THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 22/16

What Can We Learn from Student Attitudes for International Achievement Tests?

Kyle Peyton, Chris Ryan and Justin van de Ven

MELBOURNE INSTITUTE®

of Applied Economic and Social Research

What Can We Learn from Student Attitudes for International Achievement Tests?*

Kyle Peyton[†], Chris Ryan[†‡] and Justin van de Ven[†]

**[†] Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne**

[‡] ARC Centre of Excellence for Children and Families over the Life Course

Melbourne Institute Working Paper No. 22/16

ISSN 1447-5863 (Online)

ISBN 978-0-73-405219-3

July 2016

* Disclaimer: The research in this paper was commissioned by the Victorian Department of Education and Training. The findings and views in this paper are those of the author and should not be attributed to the Department of Education and Training. Thanks to Jongsay Yong for helpful suggestions. Contact: <ryan.c@unimelb.edu.au>.

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

Abstract

This study looks at whether differences in student attitudes towards mathematics and science between Victorian students and those in selected other countries can explain differences in student achievement between them. We find that they cannot. In general, in the 2011 Trends in Mathematics and Science Study (TIMSS) data used here, Victorian school students have more positive attitudes towards mathematics and science than students in high achievement countries. These results also hold where we remove any language effects from the way people respond to attitudinal questions, or any cultural or social-desirability induced elements of the responses. Further, the most reliable estimates of the relationship between attitudes and achievement point to quite small effects, suggesting any increase in achievement associated with improved student attitudes could only be small.

JEL classification: I21, I28

Keywords: International tests, achievement, student attitudes

1. Introduction and motivation

Commentators and policy-makers have been impressed by the performance of school students in a small number of countries that are consistently among the top performers in international achievement tests. The governments of various countries have announced plans to emulate those performance levels, while others have sought to identify and understand what factors are behind gaps in achievement between their own country and the top performing countries. At present, achievement gaps between Victorian students and those in high achieving countries in Years 4 and 8 are of the order of 70-80 points in mathematics and 50 points in science, using measures with average values of 500 and standard deviations of 100 (see Figures 1 and 2).

The rationale for many reforms to school systems introduced over the past decade around the world has included the intention to match achievement in these high performing countries. Some of these reforms have targeted teachers and how they teach (e.g., differentiated pay, improved performance review, better recruitment and retention, alternative pedagogies), others principals (e.g., skills and autonomy), and others parents (e.g., parenting skills, aspirations for their children). Other reforms target the students themselves; for example, improving school readiness, improving well-being and reducing the experience of bullying. This might result in students being more motivated and having better attitudes towards learning.

This paper examines whether differences in academic achievement between students in Victoria and high performing jurisdictions can be explained by variation in self-reported student attitudes, such as motivation, engagement and self-confidence in study domains. The methodological approach (known as the Blinder-Oaxaca decomposition – Oaxaca 1973 and Blinder 1973) adopted allows us to look at whether Victorian students have different attitudes towards their school subjects than students in high achieving countries. It also allows us to see if achievement has the same relationship with these attitudes across countries and whether some attitudes might be more important for achievement than others. For example, does liking a subject matter more than believing it to be important?

Although the approach is entirely descriptive, it provides an empirical foundation for assessing the relationship between comparative differences in student attitudes and academic achievement. Before we can argue that making Victorian students' attitudes look like those in high performing countries might improve their achievement, we need to know if their attitudes are in fact different. Moreover, we need to know if achievement is associated with the attitudes differently between countries. For example, maybe the “secret” of high performing countries is that even those with poor attitudes obtain high levels of achievement, while the most motivated students do comparably in all countries. Of course, causal relationships between achievement and attitudes towards subjects are likely to run in both directions: better attitudes may both influence and be affected by achievement. Before we try to estimate such causal effects, however, it is important to document the nature of the observed relationships. If we find that the attitudes are no different between students in Victoria and high performing countries, or that the association with achievement is no different, we may have a different view about whether trying to influence them will help Victorian students close the achievement gaps. This study allows us to assess whether changing attitudes might be a fruitful area to focus on in the drive to improve student achievement, or whether we can rule it out as providing any scope to make up current achievement gaps.

We find that neither differences in the way students answer relevant attitudinal questions, nor in the way such factors are associated with achievement, explain much of the differences in achievement between Victorian students and those in high performing countries (Hong Kong, Japan, Singapore, South Korea and Taiwan). Students are not less positive towards their subjects in Victoria than those in high performing countries. Further, the increments in achievement associated with better attitudes are generally similar in Victoria to the high performing countries. Instead, the better achievement in high performing countries occurs across all response categories to the attitude questions, suggesting that the reasons for the achievement gaps are largely independent of differences in student attitudes and their associations with achievement between the countries. So, in general, we should look elsewhere for mechanisms that might help close the gap in achievement between Victorian students and those in high performing countries.

The remainder of the paper is organised as follows. The next section contains a review of the relevant literature on international differences in student achievement and on the relationship between student achievement and attitudes. Section three contains descriptions of the methodology and the data used here, while section four contains the results. Concluding comments are made in section five.

2. Potential explanations for differences in average achievement

2.1. Literature on potential explanations

Cross-country (or "between-country") analysis of student achievement data offers a number of advantages over within-country analysis (see Hanushek & Woessman (2011) for a review). The chief advantage here is that comparative analysis allows us to explore the extent to which cultural (e.g. student attitudes, parental behaviour) and institutional differences (e.g. teacher accountability, school autonomy) drive variation in student achievement. The key challenge for the literature has been distinguishing between the two. How much of the variation in student achievement reflects differences in formal institutions versus differences in "culture" or prevailing social norms? Although this is impossible to answer with certainty in general terms, it motivates much of the research on international differences in student achievement. In this section we review the various institutional and cultural factors associated with between country differences in student achievement.

The primary institutional explanations for variation in student achievement focus on 1) accountability, 2) autonomy, 3) private competition, 4) tracking, 5) pre-primary education and 6) school inputs. Among these, school inputs seem to provide the least explanatory power for variation in student achievement across countries. According to the meta-analysis of cross-country empirical work conducted by Hanushek & Woessman, "it is hard to find evidence of substantial positive effects of most resource inputs, in particular class sizes and expenditure levels" (2011, p. 159). Although within country studies on this topic are decades old, this finding seems to confirm some accounts (Hanushek, 1996) and challenge others (Hedges & Greenwald, 1996; Krueger, 1999).

Although accountability and autonomy measures are strong predictors of variation in cross-country regression equations the interaction between the two seems to provide the most explanatory power. For example, Woessmann (2005) shows that the presence of external exit exams – a form of centralized accountability—and school autonomy over teacher salaries explains three quarters of a standard deviation in mathematics achievement as measured by the Trends in Mathematics and Science Study (TIMSS). However, there is substantial heterogeneity across countries and although school autonomy may improve achievement in developed countries, it likely has a negative effect on student achievement in developing countries (Hanushek, Link & Woessmann, 2013).

Private school management also correlates with cross-country differences in student achievement: students in countries with more private providers tend to have higher achievement scores. However, private schooling seems to be a proxy for competition so that the link to school achievement operates through increased competition between public and private providers. This association is far from general and varies considerably across countries (Woessman, 2009). Private competition seems to work best when sources of government funding are in place to eliminate credit constraints for the poor (Woessmenn *et al.* 2009).

Tracking and pre-primary education are both associated with inequalities in student outcomes. Hanushek & Woessmann (2006) show that educational inequality increases as students make the transition from primary to secondary school in countries with early tracking, whereas it decreases in countries without tracking. Pre-primary or early childhood education programs tend to equalize student achievement across socioeconomic backgrounds. Schuetz *et al.* (2008) demonstrate the systematic association between pre-primary education and later achievement using results from TIMSS and Schuetz (2009) finds similar patterns in the Programme for International Student Assessment (PISA) data. The cross-country results can be viewed as a subset of a wide variety of studies that support the strong link between early childhood education and later academic achievement (see Currie 2001; Blau & Currie 2006 for reviews).

Despite the strong associations between various institutional factors and student achievement, there is reason for scepticism. Cultural differences that correlate with educational achievement are one of the primary threats to the validity of between country comparisons (Hanushek & Woessmann, 2011, p. 110). Although commentators and policy makers commonly assume that variation in student achievement between countries is a function of national education policies, it does not follow that domestic policy changes will produce changes in outcomes. Persistent cultural differences may provide a stronger link to student achievement, even in the presence of institutional changes.

Feniger & Lefstein (2014) illustrate the pitfalls of what they call “PISA reasoning” by examining the mathematics scores of Chinese immigrants with at least ten years of education in Australia and New Zealand. They find that the Chinese students score considerably higher than their non-Chinese peers and on par with students in Shanghai, even after controlling for family resources. They conjecture that these differences in achievement are based on differences in extra-curricular schooling: while about 40% of Chinese students in Australia and New Zealand attended "out-of-school" enrichment classes compared to only 10% of non-Chinese students.

This finding suggests parental involvement may be a more important mechanism for student achievement than institutional differences. Houtenville & Conway (2008) also show that various measures of parental effort (e.g. attending school meetings, volunteering at the child's school, etc.) have a large effect on student achievement independent of socioeconomic family background characteristics. Of all the extra-institutional factors that influence educational outcomes, family background—measured by socioeconomic status or parental involvement—is the strongest predictor of student achievement. This is one of the few findings in the comparative literature that seems universally accepted. Although the importance of family background varies across countries, indicators of higher socio-economic status and parental involvement are associated with higher student achievement in 54 countries (Schuetz *et al.*, 2008).

Perhaps not surprisingly, the attitudes and dispositions that students develop are also correlated with academic achievement. Although there is very little evidence that specific personality attributes (e.g. anxiety, dogmatism, extraversion, locus of control, neuroticism) are associated with achievement, measures of self-efficacy, self-concept, motivation and persistence are moderately correlated with achievement (Hattie, 2009, Chapter 4). Debeer *et al.* (2014) present evidence from 2009 PISA data that suggests that examinee effort declines over the course of the (low stakes) PISA reading test in all countries, with the decline greatest among students with lower reading ability. There were also differences in average persistence between countries, with students in Singapore, Hong Kong and South Korea, but not Japan, all exhibiting higher average persistence levels than Australian students. There is also evidence that attitudes to specific school subjects (e.g. positive or negative feelings, beliefs that a subject is useful, etc.) are associated with achievement (Aitken, 1969; Ma & Kishor, 1997; Neale, 1969).

Much of the research on the link between student attitudes and achievement has focused on how attitudes toward mathematics and science predict student achievement. Students who have positive attitudes toward mathematics and science tend to be good at mathematics and science, and those who are tall tend to be good at basketball. Of course being tall does not ensure that someone will be good at basketball, in just the same way that having positive attitudes does not ensure that a student will earn high marks in mathematics and science. There is clearly a reciprocal relationship between attitudes and achievement

(Hattie, 2009, Chapter 4; Marsh & Martin, 2011) and although poor attitudes can serve as a barrier to learning, there is little reason to believe the "causal arrow" runs from attitudes to achievement in any general sense (Ma, 1997; Ma & Xu, 2004a; Ma & Xu, 2004b).

Like students elsewhere in the world, Australian students who report positive domain specific attitudes (e.g. "I like learning mathematics") tend to outperform their domestic peers on standardized tests (Thomson *et al.*, 2012a; Thomson *et al.*, 2012b). However, Australian students tend to report less favourable attitudes toward mathematics and science than the international average in TIMSS data. One appealing thought experiment is: if Australian students had better attitudes toward mathematics, would this mean better academic performance? In the next section we detail a procedure for addressing this question.

3. Methodology and data

3.1. Methodology

We rely on a high-level exploratory analysis of the role of attitudes in explaining achievement disparities between Victorian school students and students from jurisdictions that rank in the top five internationally. Our analysis uses data from the Trends in Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), which we discuss in more detail in the next sub-section. Achievement disparities are defined in terms of standardised test scores reported for students in different jurisdictions. Attitudes are described by categorical responses to survey questions, such as*:*

How much do you agree with these statements about learning mathematics?

It is important to do well in mathematics:

agree a lot *agree a little* *disagree a little* *disagree a lot*

In general, survey responses that indicate positive (negative) attitudes tend to be associated with higher (lower) test scores. That is, students who answer "agree a lot" to the example statement quoted above, tend to achieve higher test scores for mathematics than those who answer "disagree a lot". Figure 3 shows the average achievement levels of Victorian Year 4 students who gave each of the four possible responses to this question in the TIMSS data analysed here. The average mathematics score of Victorian Year 4 students was about 530. Those who agreed a lot had an average score of just over 550 in mathematics, while those who disagreed a lot had an average score in mathematics that was about 100 points lower. This pattern is replicated across other questions in mathematics, and across the learning domains of reading and science: more positive attitudes are associated with higher levels of achievement. Data, such as those that lie behind Figure 3, invite the reaction that if only we could get more children to agree a lot or even a little with that statement, we might see average mathematics achievement improve. This study takes the positive association between attitudes and achievement as given, and instead explores the question of whether jurisdictional disparities in responses to attitudinal questions help to explain differences in associated test scores.

Specifically, it is possible to apportion or decompose the gap in achievement between two jurisdictions into three contributing factors: (1) differences in the way students in the jurisdictions answer the attitude questions; (2) differences in the way answers to the attitude questions are related to achievement in the two jurisdictions; and (3) differences independent of the distribution of the answers and the way achievement moves with attitudes (see Appendix A for details). The first element will reflect whether more children in Hong Kong, for example, answered "agree a lot" to the question in Figure 3 than did Victorian Year 4 students. The second element examines whether the 100 point difference in achievement in Victoria between the "agree a lot" and "disagree a lot" groups is similar or different than in Hong Kong. The third element of this decomposition of differences in achievement is relevant where nothing about differences in the responses to the question between countries seems to be related to the observed difference in achievement.

The decomposition approach allows us to answer a number of questions. Would students in Victoria fare better if their attitudes towards education were more like students in high achieving countries? Or would they do better if attitudes were associated with achievement more like the relationship that holds in other countries? Or does the gap in achievement occur across all response categories, so that the specific question is uninformative about what contributes to the gap in achievement, other than as a useful signal to look elsewhere.

This approach is known as the Blinder-Oaxaca decomposition (Oaxaca 1973 and Blinder 1973) and is widely used in economics to study the difference in the outcomes of two distinct groups, such as determining how much of the gender wage gap reflects differences in the average characteristics of men and women and how much arises because of the way the labour market rewards those characteristics (see Altonji & Blank 1999 for a review that focuses on differences in the rewards, or "discrimination", or Fortin *et al.* 2011 for a more recent review). The approach has also been used to study achievement gaps in the economics of education literature (for example, McEwan & Marshall 2004, Krieg & Storer 2006, Ammermueller 2007, Fuchs & Woessmann 2007).

Use of the Blinder-Oaxaca decomposition approach to the study of differences in achievement across countries requires a number of assumptions. First, we must assume that prompts such as "It is important to do well in mathematics" are interpreted in the same way in the two countries. Second, the response categories of meaning "agree a lot" and so on must have the same interpretation and cultural acceptability in the two countries. Third, the approach normally assumes that any causation runs from the measured characteristics (in this case, attitudes) to the outcome (in this case, achievement).[1] Fourth, the Blinder-Oaxaca decomposition only tells us something about differences in the average scores of the two groups, not about differences in the groups across the entire distribution of achievement.[2] Finally, with a set of indicator variables as the sole explanatory variables, the decomposition approach results can be sensitive to the choice of the omitted category if the way achievement increases with attitudes is very different between countries. This can result in somewhat different allocations of the difference in achievement between the elements of the decomposition we use (Gardeazabal & Ugidos 2004).

In this study, we attempt to assess the role of these potential reservations to our approach via a number of robustness checks. We address the first two problems via a "double" differencing decomposition approach (Wellington 1993 and described in Appendix A). The approach has two conceptual elements. First we compare the responses to a prompt such as "*It is important to do well in . . .*" mathematics with those of science within one country. This removes any role for language or cultural factors, since the comparison is done within the same country. Again, it decomposes any gap in achievement between mathematics and science within the country to differences in attitudes, to the association between

---

[1] For example, in standard economic applications the individual characteristics (e.g. skills, education, gender, etc.) determine wages and the rewards or returns associated with characteristics in a wage equation can be thought of as well-measured, meaningful parameters. In the current context, however, it is more likely that causation runs in both directions, between achievement in a particular domain and student attitudes towards that domain (using US data, Ma & Xu 2004a, 2004b conclude it mostly runs from achievement to attitudes, not vice versa, but other studies point to it operating in a more dynamically reciprocal way – eg. Marsh & Martin 2011).

[2] Researchers have developed newer techniques that are more informative across the entire distribution (these are reviewed in Fortin et al. 2011).

achievement and attitudes, and to factors independent of the attitudes. In the second element of the decomposition, the differences between mathematics and science in one country are compared to the differences in another country, assessing the role of different distributions of attitudes and their associations with achievement in explaining differences in achievement between the countries.

On the third reservation, we will accept the argument that the causation runs predominantly from attitudes towards achievement, since that is necessary for interventions on attitudes to affect achievement. Further, if there is causation that runs from achievement towards attitudes, our view is that any biases in the parameters will likely be similar in both countries, so the differencing procedure will remove any effect of the biases from the main results.

The fourth reservation is correct, but misses the purpose of this analysis. Since comparisons between Victoria or Australia and other jurisdictions are mostly in terms of differences in average achievement, using a technique that is informative about differences in averages as is done here is appropriate for the research questions.

Finally, the fifth reservation concerning sensitivity of the decomposition to the choice of omitted category is accommodated in our analysis by running sensitivity tests to assess whether this issue is important in our context.

3.2. PIRLS & TIMSS Data

The study uses Australian and international data from two studies: the Trends in International Mathematics and Science Study (TIMSS) for 2011 and the Progress in International Reading Literacy Study (PIRLS) 2011 survey (see Thomson *et al.* 2012a, 2012b). PIRLS involves an assessment of the reading achievement of students in Year 4 in 2011. TIMSS involves an assessment of the mathematical and scientific achievement of students in Years 4 and 8 and is the focus of most of our analysis. For the 2011 survey, a stratified random sample of 280 primary Australian schools (42 in Victoria) participated in the data collection for both the TIMSS and PIRLS 2011 Year 4 studies. The sample of secondary schools for Year 8 TIMSS in 2011 was of 275 schools (43 in Victoria). The samples were stratified by state, school sector and rural geography. The national sample sizes of the Australian data used here are set out in Table 1, along with those for Victoria.

Some 6126 and 6146 students undertook the PIRLS and TIMSS Year 4 tests respectively (around 764 in Victoria) in 2011. Two classrooms per school were sampled, along with all of the indigenous students found in Year 4 in the sampled schools. All sampled students for TIMSS at Year 4 were also asked to take the PIRLS reading tests. More students undertook the Year 8 TIMSS tests – 7556 (958 in Victoria). There were no overlapping schools between the Year 4 and 8 samples.

The international student and school sample sizes for five high performing "countries" also appear in Table 1: Hong Kong, Japan, Singapore, South Korea and Taiwan. In general, the number of students sampled in these countries in each of Year 4 and Year 8 was in the range 4000 to 6500 students from between 120 to 180 schools – less than the Australian samples, but obviously larger than the Victorian one.

A key purpose of the PIRLS and TIMSS studies is the production of internationally comparable student achievement scales in mathematics and science (for Years 4 and 8) and reading (for Year 4 only), that are also comparable over time. These scales are all calculated to have a mean of 500 and standard deviation of 100 across students from all countries participating in the PIRLS and TIMSS studies for both Year levels.

The average scores for Victorian students versus those in high achieving countries appear in Figures 1 and 2. Figure 1 shows the average Year 4 scores in reading, mathematics and science of the countries, along with the 95 per cent confidence intervals around the respective means for countries. Figure 2 shows the average Year 8 scores in mathematics and science of the countries and the associated confidence intervals. All of the average values for Victorian students are at least 500, which indicate that the tests generally show Victorian students to be performing above the international average, if well below achievement in the high performing countries.

Victorian students perform well below students from the high achieving countries such as Hong Kong, Japan, Singapore, South Korea and Taiwan, in a statistical sense, as demonstrated by the fact that the gap between the mean values and confidence intervals around the means do not overlap in general. The rest of this paper is designed to assess whether differences in student attitudes between countries contribute to these achievement gaps.

In addition to the achievement tests, students completed questionnaires about their family backgrounds and attitudes towards reading (in PIRLS), mathematics and science (in TIMSS). These attitudinal questions contribute to various attitudinal scales used in the

PIRLS/TIMSS environment.[3] These scales indicate how much students *like* a study domain (reading, mathematics or science), are *confident* about learning it, *value* it in terms of their future study and work careers, or are *engaged* by the classes in which they study it. Rather than study the scales, this paper contains analysis of the responses to the individual questions to assess how different they are across countries.[4] The individual items were used, rather than the scales, to increase the chances of isolating important attitudes. Students were asked to respond to the individual questions or prompts by indicating how much they *strongly disagree*, *disagree*, *agree* or *strongly agree* with the specific questions.

Despite the best efforts of the study designers, it is possible that language and cultural differences across countries might mean that the language of the questions and the possible responses are interpreted differently across countries, or that social desirability pressures differ across countries. This could induce different response patterns between countries where they do not really exist, or mask differences where they do. Below, we address the possible implications of this for our analysis where we exploit the fact that, in general, the individual questions used to elicit responses from individuals about their attitudes towards mathematics and science are identical, except for the exchange of the words "mathematics" and "science".

4. Results

4.1. Simple intuition of the nature of the results

As a way of introducing the overwhelming pattern found in the results, we consider first one question asked of Year 8 students in TIMSS and look at how different the responses are between Victorian students and those in high achieving countries. The question is "*How much do you agree that you like mathematics*?" with possible responses "disagree a lot", "disagree", "agree" and "agree a lot".

As noted, the Blinder-Oaxaca approach decomposes the gap in average achievement between two jurisdictions into three contributing factors: (1) differences in the way students in the jurisdictions answer the attitude question; (2) differences in the way answers to the

---

[3] The attitudinal scales used in other international testing regimes, such as PISA differ from those in the TIMSS environment, though all of the TIMSS scales have somewhat equivalent scales in PISA other than the *engaged* scale.

[4] Analysis of the scales is carried out for robustness purposes, with results described in the relevant section below.

attitude question are related to achievement in the two jurisdictions; and (3) differences independent of the distribution of the answers and the way achievement moves with attitudes.

In the first instance, we show the elements of the decomposition for Year 8 Victorian students compared with those in Hong Kong. In Figure 4, the distribution of responses across the four categories in Victoria and Hong Kong are presented. The proportion of students in each category are almost identical, with the proportion indicating that they "agree a lot" that they like mathematics is marginally larger in Hong Kong than Victoria but there is no statistical difference in the two distributions. In terms of the first element of the Blinder-Oaxaca decomposition, it seems that differences in the way students answer how much they like mathematics in Victoria and Hong Kong can contribute very little to understanding why average achievement might be different between them.[5]

In Figure 5, the average achievement in mathematics associated with each of the four attitude categories are presented for Victoria and Hong Kong. These are different between Victoria and Hong Kong, being substantially higher for each response category in Hong Kong than Victoria. The difference is typically of the order of about 80 points. In fact, the gap in average achievement is remarkably constant across the four response categories. The average achievement of Victorian students is about 80 points worse than those of students in Hong Kong regardless of which response they give to how much they like mathematics. In terms of the second element of the Blinder-Oaxaca decomposition, it seems that relative differences in the way achievement is associated with the answers students give about how much they like mathematics between Victoria and Hong Kong also contributes very little to understanding why average achievement might be different between them.[6]

Instead, it appears that differences in average achievement between Victoria and Hong Kong are independent of the distribution of the answers students give and the way achievement moves with how much student like mathematics. That is, the difference in average achievement is not explained in any way by differences in the answers students give

[5] In the terms of equation (3) of the appendix, $(\bar{\boldsymbol{p}}_2 - \bar{\boldsymbol{p}}_1) \approx 0$, so the second term on the right hand side of equation (3) is also approximately zero.

[6] In the terms of equation (3) of the appendix, $(\hat{\boldsymbol{\beta}}^2 - \hat{\boldsymbol{\beta}}^1) \approx 0$, so the third term on the right hand side of equation (3) is also approximately zero.

about how much they like mathematics between the jurisdictions or in the way achievement moves with how much student like mathematics in the two jurisdictions.[7]

In fact, this type of result is representative of all of the results decomposing average achievement gaps between Victorian students and those in high performing countries across mathematics and science, for Years 4 and 8, and across the multitude of attitude questions. The gaps in average achievement are not explained by differences in the way students respond to questions on how they view mathematics or science, or to different patterns in the way achievement and attitudes are related. These attitudes have little to contribute to understanding achievement gaps between Victorian students and those in high performing countries.

In Figure 6, the elements of the formal decomposition of the gaps in average achievement between Victorian students and those in all of the high performing countries we study throughout the rest of the paper, Hong Kong, Japan, South Korea, Singapore and Taiwan, are presented for the question analysed in Figures 4 and 5. The contributions towards the achievement gaps of the elements of the decompositions undertaken for each country for that question are shown, divided by the achievement gap between Victoria and each jurisdiction to make the results comparable. The contribution of differences in the answers to how much students like mathematics is shown in blue in each bar graph, the contribution of the differences in the relationship between achievement and the responses in red, and the part that is independent on how much students like mathematics in green (referred to in the Figure as the "residual"). Since the contributions are all divided by the achievement gap between Victorian students and those in each jurisdiction, the three contributions are scaled to sum to one. The first bar shows the formal results for Hong Kong described already: that the differences in attitudes (in blue) and in the relationship between achievement and the responses (in a reddish-brown colour) make only small contributions to the total gap in achievement, jointly less than 5 per cent of the gap. Instead, there is just a gap in achievement that is overwhelmingly independent of how students answer the question about how much they like mathematics.

The second bar shows the contributions to the average achievement gap between Victorian students and those in Japan. The negative contribution from differences in attitudes

[7] In the terms of equation (3) of the appendix, $\bar{y}^2 - \bar{y}^1 \approx (\hat{\alpha}^2 - \hat{\alpha}^1)$, so the first term on the right hand side of equation (3) overwhelmingly "explains" the gap in average achievement .

(in blue) indicates that Victorian students have more positive attitudes towards mathematics, in terms of their responses to how much they like it, relative to the Japanese students. Japanese students, however, perform better than Victorian students in mathematics, despite seeming to like it less. But in general, the large green contribution means that there is just a gap in average achievement in mathematics between students in Victoria and Japan that is unrelated to how much students indicate they like mathematics. The same conclusion holds broadly for the decompositions of the gaps in achievement between Victorian students and those in South Korea, Singapore and Taiwan – large green contributions mean that the gap in average achievement in mathematics between students in Victoria and the other countries is not related to how much students indicate they like mathematics. Like the comparison with Japan, Victorian students have more positive attitudes towards mathematics than students in both South Korea and Taiwan, so it seems getting Victorian students to be more positive about mathematics cannot be very helpful in closing the gaps in achievement between those countries.

Moreover, the results for the decomposition of the achievement gap between Victorian students and those of high performing countries are qualitatively similar for attitudes other than how much students say they like mathematics. Figure 7 contains the decomposition results relative to the set of five high performing countries for the like mathematics question and a further set of four questions that together make up the Year 8 *Students like learning mathematics* scale in TIMSS (the other items involve responses to: I enjoy learning mathematics, I wish have not to study mathematics, Mathematics is boring and I learn interesting things in mathematics). The first four bars show the decomposition results against Hong Kong for these four items. The fifth bar repeats the decompositions for the "like mathematics" question and the sixth bar shows the decompositions for the responses to the prompt that "It is important to do well in mathematics" for Hong Kong. The next six bars show the decompositions for Japan, and so on for South Korea, Singapore and Taiwan. What should be clear from Figure 7, like Figure 6, is that the green bars dominate the decompositions, confirming that the gap in achievement is not explained by differences in how students answer the question about their attitudes towards mathematics in Victoria compared to high performing countries or how attitudes are related to achievement.

Figure 8 extends the analysis still further. It contains the decompositions of average mathematical achievement for forty-one prompts for Victorian Year 8 students compared with those in the five high performing countries. The prompts include the complete set of mathematics attitudinal variables used in the *like learning*, *confidence in learning*, *engaged* and *value* mathematics scales, as well a set of prompts about books in the home and other home possessions, school satisfaction and the experience of bullying and educational attainment aspirations. Once more, the green bars dominate the decompositions, confirming that the gap in achievement is not explained by differences in how students answer the question about their attitudes towards mathematics in Victoria compared to high performing countries or how attitudes are related to achievement. In fact, in many of the cases, Victorian students have attitudes or characteristics that are already in their favour, in terms of achievement. This is so for factors such as home possessions and the number of books in their households, as well as better attitudes towards mathematics as already seen.

Without labouring the point more than has already been done, it suffices to say that the decomposition results for Year 8 science are similar to those shown for mathematics and the Year 4 results similar to those for Year 8 – confirming that the gap in achievement is typically not explained by differences in how students answer the question about their attitudes towards mathematics in Victoria compared to high performing countries or how attitudes are related to achievement.

4.2. Detailed results

The results presented to date suggest in a convincing, yet informal, way that the attitudes studied here do not contribute much to an understanding of what explains differences in average achievement between Victorian students and those from high performing countries. We now focus more on the formal significance of the elements of the decompositions undertaken here, albeit in a visual way intended to summarise the results. This is done in Figures 9 and 10 for Year 8 mathematics and science.

White cells in Figures 9 and 10 indicate that differences in the individual attitudes listed (along with other measures of parental engagement with school, school satisfaction and the experience of bullying) or the way they are associated with achievement appear to contribute significantly to the observed differences in achievement between Victorian students and those from high performing countries. Other colours (see the legend) either indicate that the estimated effects are not significant or that their signs work in the wrong direction to explain the observed differences in achievement. There are very few white cells in the tables and no obvious systematic patterns of note. There are more white cells in the Singapore column than for other countries, especially for attitudes around liking science and valuing its importance, and a few instances where an attitude is significant for more than one country, but in general systematic patterns are difficult to discern.

Similar figures showing the role of attitudes of other respondents captured in the data show even fewer white cells than those presented for Year 8 students. These other groups included Year 4 students (for mathematics, science and reading), along with the decomposition of achievement according to teacher attitudes and practices.

Figure 11 shows the distribution of contributions from differences in attitudes towards differences in achievement in mathematics between Victorian Year 8 students and those from high performing countries. The distribution is skewed into negative territory, indicating that in general, differences in attitudes exist in the wrong direction to help explain differences in achievement very much. That is, Victorian students tend to have more positive attitudes towards mathematics than do students in high performing countries despite comparatively low performance.

Figure 12 shows the distribution of the K-statistic (described in the Appendix), a statistic that tells us how similar the associations between mathematics attitudes and achievement are for Victorian Year 8 students and those from high performing countries. Values around zero indicate that the associations are the same in Victoria as in the other countries. The modal value is very close to zero, but the distribution is skewed somewhat to the right, which suggests that there is a tendency for the gap in achievement associated with the most and least positive attitude responses to be greater in high performing countries than it is in Victoria.

4.3. Robustness of general results

In this section we discuss three sets of results. The first involves the results from the analysis that attempts to remove any effects from potential differences in linguistic meaning between the responses given in different countries as well as any cultural factors that might make some responses more socially acceptable than others. The second set of results looks at whether there might still be a relationship between attitudes and achievement that could still be exploited within countries to improve achievement. That is, even if differences in attitudes do not really explain the differences in achievement between students in Victoria and high performing countries, it may still be the case that just improving attitudes among Victorian students might still have the potential to improve achievement too. The third set of results are the decomposition estimates generated where we use the estimated *like*, *confident*, *engaged* and *value* scales simultaneously, rather than every individual item that goes into the scales.

The approach to dealing with the first issue (described on page 8) has already been described in Section 3.1 and involves a "double" differencing decomposition approach (Wellington 1993 and described in Appendix A). Essentially, we decompose the gap between a high performing country's difference in average achievement between mathematics and science from that of Victoria's into a part that reflects differences in attitudes, a part that reflects differences in associations between achievement and attitudes, and a part unrelated to both. We do this for each attitude and for the comparison between Victoria and each high performing country. One way of summarising the results is just to show the distribution of all the estimated contributions associated with differences in attitudes. This is done in Figure 13, where the Years 4 and 8 distributions are shown. The contributions have been normalised to a proportion by dividing them by the gaps in the difference in average achievement between mathematics and science between high performing countries and Victoria. This presentational approach essentially asks, is the difference in average achievement between mathematics and science larger (smaller) in Hong Kong than Victoria, for example, because the difference in attitudes towards mathematics and science are larger (smaller) there? A positive estimate of the attitude effect would be consistent with attitudes being important in explaining differences between countries in average achievement.

In Figure 13, while the Year 4 distribution is centred on zero, most of the Year 8 values are negative. That is, the contribution of attitudes appears to be largely zero or move in the wrong direction for attitudes to play much of a role in explaining differences in average achievement between countries. The difference in average achievement between mathematics and science are not larger (smaller) in Hong Kong than Victoria because the difference in attitudes towards mathematics and science are larger (smaller) there. Instead, the difference in average achievement between mathematics and science appears largely independent of differences in attitudes or the way they are associated with achievement between countries.

The second piece of analysis involves estimation of a series of regression equations. A fixed effects specification is used in the regression of achievement in both mathematics and science on student attitudes towards those subjects. The approach is specified in the Appendix and makes use of the fact that the wording of the attitude questions for mathematics and science are identical, except for the change in the name of the subject. Separate equations are estimated for each similarly worded attitude, and for the scales available in the data that summarises the individual measures, along with one specification that includes all of the scales.

The feature of the fixed effects specification that is important is that it removes the impact of common factors that might influence the reported attitudes to both mathematics and science, such as individual ability and motivation, family background and personality factors that might induce individuals to be similarly positive or negative about both subjects. The approach changes the nature of the comparison being made. A simple regression involves a comparison between individuals with positive attitudes (whose achievement will be higher) with individuals with more negative attitudes (whose achievement will tend to be lower). But these individuals may also be different in other ways, not just in that their attitudes differ, and this "unobserved heterogeneity" may result in too large a change in achievement being attributed to differences in attitudes than is really the case. In contrast, the fixed effects estimation involves a comparison within individuals, asking for those individuals whose attitudes towards mathematics and science appeared to differ, how much did their achievement in the two subjects differ. This seems much closer to the research question actually being addressed here, of how much achievement might increase if only student attitudes could be increased. A common feature of regression estimates that allow for such fixed effects that deal with unobserved heterogeneity between individuals is that they point to weaker relationships between phenomena than is apparent from simple estimates that make comparisons across individuals.

The results of this exercise for Years 4 and 8 are presented in Tables 2 and 3 respectively. To simplify the presentation, the attitude variables were specified in a linear form, with the value 1 representing the least positive response to the prompt and the value 4 representing the most positive. If a parameter value on an attitude variable was of the order of 30, this would mean that the difference in achievement values associated with improved attitudes varying between the bottom and the top response would be of the order of 90 points, a range in achievement a little smaller than those apparent for the attitudes shown in Figures 3 and 5 for Victorian students, for example.

The three columns of numbers in Tables 2 and 3 contain the fixed effect regression estimate, its standard error and its value as a proportion of the average estimate if the equation was estimated separately for mathematics and science, without allowing for the fixed effects for individuals that influence their responses to questions about both subjects. The third column essentially provides an indication of how misleading simple analyses of the relationship between attitudes and achievement are, and the first column what a more plausible estimate of how achievement might change with improved attitudes.

The estimated parameters in Tables 2 and 3 are much smaller than 30, with only two being greater than 15. Many are substantially smaller and some, notably those that deal with the value of studying mathematics or science among Year 8 students, are not significantly different from zero. The largest effects involve attitudes associated with the *like* and *confidence* scales. For the Year 8 estimates, the parameters in the individual mathematics and science equations were often of a similar magnitude, with the fixed effects estimate being substantially smaller, generally less than half that magnitude. For Year 4, the parameters for the science equations were typically much smaller than those for mathematics, such that the fixed effects parameters were much closer to the average of the other two, hence the values in the third column of Table 2 are often larger than those in the third column of Table 3.

In general, the magnitudes of the parameters point to modest potential achievement gains from improving the attitudes of Victorian students. The largest effect of any individual parameter of moving students up one category would be of the order of 15 points, far less than the gap between the high performing countries and Victoria. The jointly estimated scale parameters tell us what the effect of changing all the attitudes might be. If it was possible to improve attitudes by as much as a standard deviation (a change of three units in the scales), average achievement might increase by somewhere from 18-21 points, putting Victoria still well short of the performance of the high performing countries.

The results for the third piece of analysis are summarised in Figure 14. It shows the decomposition estimates when we use the estimated *like*, *confident*, *engaged* and *value* scales simultaneously for each country, rather than the individuals item that go into the scales. The results are qualitatively similar to those already presented. Attitudes work in the wrong direction to contribute to the gaps in achievement between Victoria and high performing countries and most of the gaps in achievement reflect the constant or residual, that is, from factors independent of the attitudes captured in the scales.

4.4. Where there might be scope to influence achievement

With so many "results", it is possible that some important, and potentially exploitable, relationships between attitudes and achievement might have been overlooked. In general, though, there seem to be few regularities in the analysis of the gaps in achievement between the high performing countries and Victoria and whether any part of the gaps might be explained by differences in student attitudes towards mathematics and science.

One set of estimates that is close to consistent points to differences in attitudes to science in Year 8 between Victorian and Singaporean students. These are evident in Figure 10 and relate to differences in attitudes used in the *like* and *value* science scales. While the effects are significant, they are also quite small, with the attitude effects contributing less than 20 per cent of the achievement gap between Victoria and Singapore in both the *like* and *value* science items.[8]

The analysis has thrown up some curiosities, notably in how confident Victorian students are about their learning in mathematics and science compared to students in high performing countries.[9] Such confidence clearly seems misplaced, but there is no evidence that it actually gets in the way of improving the skills of Victorian students. And to the extent that the differences reflect classroom pressures and reduced student well-being in high performing countries, there may be no call to address this as a particular "problem".

5. Concluding remarks

Differences in attitudes towards mathematics and science between Victorian students and those in high performance countries are not important in explaining differences in achievement. Victorian students report very positive attitudes towards their subjects, often more positive attitudes than students in high performing countries, even if these attitudes are generally a bit less positive than those in other countries not part of the top tier. Succeeding in getting them to be even more positive towards mathematics and science does not appear to be a potentially fruitful way of closing the gap in achievement with students in high performing

---

[8] Another result, not presented here, is that aspirations for future educational attainment are positively related to achievement and tend to be slightly higher in high achievement countries among Year 8 students than in Victoria. Educational aspirations are clearly important, and translate into actual attainment (Homel & Ryan 2014), so strategies to raise them, particularly among students from disadvantaged backgrounds, are important in their own right. The patterns observed here likely reflect reverse causality (achievement determines aspirations, with any causal effect operating in the other direction likely to be much weaker) and how selective entrance to university is in the different countries.

[9] In this regard, Victorian students are representative of those in other Australian states, who are similarly confident.

countries. At least this negative result rules out using attitudes as a vehicle through which to influence achievement.

Instead, research on interventions in other areas should be pursued. It is possible that evidence for or against such potential interventions can be found in international achievement test data, just not from the attitudinal data they contain. As was noted in the introduction, many countries have responded to the international test results by introducing reforms to their education systems. A potentially productive direction for research on how to improve achievement would be to document these reforms in a very detailed way so that it is possible to assess how different types of reforms within countries influence achievement relative to the high performing countries.

References

Aitken, JR (1969). *A study of attitudes and attitudinal change of institutionalized delinquents through group guidance techniques.* Unpublished Ed.D., The University of Southern Mississippi, MS.

Altonji, JG & Blank, RM, (1999), "Race and gender in the labor market", in Ashenfelter, OC. & Card, D (Eds.), *Handbook of Labor Economics*, vol. 3. Elsevier.

Ammermueller, A (2007), "PISA: What makes the difference? Explaining the gap in test scores between Finland and Germany", *Empirical Economics*, 33, pp. 263 - 287.

Blau, D, & Currie, J (2006), "Pre-school, day care, and after-school care: who's minding the kids?." *Handbook of the Economics of Education* 2: 1163-1278.

Blinder, A (1973), 'Wage discrimination: reduced form and structural estimates', *Journal of Human Resources*, 8, 436–455.

Currie, J (2001), "Early childhood education programs." *Journal of Economic perspectives*: 213-238.

Debeer, D, Buchholz, J, Hartig J & Janssen, R, (2014), "Student, School, and Country Differences in Sustained Test-Taking Effort in the 2009 PISA Reading Assessment", *Journal of Educational and Behavioral Statistics,*39(6), 502–523.

Feniger, Y & Lefstein, A (2014), "How not to reason with PISA data: an ironic investigation." *Journal of Education Policy* 29.6:845-855.

Fortin, N., Limieux, T. & Firpo, S. (2011) 'Decomposition Methods in Economics', in Ashenfelter, O & Card D. (eds), *Handbook of Labor Economics*, Vol IVa, Elsevier, Amsterdam.

Fuchs, T. & Woessmann, L. (2007), "What accounts for international differences in student performance? A re-examination using PISA data ", *Empirical Economics*, 32, pp. 433 - 464.

Gardeazabal, J & Ugidos, A (2004), 'More on identification in detailed wage decompositions', *The Review of Economics and Statistics,* 86, 1034–1036.

Hanushek, EA (1996),"School resources and student performance." *Does money matter,* 43-73.

Hanushek, EA (2006),"Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries*", *The Economic Journal,* 116, 510 C63-C76.

Hanushek, EA & Woessman, L. (2011) 'The Economics of International Differences in Educational Achievement', in Hanushek, EA, Machin, S & Woessman, L (eds), *Handbook of the Economics of Education*, Vol III, North Holland, Amsterdam.

Hanushek, EA, Link, S & Woessmann, L, (2013),"Does school autonomy make sense everywhere? Panel estimates from PISA", *Journal of Development Economics,* 104, 212-232.

Hattie, J (2009), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hedges, LV, & Greenwald, R, (1996), "Have times changed? The relation between school resources and student performance." *Does money matter* 74-92.

Homel, J. & Ryan C. (2014), *Educational outcomes: the impact of aspirations and the role of student background characteristics* Longitudinal Surveys of Australian Youth Research Report No. 65, National Centre for Vocational Education Research, Adelaide

Houtenville, AJ, & Smith Conway, K (2008), "Parental effort, school resources, and student achievement." *Journal of Human resources* 43.2: 437-453.

Krieg, JM & Storer, P (2006), "How much do students matter? Applying the Oaxaca decomposition to explain determinants of adequate yearly progress", *Contemporary Economic Policy* 24, 563–581.

Krueger, AB (1999), "Experimental Estimates of Education Production Functions."*Quarterly Journal of Economics,* 497-532.

Ma, X & Kishor, N. (1997), "Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis." *Journal for Research in Mathematics Education*, 28(1): 26-47.

Ma, X & Xu, J (2004a), "Determining the Causal Ordering between Attitude toward Mathematics and Achievement in Mathematics" *American Journal of Education*, 110(3), 256-280

Ma, X & Xu, J (2004b), "The causal ordering of mathematics anxiety and mathematics achievement: a longitudinal panel analysis" *Journal of Adolescence,* 27, 165–179

McEwan, PJ & Marshall, JH (2004), 'Why does academic achievement vary across countries? Evidence from Cuba and Mexico', *Education Economics*, 12, 205–217

Marsh, HW & Martin, AJ (2011), 'Academic self-concept and academic achievement: Relations and causal ordering', *British Journal of Educational Psychology*, 81, 59–77

Neale, DD (1969), "The role of attitudes in learning mathematics", *Arithmetic Teacher*, 16(8): 631-640.

Oaxaca, R (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review,* 14, 693–709.

Schütz, G (2009), *Does the quality of pre-primary education pay off in secondary school? An international comparison using PISA 2003*.. Ifo Working Paper No. 68.

Schütz, G, Ursprung, HW & Woessmann, L (2008), "Education policy and equality of opportunity." *Kyklos,* 61, 279-308.

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

Thomson S, Hillman, K, Wernert, N, Schmid M, Buckley, S & Munene, A (2012b) *Monitoring Australian Year 4 student achievement internationally: TIMSS and PIRLS 2011*, Australian Council for Educational Research, Camberwell.

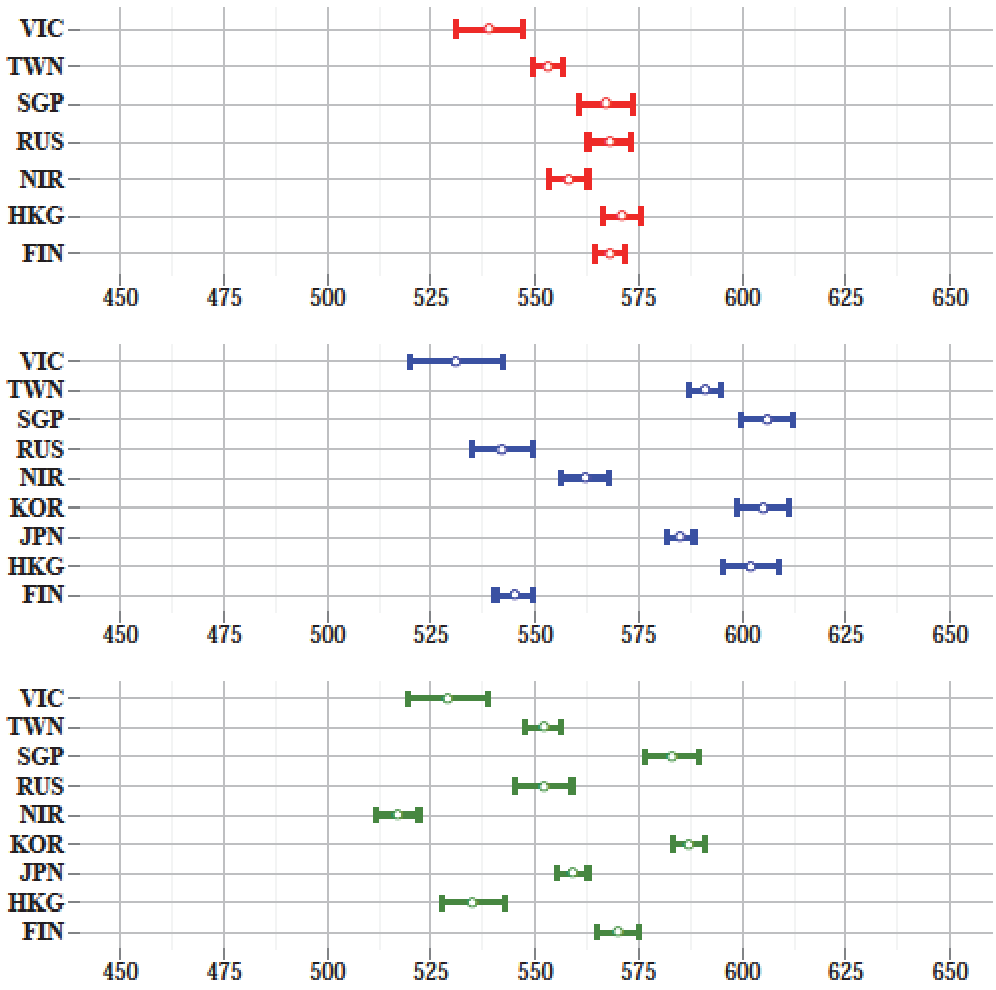
Wellington, AJ (1993), 'Changes in the Male/Female Wage Gap, 1976-85', *Journal of Human Resources*, 28, 383–411.

Woessmann, L (2005), "The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA." *Education Economics* 13.2: 143-169.

Woessmann, L (2009), "Public-private partnerships and student achievement: A cross-country analysis." In: Chakrabarti, R, Peterson, PE (Eds.), *School Choice International: Exploring Public-Private Partnerships.* The MIT Press, pp. 13-45.

Woessmann, L, Luedemann, E, Schuetz, G & West, MR, (2009), *School Accountability, Autonomy and Choice Around the World*, Edward Elgar.

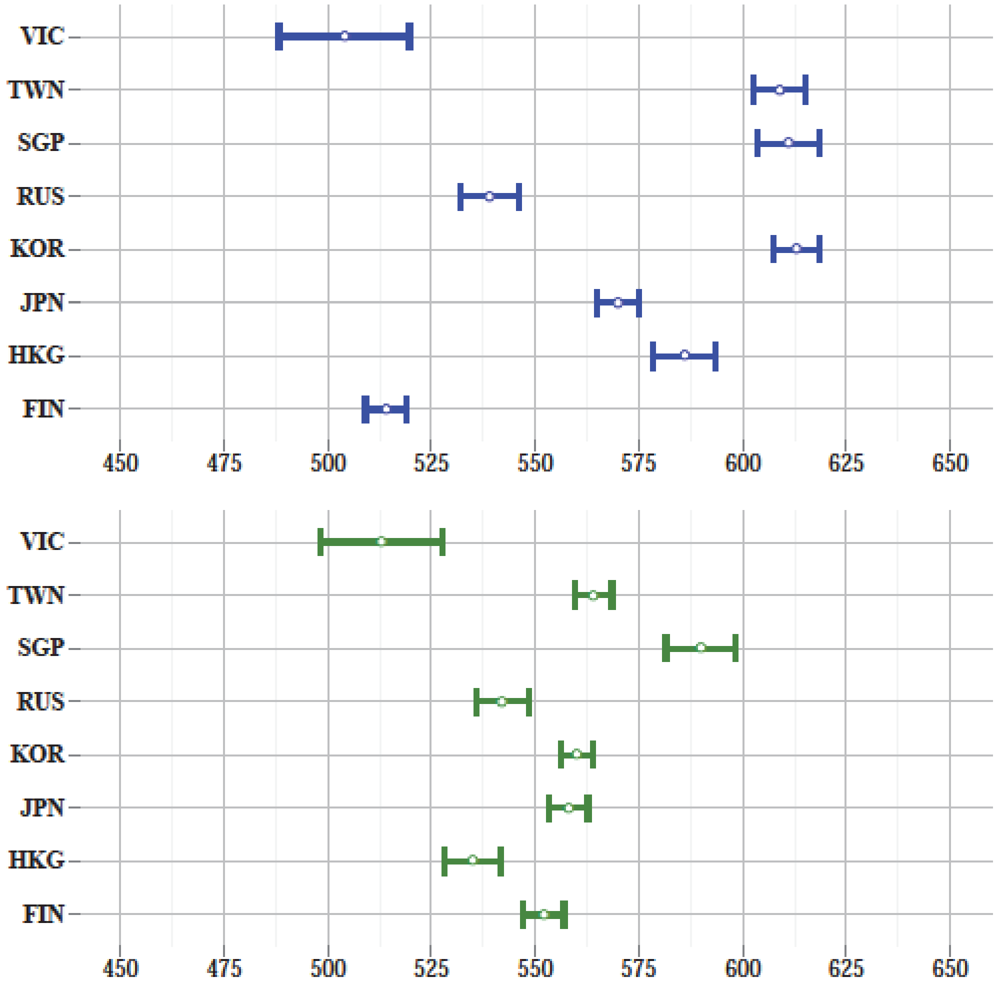
Figure 1: Average achievement[a]: Victoria versus high performing countries in reading (top panel), mathematics (middle panel) and science (lower panel), Year 4 2011



Source: PIRLS/TIMSS 2011

(a) The 95 per cent confidence intervals shown around the average values were estimated taking into account the complex survey design used for each country via use of jackknife repeated replication (JRR) methodology.

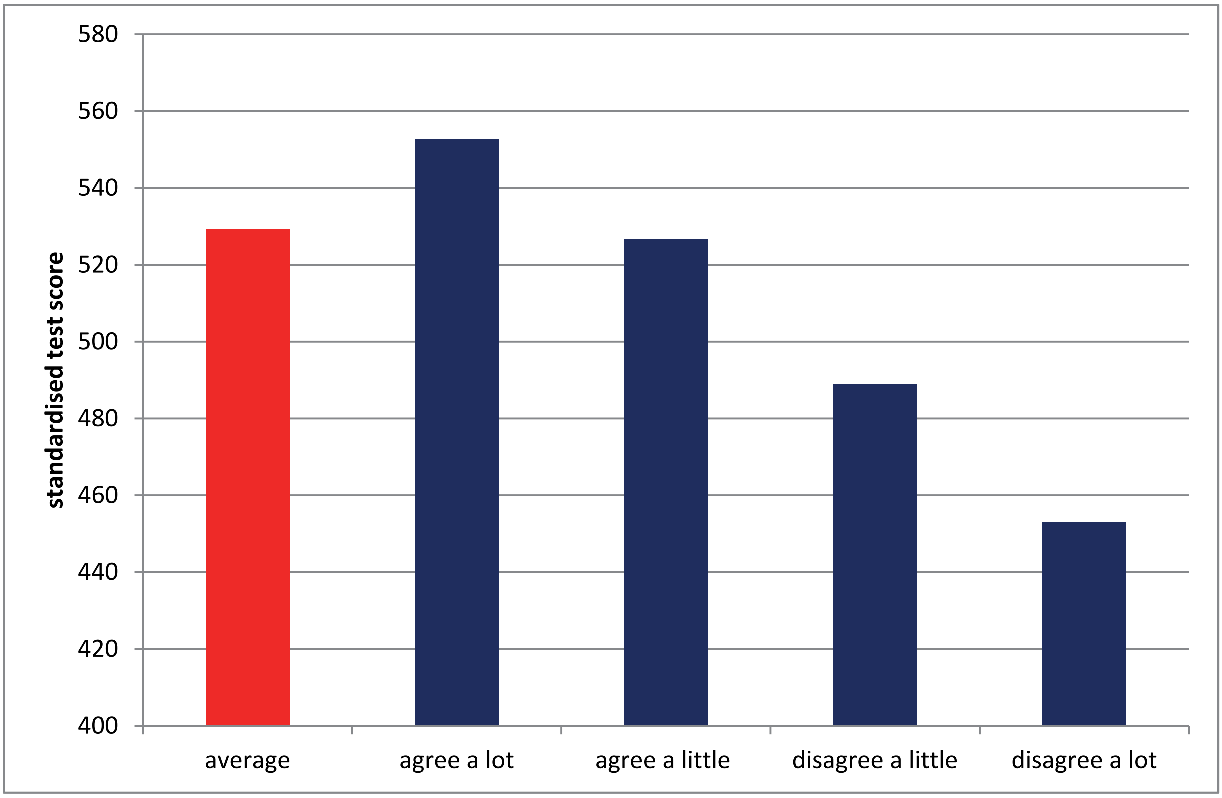
Figure 2: Average achievement[(a)]: Victoria versus high performing countries in mathematics (top panel) and science (lower panel), Year 8 2011



Source: TIMSS 2011

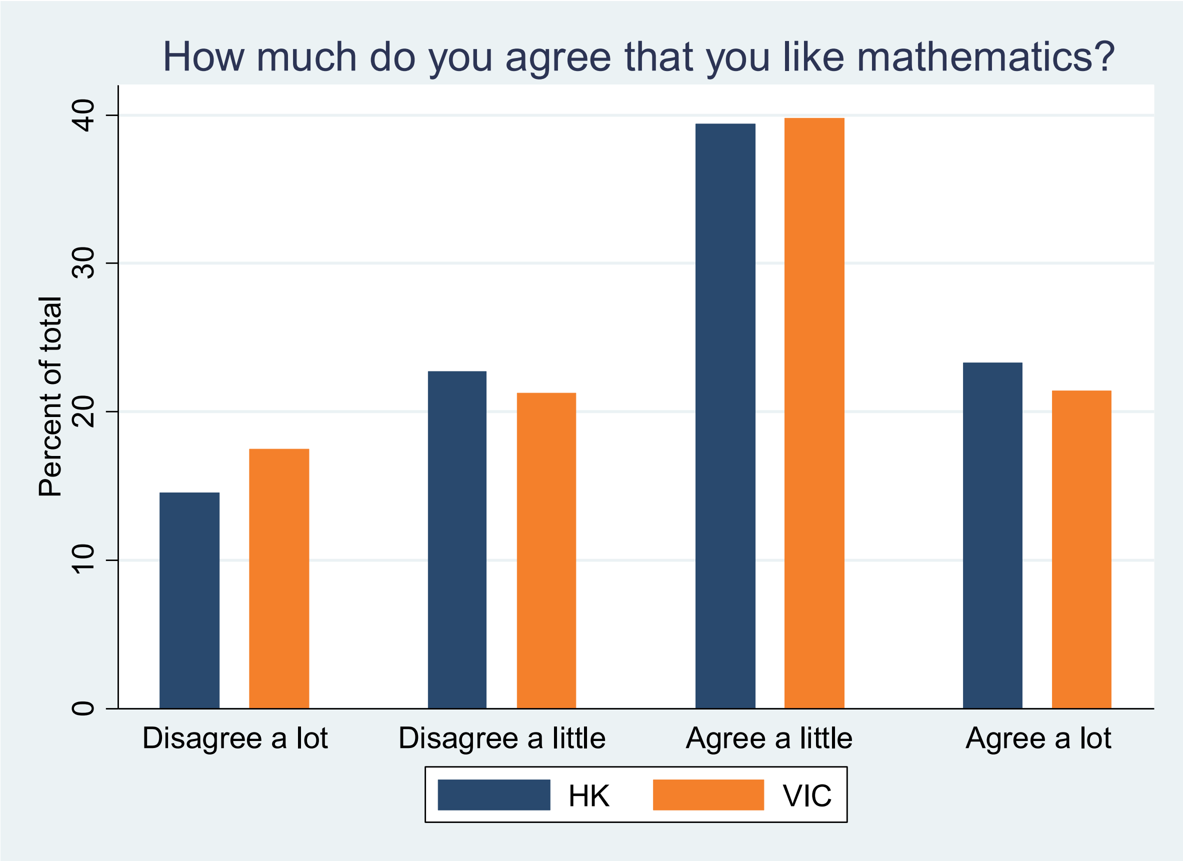
(a) The 95 per cent confidence intervals shown around the average values were estimated taking into account the complex survey design used for each country via use of jackknife repeated replication (JRR) methodology.

Figure 3: Achievement levels[a] associated with student responses to the statement "It is important to do well in mathematics", Victorian students in Year 4 2011



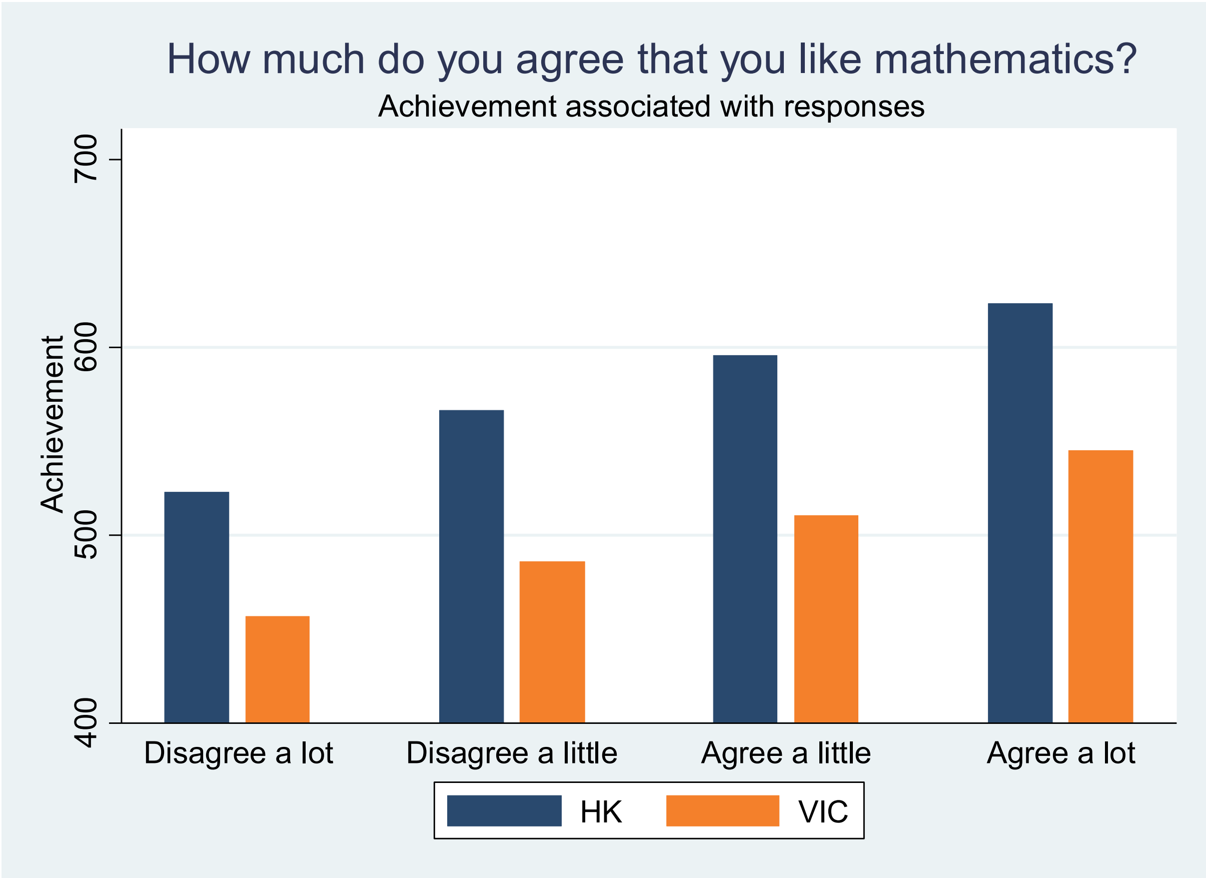
Source: TIMSS 2011
(a) The "average" bar shows the average mathematics achievement level of Victorian Year 4 students. It is not a response category to the question asked.

Figure 4: Victorian and Hong Kong student responses about liking mathematics, Year 8 2011



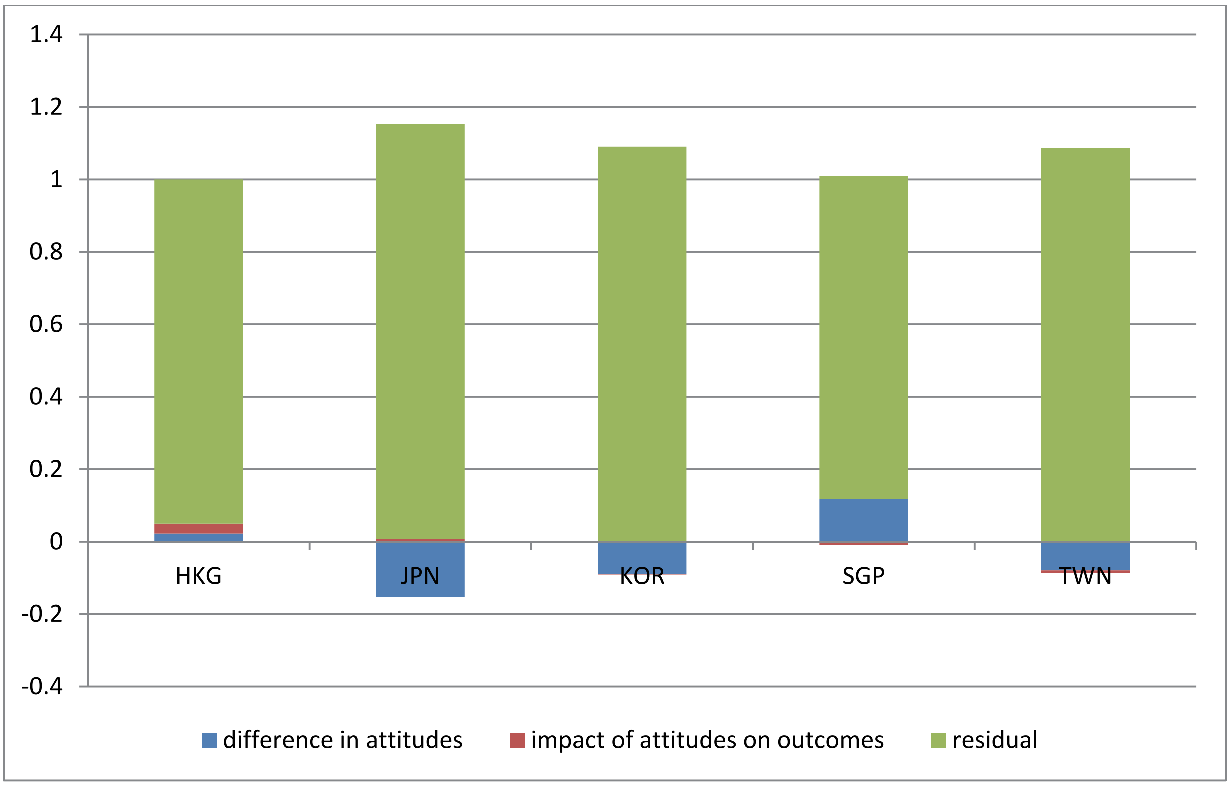
Source: TIMSS 2011

Figure 5: Average mathematics achievement by response category to question about liking mathematics - Victorian and Hong Kong, Year 8 2011



Source: TIMSS 2011

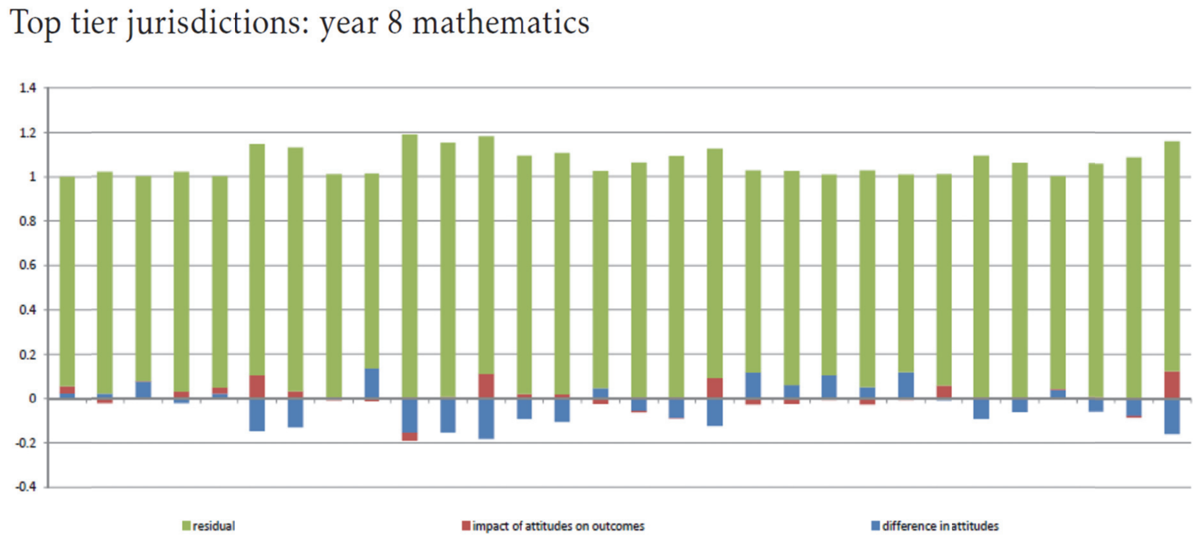
Figure 6: Decomposition results for students in high performing countries compared to Victorian Year 8 responses to "How much do you agree that you like mathematics?"[(a)]



Source: data from TIMSS 2011, authors' calculations

(a) The vertical axis shows the contribution of each element of the decomposition to the difference in achievement between Victorian students and those in other countries, divided by the observed difference in achievement between the sets of students. Negative values mean either the differences in attitudes or their impact on outcomes worked in favour of Victorian students having higher achievement relative to students in the relevant countries than they actually did. Such negative values push the contribution from the difference in the residuals or constants of the two countries above unity.

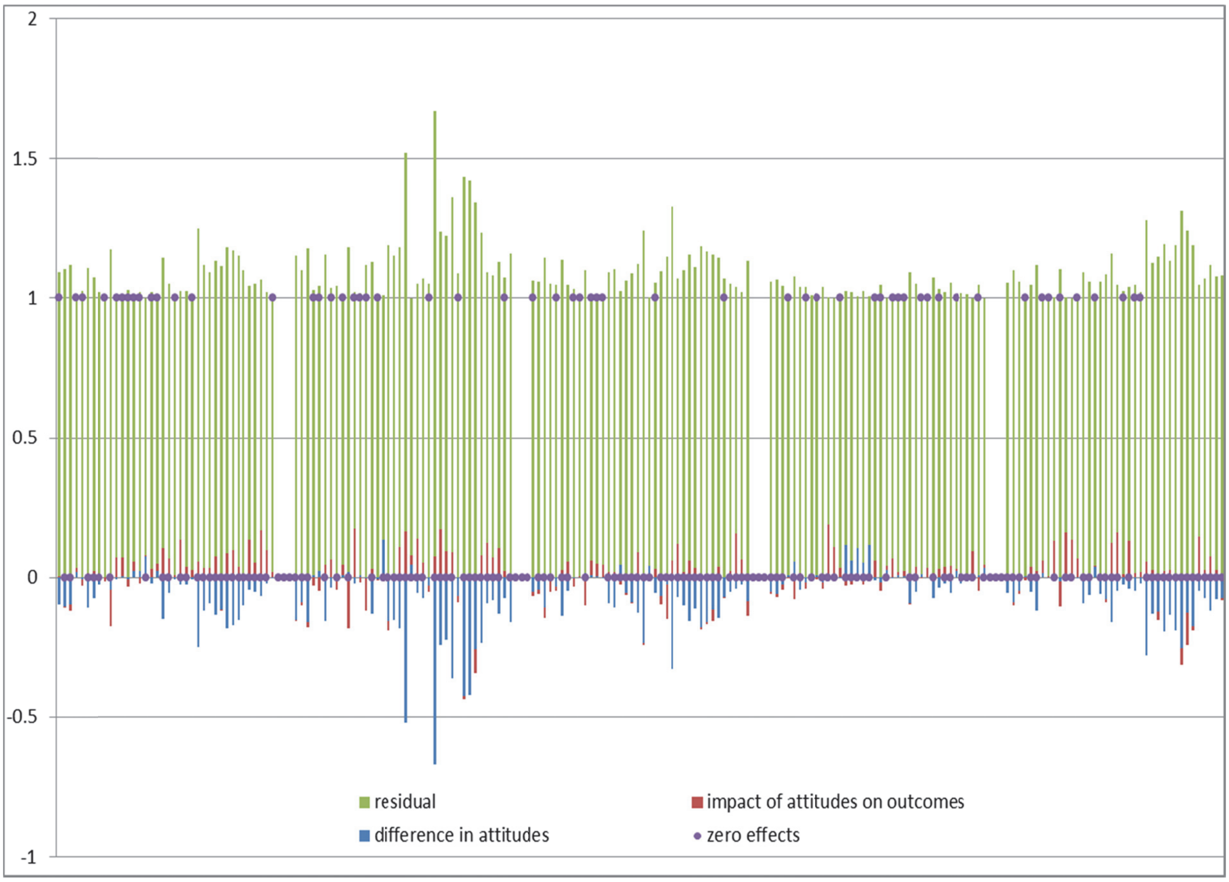
Figure 7: Decomposition results for students in high performing countries compared to Victorian Year 8 responses to questions that form the TIMSS *like* mathematics scale



Notes: Six questions for each jurisdiction form a battery of items that measure student attitudes toward mathematics.

(a) The vertical axis shows the contribution of each element of the decomposition to the difference in achievement between Victorian students and those in other countries, divided by the observed difference in achievement between the sets of students. Negative values mean either the differences in attitudes or their impact on outcomes worked in favour of Victorian students having higher achievement relative to students in the relevant countries than they actually did. Such negative values push the contribution from the difference in the residuals or constants of the two countries above unity.

Figure 8: Decomposition results for students in high performing countries compared to Victorian Year 8 responses to questions that form the TIMSS *confidence*, *like*, *engaged* and *value* mathematics scales



(a) The vertical axis shows the contribution of each element of the decomposition to the difference in achievement between Victorian students and those in other countries, divided by the observed difference in achievement between the sets of students. Negative values mean either the differences in attitudes or their impact on outcomes worked in favour of Victorian students having higher achievement relative to students in the relevant countries than they actually did. Such negative values push the contribution from the difference in the residuals or constants of the two countries above unity.

Figure 9: Mathematics – significance of decomposition results for students in high performing countries compared to Victorian Year 8 students by country

lable var	jurisvar	HKG	JPN	KOR	SGP	TWN
fraction of questions for which hypothesis of zero explanatory power rejected		0.02	0.02	0.00	0.10	0.02
GEN\HOW OFTEN\HOME\PARENTS ASK LEARNING	BSBG11A					
GEN\HOW OFTEN\HOME\TALKING ABOUT SCHOOL	BSBG11B					
GEN\HOW OFTEN\HOME\PARENTS MAKE SURE	BSBG11C					
GEN\HOW OFTEN\HOME\PARENTS CHECK HOMEWOR	BSBG11D					
GEN\AGREE\BEING IN SCHOOL	BSBG12A					
GEN\AGREE\SAFE AT SCHOOL	BSBG12B					
GEN\AGREE\BELONG AT SCHOOL	BSBG12C					
GEN\HOW OFTEN\MADE FUN OF	BSBG13A					
GEN\HOW OFTEN\LEFT OUT OF GAMES	BSBG13B					
GEN\HOW OFTEN\SPREADING LIES ABOUT ME	BSBG13C					
GEN\HOW OFTEN\STEALING STH FROM ME	BSBG13D					
GEN\HOW OFTEN\HURT BY OTHERS	BSBG13E					
GEN\HOW OFTEN\FORCE TO DO STH	BSBG13F					
MAT\AGREE\ENJOY LEARNING MATHEMATICS	BSBM14A					
MAT\AGREE\WISH HAVE NOT TO STUDY MATH	BSBM14B					
MAT\AGREE\MATH IS BORING	BSBM14C					
MAT\AGREE\LEARN INTERESTING THINGS	BSBM14D					
MAT\AGREE\LIKE MATHEMATICS	BSBM14E					
MAT\AGREE\IMPORTANT TO DO WELL IN MATH	BSBM14F					
MAT\AGREE\TEACHER EXPECTS TO DO	BSBM15A					
MAT\AGREE\THINGS NOT RELATED TO LESSON	BSBM15B					
MAT\AGREE\TEACHER IS EASY TO UNDERSTAND	BSBM15C					
MAT\AGREE\INTERESTED IN WHAT TCHR SAYS	BSBM15D					
MAT\AGREE\INTERESTING THINGS TO DO	BSBM15E					
MAT\AGREE\USUALLY DO WELL IN MATH	BSBM16A					
MAT\AGREE\MATHEMATICS IS MORE DIFFICULT	BSBM16B					
MAT\AGREE\MATHEMATICS NOT MY STRENGTH	BSBM16C					
MAT\AGREE\LEARN QUICKLY IN MATHEMATICS	BSBM16D					
MAT\AGREE\MAT MAKES CONFUSED AND NERVOUS	BSBM16E					
MAT\AGREE\GOOD AT WORKING OUT PROBLEMS	BSBM16F					
MAT\AGREE\I CAN DO WELL IN MATHEMATICS	BSBM16G					
MAT\AGREE\I AM GOOD AT MATHEMATICS	BSBM16H					
MAT\AGREE\MATHEMATICS HARDER FOR ME	BSBM16I					
MAT\AGREE\MATHEMATICS WILL HELP ME	BSBM16J					
MAT\AGREE\NEED MAT TO LEARN OTHER THINGS	BSBM16K					
MAT\AGREE\NEED MATH TO GET INTO <UNI>	BSBM16L					
MAT\AGREE\NEED MAT TO GET THE JOB I WANT	BSBM16M					
MAT\AGREE\JOB INVOLVING MATHEMATICS	BSBM16N					

NOTES:

- cannot reject at 95% confidence interval the joint hypothesis that differences in country outcomes are unrelated to differential responses to attitudes or to the relationship between attitudes and outcomes
- cannot reject at the 95% confidence interval the hypothesis that differences in country outcomes are equal to differences in intercept terms of associated linear regression equations
- reject both hypotheses identified above, but relative responses to question exaggerate reported differences in test outcomes between jurisdictions
- reject the above options, but differences between jurisdictions in the relationship between the question and test outcomes work against explaining differences between the jurisdictions

Figure 10: Science – significance of decomposition results for students in high performing countries compared to Victorian Year 8 students by country

lable var	jurisvar	JPN	KOR	SGP	TWN	TWN
fraction of questions for which hypothesis of zero explanatory power rejected		0.05	0.00	0.28	0.05	0.05
GEN\HOW OFTEN\HOME\PARENTS ASK LEARNING	BSBG11A					
GEN\HOW OFTEN\HOME\TALKING ABOUT SCHOOL	BSBG11B					
GEN\HOW OFTEN\HOME\PARENTS MAKE SURE	BSBG11C					
GEN\HOW OFTEN\HOME\PARENTS CHECK HOMEWOR	BSBG11D					
GEN\AGREE\BEING IN SCHOOL	BSBG12A					
GEN\AGREE\SAFE AT SCHOOL	BSBG12B					
GEN\AGREE\BELONG AT SCHOOL	BSBG12C					
GEN\HOW OFTEN\MADE FUN OF	BSBG13A					
GEN\HOW OFTEN\LEFT OUT OF GAMES	BSBG13B					
GEN\HOW OFTEN\SPREADING LIES ABOUT ME	BSBG13C					
GEN\HOW OFTEN\STEALING STH FROM ME	BSBG13D					
GEN\HOW OFTEN\HURT BY OTHERS	BSBG13E					
GEN\HOW OFTEN\FORCE TO DO STH	BSBG13F					
SCI\AGREE\ENJOY LEARNING SCIENCE	BSBS17A					
SCI\AGREE\WISH HAVE NOT TO STUDY SCIENCE	BSBS17B					
SCI\AGREE\READ ABOUT SCIENCE SPARE TIME	BSBS17C					
SCI\AGREE\SCIENCE IS BORING	BSBS17D					
SCI\AGREE\LEARN INTERESTING THINGS	BSBS17E					
SCI\AGREE\LIKE SCIENCE	BSBS17F					
SCI\AGREE\IMPORTANT TO DO WELL IN SCI	BSBS17G					
SCI\AGREE\TEACHER EXPECTS TO DO	BSBS18A					
SCI\AGREE\THINGS NOT RELATED TO LESSON	BSBS18B					
SCI\AGREE\TEACHER EASY TO UNDERSTAND	BSBS18C					
SCI\AGREE\INTERESTED IN WHAT TCHR SAYS	BSBS18D					
SCI\AGREE\INTERESTING THINGS TO DO	BSBS18E					
SCI\AGREE\USUALLY DO WELL IN SCIENCE	BSBS19A					
SCI\AGREE\SCIENCE IS MORE DIFFICULT	BSBS19B					
SCI\AGREE\SCIENCE NOT MY STRENGTH	BSBS19C					
SCI\AGREE\LEARN QUICKLY IN SCIENCE	BSBS19D					
SCI\AGREE\SCI MAKES CONFUSED AND NERVOUS	BSBS19E					
SCI\AGREE\GOOD AT WORKING OUT PROBLEMS	BSBS19F					
SCI\AGREE\I CAN DO WELL IN SCIENCE	BSBS19G					
SCI\AGREE\I AM GOOD AT SCIENCE	BSBS19H					
SCI\AGREE\SCIENCE IS HARDER FOR ME	BSBS19I					
SCI\AGREE\SCIENCE WILL HELP ME	BSBS19J					
SCI\AGREE\NEED SCI TO LEARN OTHER THINGS	BSBS19K					
SCI\AGREE\NEED SCI TO GET INTO <UNI>	BSBS19L					
SCI\AGREE\NEED SCI TO GET THE JOB I WANT	BSBS19M					
SCI\AGREE\JOB INVOLVING SCIENCE	BSBS19N					

NOTES:

- cannot reject at 95% confidence interval the joint hypothesis that differences in country outcomes are unrelated to differential responses to attitudes or to the relationship between attitudes and outcomes
- cannot reject at the 95% confidence interval the hypothesis that differences in country outcomes are equal to differences in intercept terms of associated linear regression equations
- reject both hypotheses identified above, but relative responses to question exaggerate reported differences in test outcomes between jurisdictions
- reject the above options, but differences between jurisdictions in the relationship between the question and test outcomes work against explaining differences between the jurisdictions

Figure 11: Mathematics – distribution of contribution of attitudes from decomposition results, all countries, Year 8 2011

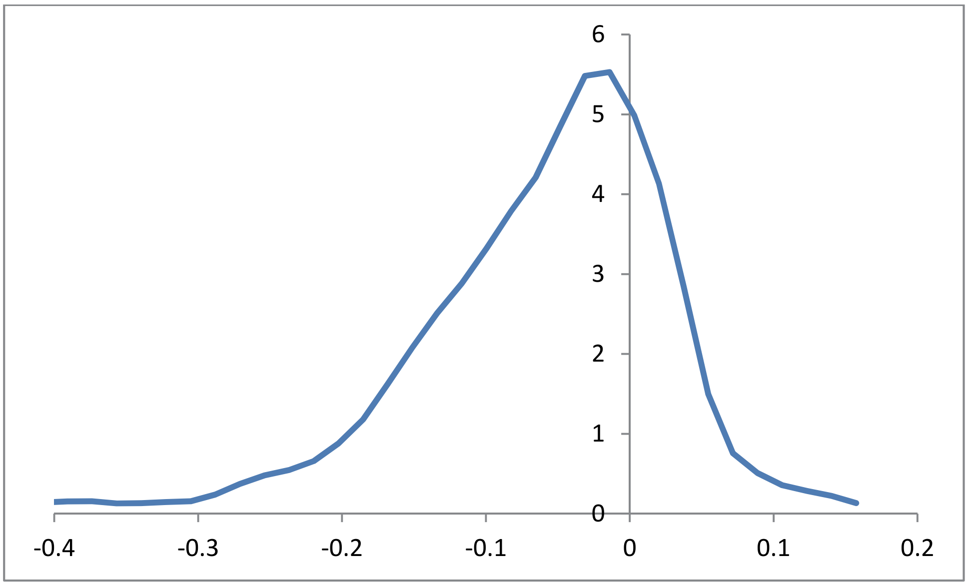
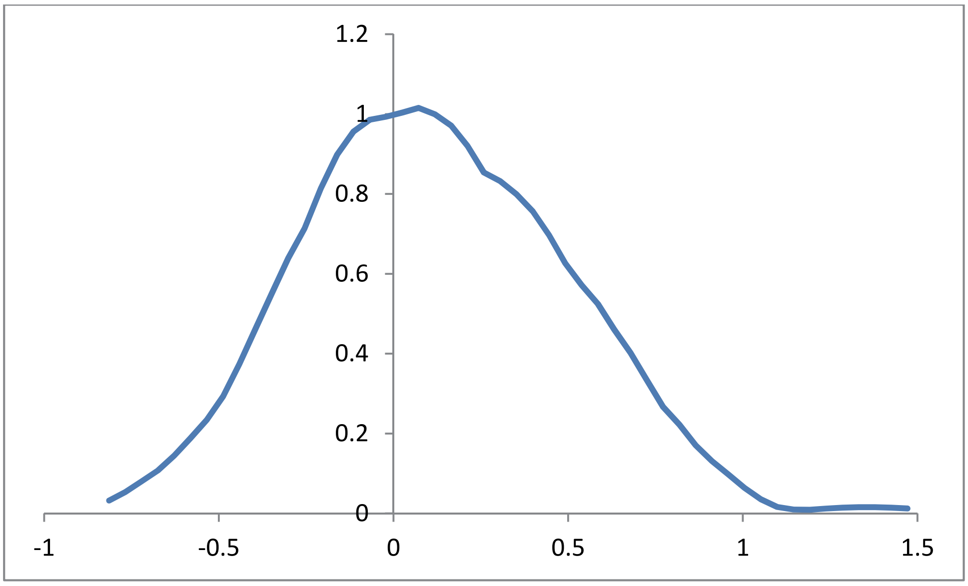


Figure 12: Mathematics – distribution of statistic showing similarity of relationship between attitudes and achievement in high performing countries and Victoria, Year 8 2011



Notes: Zero implied the same relationship, positive (negative) values that achievement increases more (less) in high performing countries with the attitude responses than in Victoria.

Figure 13: Distribution of the contribution of attitudes to the gap in achievement between mathematics and science in high performing countries compared to Victoria, Years 4 and 8 2011

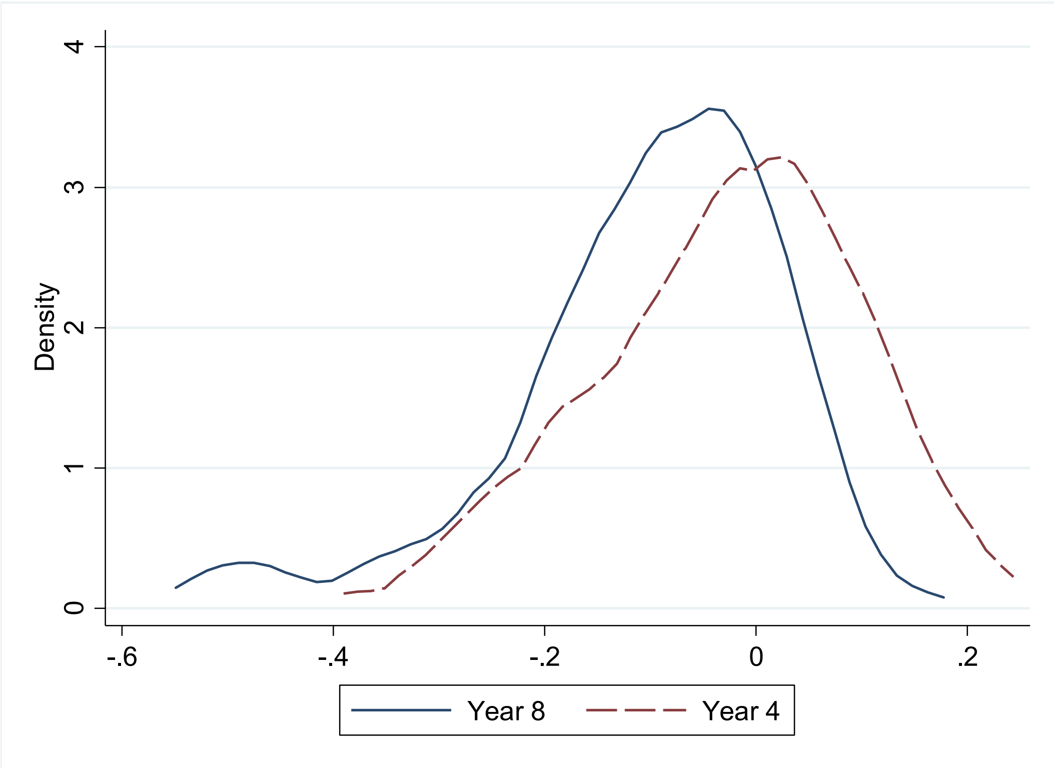
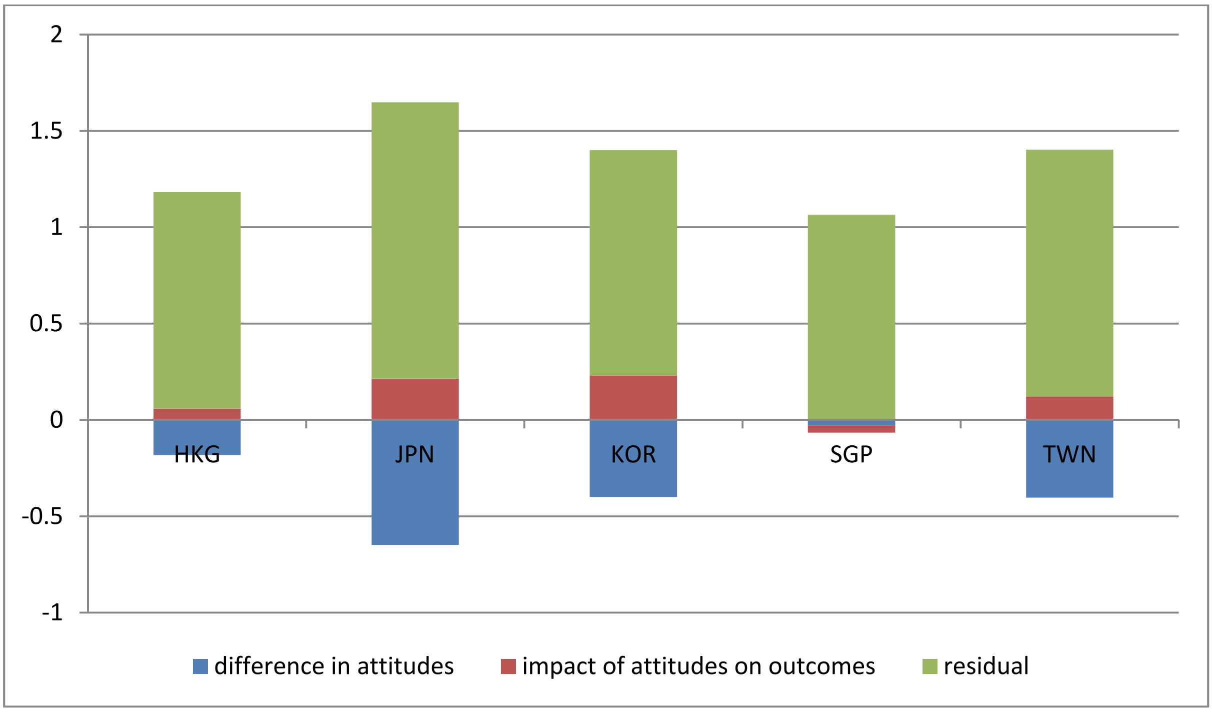


Figure 14: Decomposition results for students in high performing countries compared to Victorian Year 8 in mathematics, using the *confidence*, *like*, *engaged* and *value* scales simultaneously[(a)]



(a) The vertical axis shows the contribution of each element of the decomposition to the difference in achievement between Victorian students and those in other countries, divided by the observed difference in achievement between the sets of students. Negative values mean either the differences in attitudes or their impact on outcomes worked in favour of Victorian students having higher achievement relative to students in the relevant countries than they actually did. Such negative values push the contribution from the difference in the residuals or constants of the two countries above unity.

Table 1: Sample sizes - PIRLS/TIMSS 2011

	Schools		Students	
	Year 4	Year 8	Year 4	Year 8
Australia				
PIRLS	280		6126	
TIMSS	280	275	6146	7556
Victoria				
PIRLS	42		764	
TIMSS	42	43	760	958
Hong Kong				
TIMSS	136	117	3957	4015
Japan				
TIMSS	149	138	4411	4414
Singapore				
TIMSS	176	165	6368	5927
South Korea				
TIMSS	150	150	4334	5166
Taiwan				
TIMSS	150	150	4284	5042

Table 2: Fixed effects regression estimates of attitudes on achievement, Year 4

		Parameter	S.E.	FE Effect[a]
Enjoy Learning Mathematics/Science		13.0***	(1.4)	1.12
Wish Have Not To Study Mathematics/Science		7.8***	(1.3)	0.45
Mathematics/Science Is Boring		8.6***	(1.4)	0.61
Learn Interesting Things		10.8***	(1.6)	0.64
Like Mathematics/Science		10.6***	(1.3)	0.83
Important To Do Well In Mathematics/Science		8.1***	(2.1)	2.18
Teacher Expects To Do		6.2**	(2.5)	1.94
Things Not Related To Lesson		5.1***	(1.7)	0.32
Teacher Is Easy To Understand		7.0***	(2.4)	0.81
Interested In What Teacher Says		5.3**	(2.2)	0.66
Interesting Things To Do		6.3***	(2.1)	1.41
Usually Do Well In Mathematics/Science		13.1***	(1.6)	0.59
Mathematics/Science Harder For Me Than For Others		10.7***	(1.4)	0.43
Just Not Good At Mathematics/Science		12.1***	(1.4)	0.47
Learn Quickly In Mathematics/Science		11.3***	(1.5)	0.64
I Am Good At Mathematics/Science		12.0***	(1.5)	1.14
Mathematics/Science Is Harder For Me		10.5***	(1.1)	0.60
Separate:	Likes Mathematics/Science	5.6***	(0.6)	0.64
	Confidence In Mathematics/Science	7.2***	(0.6)	0.64
	Engaged In Mathematics/Science	4.1***	(1.0)	0.68
Joint estimate:	Likes Mathematics/Science	2.0**	(0.8)	1.59
	Confidence In Mathematics/Science	6.4***	(0.8)	0.70
	Engaged In Mathematics/Science	-1.5	(1.0)	-1.53

"***", "**" and "*" indicate significant at the 1, 5 and 10 per cent levels.
(a) This column is equal to the parameter divided by the average coefficient on the variable in the separate mathematics and science equations.

Table 3: Fixed effects regression estimates of attitudes on achievement, Year 8

		Parameter	S.E.	FE Effect[(a)]
Enjoy Learning Mathematics/Science		14.2***	(1.2)	0.50
Wish Have Not To Study Mathematics/Science		11.9***	(1.2)	0.46
Mathematics/Science Is Boring		11.1***	(1.3)	0.46
Learn Interesting Things		11.1***	(1.4)	0.46
Like Mathematics/Science		13.7***	(1.2)	0.47
Important To Do Well In Mathematics/Science		2.1	(1.7)	0.09
Teacher Expects To Do		6.4***	(1.9)	0.35
Things Not Related To Lesson		9.9***	(1.8)	0.62
Teacher Is Easy To Understand		5.1***	(1.3)	0.40
Interested In What Teacher Says		8.7***	(1.5)	0.49
Interesting Things To Do		8.3***	(1.3)	0.66
Usually Do Well In Mathematics/Science		15.8***	(1.4)	0.38
Mathematics/Science Is More Difficult		13.1***	(1.2)	0.38
Mathematics/Science Not My Strength		14.8***	(1.1)	0.44
Learn Quickly In Mathematics/Science		15.1***	(1.3)	0.40
Mathematics/Science Makes Confused And Nervous		12.9***	(1.4)	0.42
Good At Working Out Problems		13.3***	(1.4)	0.39
I Can Do Well In Mathematics/Science		11.4***	(1.5)	0.45
I Am Good At Mathematics/Science		8.4***	(1.4)	0.46
Mathematics/Science Harder For Me		13.0***	(1.0)	0.41
Mathematics/Science Will Help Me		2.2	(1.5)	0.13
Need Mathematics/Science To Learn Other Things		1.9	(1.6)	0.14
Need Mathematics/Science To Get Into University		1.7	(1.5)	0.09
Need Mathematics/Science To Get The Job I Want		1.3	(1.4)	0.10
Job Involving Mathematics/Science		6.3***	(1.3)	0.31
Separate:	Likes Mathematics/Science	8.0***	(0.7)	0.53
	Confidence In Mathematics/Science	7.4***	(0.6)	0.41
	Engaged In Mathematics/Science	5.1***	(0.8)	0.53
	Values Mathematics/Science	3.4*	(1.8)	0.14
Joint estimate:	Likes Mathematics/Science	5.6***	(1.0)	1.00
	Confidence In Mathematics/Science	5.0***	(0.7)	0.31
	Engaged In Mathematics/Science	-1.5*	(0.9)	0.44
	Values Mathematics/Science	-3.1*	(1.8)	-0.91

"***", "**" and "*" indicate significant at the 1, 5 and 10 per cent levels.
(a) This column is equal to the parameter divided by the average coefficient on the variable in the separate mathematics and science equations.

Appendix A: Technical Appendix

The Blinder-Oaxaca decomposition

Our interest is in the relationship between attitudes and achievement

(1) $y_i = \delta_1 I_1 + \delta_2 I_2 + \delta_3 I_3 + \delta_4 I_4 + e_i$

where y_i is the achievement of individual i; I_j equals one if the individual gives response j and is equal to zero otherwise, and the parameters δ_j represent the average achievement level of individuals who give response j (so it could be written as $\bar{y}_j$); and e_i is a residual component: $E[e_i] = 0$. A number of normalisations of equation (1) are possible – one is to make the first category parameter the intercept and to estimate $(\delta_2 - \delta_1)$, $(\delta_3 - \delta_1)$ and $(\delta_4 - \delta_1)$ as the increments to achievement with each higher level response. Another is to estimate parameters $(\delta_1 - \delta_0)$, $(\delta_2 - \delta_0)$, $(\delta_3 - \delta_0)$ and $(\delta_4 - \delta_0)$ such that they sum to zero (Gardeazabal & Ugidos 2004).

If we adopt the first approach, then we can write the equation for jurisdiction 1 (denoted by the superscript) as

(2) $y_i^1 = \alpha^1 + \beta_2^1 I_2^1 + \beta_3^1 I_3^1 + \beta_4^1 I_4^1 + e_i$

Where $\beta_2^1 = (\delta_2^1 - \delta_1^1), \beta_3^1 = (\delta_3^1 - \delta_1^1), \beta_4^1 = (\delta_4^1 - \delta_1^1)$. Then, a Blinder-Oaxaca decomposition of the gap in average achievement between jurisdictions 2 and 1, denoted by $\bar{y}^2 - \bar{y}^1$, is given by

(3) $\bar{y}^2 - \bar{y}^1 = (\hat{\alpha}^2 - \hat{\alpha}^1) + \hat{\boldsymbol{\beta}}^{\mathbf{1}\prime}(\bar{\boldsymbol{p}}^{\mathbf{2}} - \bar{\boldsymbol{p}}^{\mathbf{1}}) + (\hat{\boldsymbol{\beta}}^{\mathbf{2}} - \hat{\boldsymbol{\beta}}^{\mathbf{1}})'\bar{\boldsymbol{p}}^{\mathbf{2}}$

where $\bar{\boldsymbol{p}}^{\mathbf{2}}$ and $\bar{\boldsymbol{p}}^{\mathbf{1}}$ are matrices of the proportions who respond to the attitudinal question in the two jurisdictions in categories 2, 3 and 4 and $\hat{\alpha}^j$ and $\hat{\boldsymbol{\beta}}^{\boldsymbol{j}}$ are the parameter estimates (matrices of parameters in the case of $\hat{\boldsymbol{\beta}}^{\boldsymbol{j}}$) from the two countries. The three terms allow the total gap in average achievement to be decomposed into the contribution from the constants in the two jurisdictions (measured from the least favourable attitude response); from differences in the distributions across other categories of answers given to the attitudinal question in the two jurisdictions; and from differences in the way those answers are associated with increased achievement in the two jurisdictions.

In economics, the second term in equation 3, $\hat{\boldsymbol{\beta}}^{\mathbf{1}\prime}(\bar{\boldsymbol{p}}^{\mathbf{2}} - \bar{\boldsymbol{p}}^{\mathbf{1}})$, is called the "explained" difference, since it reflects the contribution of observed characteristics, or in this case reported attitudes. The sum of the first and third terms is often called the "unexplained" component, since it reflects differences in parameters between the countries.

Since it is arbitrary which category is used to estimate the "constant" between the two countries in equation (2), it is somewhat arbitrary exactly how much of the unexplained difference in achievement is allocated between the first and the third elements of the decomposition. To assess how much this might matter, we also make use of a statistic $k = (\hat{\beta}_4^2 - \hat{\beta}_4^1)/(\bar{y}^2 - \bar{y}^1)$ which summarises how different the relationship between achievement and improved attitudes are across the two countries, normalised by the gap in average achievement. Values that depart substantially from zero point to major differences between countries in the way achievement and attitudes are related in the countries and that the use of the lowest category attitudinal response may not be a good way of capturing the typical difference in achievement across the countries. If $k \cong 0$, it does not matter much which normalisation is used to measure the constant since the gap in achievement between the countries is similar in the top and bottom attitude categories.

If responses to the question in the two jurisdictions and the way answers are associated with achievement are the same, then the whole of the gap will be explained by differences in the constant. This implies that the relevant attitudinal question does not explain differences in average achievement between the two jurisdictions. This does not mean the attitude does not help explain achievement within the jurisdiction, only that operates in the same way in the two jurisdictions. There is also an alternative decomposition to that given by equation (3), and sometimes the results from the two decompositions provided are averaged. The alternative decomposition is given by

$$(4)\ \bar{y}^2 - \bar{y}^1 = (\hat{\alpha}^2 - \hat{\alpha}^1) + \hat{\boldsymbol{\beta}}^{2\,\prime}(\bar{\boldsymbol{p}}^2 - \bar{\boldsymbol{p}}^1) + (\hat{\boldsymbol{\beta}}^2 - \hat{\boldsymbol{\beta}}^1)'\bar{\boldsymbol{p}}^1$$

Fixed effect regression

The fixed effect regression involves the regression analysis of achievement in mathematics and science for individual *i* against their reported attitudes towards that subject, denoted by *s* below.

$$(5)\ y_{is} = \delta_1 I_{1s} + \delta_2 I_{2s} + \delta_3 I_{3s} + \delta_4 I_{4s} + \mu_i + e_{is}$$

or with just two subjects, *m* and *s*

$$(6)\ y_{im} - y_{is} = \delta_1(I_{1m} - I_{1s}) + \delta_2(I_{2m} - I_{2s}) + \delta_3(I_{3m} - I_{3s}) + \delta_4(I_{4m} - I_{4s}) + e_{im} - e_{is}$$

Factors that affect mathematics and science similarly, such as unobserved ability and motivation levels, family background and general attitudes towards education, examinee effort levels in exams, psychological factors that induce individuals to give the same answers about similarly worded questions and so on are removed by the fixed effects estimation, so only the relationship between differences in achievement and differences in attitudes is measured. Only individuals whose reported attitudes to the two subjects are in different response categories contribute to the estimated parameters of equation (4). This approach is quite different from the decomposition in that the gap in achievement between mathematics and science is assumed to be driven by differences in attitudes and a constant, not by differences in the parameters on the attitudes.

Differencing out cultural factors: Mathematics versus science

The Blinder-Oaxaca decomposition equivalent to equation (6) of the difference in average mathematics and science achievement within Victoria would be given by

$$(7) \qquad (\bar{y}_m^1 - \bar{y}_s^1) = [\,\hat{\alpha}_m^1 - \hat{\alpha}_s^1] + \; +\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{1}\prime}(\bar{\boldsymbol{p}}_{\boldsymbol{m}}^{\boldsymbol{1}} - \bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{1}}) + \left(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{1}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{s}}^{\boldsymbol{1}}\,\right)'\bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{1}}$$

Note that this equation removes any language or cultural effects from the way questions are answered in Victoria by focussing on how the differences in answers might be associated with differences in average achievement. By extension, differences in the gaps in average achievement between a high performing country and Victoria would be given by

$$(8) \qquad (\bar{y}_m^2 - \bar{y}_s^2) - (\bar{y}_m^1 - \bar{y}_s^1) = [(\hat{\alpha}_m^2 - \hat{\alpha}_m^1) - (\hat{\alpha}_m^2 - \hat{\alpha}_s^1)] +$$
$$[\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{2}}{}'(\bar{\boldsymbol{p}}_{\boldsymbol{m}}^{\boldsymbol{2}} - \bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{2}}) - \hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{1}}{}'(\bar{\boldsymbol{p}}_{\boldsymbol{m}}^{\boldsymbol{1}} - \bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{1}})] + [\,(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{2}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{s}}^{\boldsymbol{2}})'\bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{2}} - (\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}^{\boldsymbol{1}} - \hat{\boldsymbol{\beta}}_{\boldsymbol{s}}^{\boldsymbol{1}})'\bar{\boldsymbol{p}}_{\boldsymbol{s}}^{\boldsymbol{1}}]$$

where the subscripts 'm' and 's' represent mathematics and science respectively and the high performing country is superscripted by '2' and Victoria by '1'. Then, the first term is the contribution of differences in the constants on mathematics and science across the two countries, the second term is the contribution from differences in the distributions of mathematics and science responses within the high performing country versus differences in those distributions within Victoria. The third term is the contribution of differential returns to mathematics and science in the two countries. This approach asks whether different response patterns to the prompts and categories within a country help explain differential achievement patterns between mathematics and science within the country compared to another. Any language or cultural differences in interpreting the prompts and responses are removed by making part of the comparison between mathematics and science within the country.