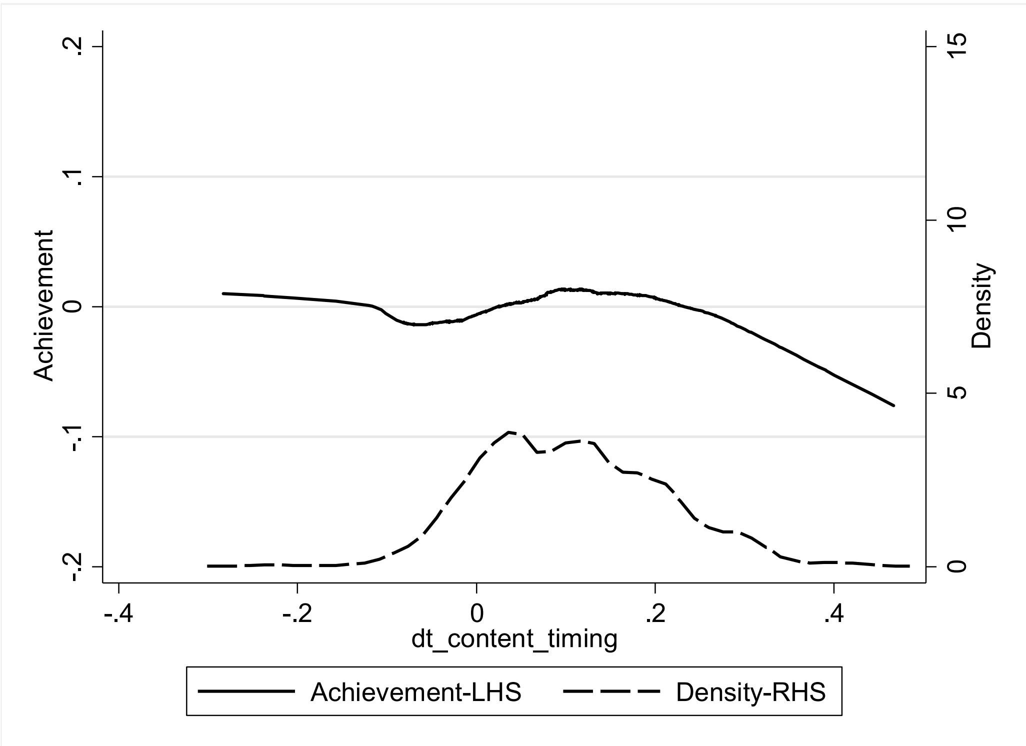


Figure 10: Timing of TIMSS test content and relationship with achievement





Appendix C – Teacher skills self-efficacy questions (mathematics teacher questionnaire)

How well prepared do you feel you are to teach the following mathematics topics? If a topic is not in the <eighth-grade> curriculum or you are not responsible for teaching this topic, Please choose “Not applicable.” – Possible responses: “Not applicable” (excluded from aggregate measure), “Not very well prepared” (= 1); “somewhat prepared” (= 2); “very well prepared” (= 3)

Science teacher question similar in structure around headings Biology, Chemistry, Physics and Earth Science

A. Number

a) Computing, estimating, or approximating with whole numbers

b) Concepts of fractions and computing with fractions

c) Concepts of decimals and computing with decimals

d) Representing, comparing, ordering, and computing with integers

e) Problem solving involving percents and proportions

B. Algebra

a) Numeric, algebraic, and geometric patterns or sequences (extension, missing terms, generalization of patterns)

b) Simplifying and evaluating algebraic expressions

c) Simple linear equations and inequalities

d) Simultaneous (two variables) equations

e) Representation of functions as ordered pairs, tables, graphs, words, or equations

C. Geometry

a) Geometric properties of angles and geometric shapes (triangles, quadrilaterals, and other common polygons)

b) Congruent figures and similar triangles

c) Relationship between three–dimensional shapes and their two–dimensional representations

d) Using appropriate measurement formulas for perimeters, circumferences, areas, surface areas, and volumes

e) Points on the Cartesian plane

f) Translation, reflection, and rotation

D. Data and Chance

a) Reading and displaying data using tables, pictographs, bar graphs, pie charts, and line graphs

b) Interpreting data sets (e.g., draw conclusions, make predictions, and estimate values between and beyond given data points

c) Judging, predicting, and determining the chances of possible outcomes