

Mahuteau, Stéphane; Mavromaras, Kostas G.

Working Paper

An Analysis of the Impact of Socioeconomic Disadvantage and School Quality on the Probability of School Dropout

IZA Discussion Papers, No. 7566

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Mahuteau, Stéphane; Mavromaras, Kostas G. (2013) : An Analysis of the Impact of Socioeconomic Disadvantage and School Quality on the Probability of School Dropout, IZA Discussion Papers, No. 7566, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/89877>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 7566

**An Analysis of the Impact of Socioeconomic
Disadvantage and School Quality on the
Probability of School Dropout**

Stéphane Mahuteau
Kostas Mavromaras

August 2013

An Analysis of the Impact of Socioeconomic Disadvantage and School Quality on the Probability of School Dropout

Stéphane Mahuteau

*NILS, Flinders University
and IZA*

Kostas Mavromaras

*NILS, Flinders University
and IZA*

Discussion Paper No. 7566
August 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

An Analysis of the Impact of Socioeconomic Disadvantage and School Quality on the Probability of School Dropout

PISA scores are an internationally established indicator of student and school performance. This paper builds on the evidence that better PISA scores are known to be associated with better later life outcomes. It uses the Australian PISA micro-level data in combination with its longitudinal continuation in the LSAY data, to measure the degree to which individual PISA scores are associated with individual early school dropouts. It distinguishes between student and school factors and estimates a model of the propensity to drop out from school between ages 15 and 18. The paper finds that PISA scores are a good predictor of early dropout, and that individual and social disadvantage plays a crucial role in this relationship both directly and indirectly.

JEL Classification: I24, I21

Keywords: PISA, socioeconomic disadvantage, school dropout, student outcomes, multilevel modelling

Corresponding author:

Stephane Mahuteau
National Institute of Labour Studies
Flinders University
Adelaide, South Australia 5001
Australia
E-mail: stephane.mahuteau@flinders.edu.au

1. Introduction

A wide range of empirical studies shows that people who begin their lives with lower school performance, are more likely to end up with worse longer term economic and general life outcomes. Although research has been suggesting for a long time and consistently that earlier intervention is more cost effective than later intervention, it is still the case that intensely negative schooling outcomes present themselves during adolescence with severe negative consequences for those young people concerned. One such outcome is dropping out from school before the completion of the customary twelve years of schooling without continuing in another form of complementary education pathway. There is strong evidence that dropping out from school before year 12 is strongly associated with poor later labour market outcomes (Rumberger & Lamb 2003, Oreopoulos 2007). There is also evidence that dropping out is associated with several forms of pre-existing or co-existing social and individual disadvantage, with cumulative and lasting negative consequences for future wages, unemployment, job and life satisfaction and much more. It is of policy value, therefore, that the process that leads to school dropout is better understood and that adequate predictors are developed to inform dropout prevention efforts.

This paper sets out to estimate how well individual PISA school scores predict school dropouts. Our definition of dropout is leaving school between 15 and 18 *and* not engaging in any other educational activity at all. PISA scores are collected at the age of 15 in a way that allows considerable comparability between individuals in many different countries, and they are often used as a measure of overall national school performance. Moreover, the methodology behind the computation of standardised test scores like PISA makes them comparable over time, across PISA waves. The paper is motivated by the fact that PISA scores contain probably the most extensive standardised information on individual performance at school that will be collected in the lifetime of an adolescent person. In order to utilise this information, the paper builds on the unique additional information that the Longitudinal Survey of Australian Youth (LSAY) offers by following up the Australian PISA 15-year olds for several years afterwards. The paper estimates the probability that the PISA 15 year olds have dropped out of school by their 18th birthday. The paper then investigates the degree to which the recorded individual PISA scores, may be used as a predictor of the probability of dropout. This exercise could inform the question of whether standardized PISA test scores have the potential to provide an early warning signal of a student's propensity to drop out of school early. The paper identifies several crucial factors that may affect both standardized scores and early dropout, including socioeconomic status, indigenous status, remoteness and school quality. There are several important conceptual issues in using PISA scores as a predictor of future dropout. While the PISA scores may influence the probability of a dropout (in that students with a good performance are less likely to be discouraged and leave school early), it may also be that a high probability of dropout may influence school achievement reflected in the PISA scores (in that students who plan to leave school early are less likely to try hard for a better PISA score). This type of relationships would describe a common case of two-way causality. It could also be that both the PISA scores and the probability to dropout may covary, not because they influence one another, but because both may be determined by other factors, such as ability or other socioeconomic barriers.

This paper utilises the information contained in the PISA data and combines it with the longer term information in the LSAY. It also utilises a multilevel regression model which allows the explicit estimation of the different components of the PISA scores which may influence the

dropout probability, highlighting in particular the difference between student and school, as well as between observed and unobserved characteristics. The paper focusses on several specific hypotheses of policy importance, where the combination of data and estimation method offers novel findings. For example, the paper tests the underlying hypothesis that bad PISA results may be able to give a reliable warning sign that students concerned are at risk of dropping out of school before completing year 12. The paper also tests the hypothesis that the estimated school ‘quality’ (or value-added) plays a part in determining the schooling outcomes of students in terms of participation and completion. An important aspect of this paper is the investigation of whether all factors examined have an effect on the probability to dropout over and above that on the test score. Being able to distinguish between the direct and indirect effects of socio economic factors on school dropout is highly relevant to defining appropriate policy responses by suggesting targets for early intervention and identifying types of disadvantage that may require a longer term policy approach.

The main innovation of this paper is that it uses a unique combination of data sets (the PISA set and its Australian longitudinal continuation, the LSAY) to investigate the degree to which the PISA scores can be an accurate predictor of later performance, measured by the rate of school dropouts. It also decomposes the impact of socioeconomic disadvantages into a direct effect on the propensity of dropping out and an indirect effect through the standardised test scores. Three aspects of the paper’s innovation are worth individual attention. First, it uses of multilevel econometric modelling, which allows us to distinguish between individual and school-specific factors. Second, it uses multivariate analysis, which enables us to investigate the role of different types of disadvantage in the relationship between PISA scores and school dropout. Finally, it uses a unique combination of Australian data sets, which allows us to derive a precise measure of school outcomes three years after the PISA scores were collected. This last innovation may be Australia-specific, but has wider implications as it allows us to understand the broader capabilities of the PISA score as a predictor of future performance, which may also apply to other countries where an additional data set like the LSAY may not be available. The remainder of the paper is structured as follows. Section 2 presents the two data sets we use. Section 3 develops the multivariate regression and multilevel methodology of the estimations and presents the final econometric model. Section 4 presents the results and discusses their implications. Section 5 concludes.

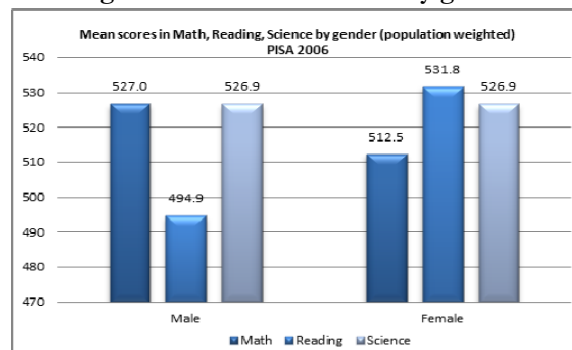
2. The Data

The paper begins with matching the data from the Australian PISA 2006 standardized test scores with the 2006 Longitudinal Survey of Australian Youth (LSAY). The first wave of the 2006 LSAY consists of the students in the PISA 2006. The matched data preserves all the information on students and schools available in the 2006 PISA while adding extra LSAY information about the school attainment of students up to year 2009, the year of the last LSAY interview that is currently available for the 2006 cohort. By 2009 many of the 2006 students had completed high school and had gone on to further education or to the labour force. The PISA/LSAY combined data contains rich socio-demographic and economic information. It also contains rich information through a number of measures on socio economic disadvantage, both at the student and the school levels. The data allows us to link socio economic disadvantage, scores and school estimated ‘quality’ with the schooling attainments of the students. More specifically the paper calculates the probability that a student may drop out of school before completing year 12 or equivalent. Dropout is defined as the state where the student has left school and has not engaged in any further education activity. Those students who left school early (in the Australian context

that is before the completion of Year 12) and engaged in Vocational Education and Training (VET), are not counted as dropouts.¹ The data set consists of 14,170 students attending 356 schools initially surveyed in the 2006 PISA data. The LSAY data collection surveyed these students yearly until 2009, until they reached their Year 12. Whether they remain at school until graduation, drop out, or leave school to undertake vocational education before completing year 12, these students were sent the LSAY questionnaire. There is some attrition from one LSAY wave to the next, mostly after the first year where 33% of the students did not return their questionnaire. The attrition rate for the remaining two waves of the survey is much lower. Altogether, by 2009, we are left with 51% of the original PISA sample, which makes a total of 7,299 students. The methodology of the estimations takes this attrition into account so as to maintain the representativeness of the results presented in the paper.

In the remainder of this section, we present a number of relevant descriptive statistics of the PISA and LSAY samples upon which we base our study. The following graph shows the mean PISA scores of students in Reading, Maths and Science by gender (population weighted). The scores are significantly above the OECD average of respectively 492 in Reading, 498 in Maths and 500 in Science. A notable exception is the reading scores of Australian males, which are not significantly higher than the OECD average. We observe that males have higher scores than females in Maths but a much lower score in Reading. The scores in Science are comparable.

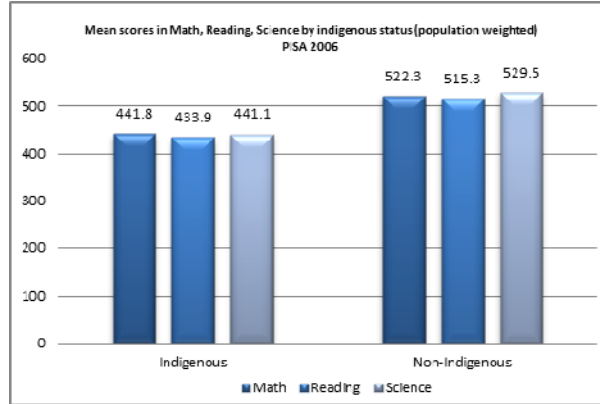
Figure 1: Mean PISA scores by gender



The following graph illustrates the large, statistically significant, differences that exist between indigenous and non-indigenous students with respect to PISA scores. The scores of indigenous students are almost 20% lower, on average, than for non-indigenous students.

¹ Those who completed a VET course at Certificate I or II levels remain accounted as dropouts since this type of training does not provide individuals with substantial extra qualifications compared to the level they would normally have after dropping out.

Figure 2: Mean PISA scores by indigenous status



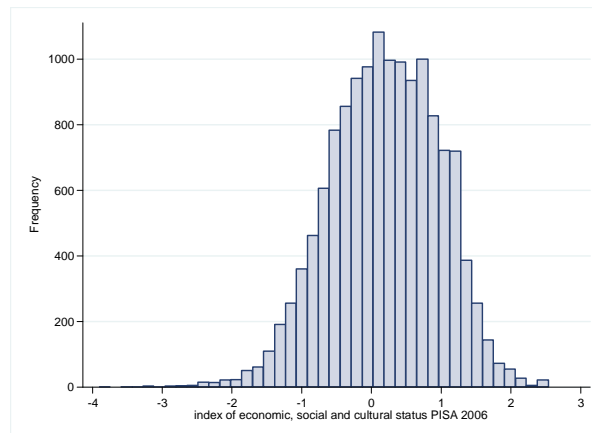
The disadvantage of indigenous students extend to the proportion of school dropout which is at least twice as high as non indigenous students as illustrated by the following table.

Table 1: School dropout rates by indigenous status

Dropout before completing year 12 (pop weighted)		
Indigenous status	dropout (all initial PISA students)	dropout (non attriters)
Non-Indigenous	0.045	0.083
Indigenous	0.070	0.181

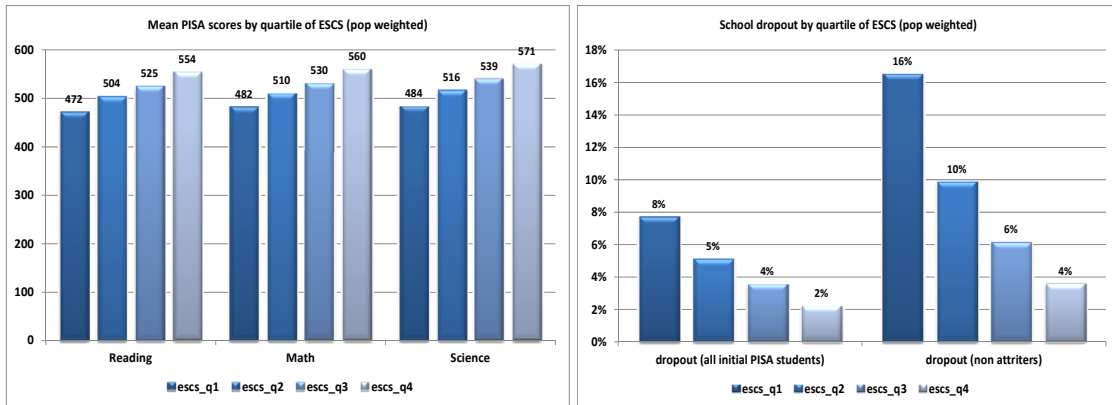
Parental socioeconomic status (denoted ESCS) is captured by the composite index available in the PISA data (see OECD, 2007) which is obtained by principal component analysis and crosses five dimensions of the background of students (captured by parental information on highest occupation, education, family wealth, cultural possessions and home educational resources). The index has been computed to allow over time as well as cross country comparisons and has been standardised so that the mean ESCS of all OECD countries is 0. Australian students surveyed in PISA 2006 have a population weighted mean ESCS of 0.21 which suggests that their mean socioeconomic status is above the OECD average. The ESCS score ranges from -3.9 to 2.5 in our sample and its distribution is presented in the following graph.

Figure 3: Economic, social, and cultural status



The following two graphs show the mean PISA scores (in Reading, Maths and Science) and early school dropout proportions by the socioeconomic status quartile (ESCS) of the student. Looking at reading scores, the highest ESCS quartile students achieve a score that is 17% higher than the lowest quartile students. The gap in terms of proportion of early dropout between top and bottom quartiles is even more obvious.

Figure 4: PISA scores and School dropout by ESCS quartile



3. Estimation Methodology

The PISA data used for estimation comes in a hierarchical structure, comprising students at the first level who are nested in schools at the second level. Having reliable and complete information for both levels enables us to distinguish between (i) the variation between all student scores within each school (which we call *within school variation*) and (ii) the variation between all average scores of students in different schools (which we call *between school variation*). This offers more opportunity to control for the degree to which unobserved heterogeneity (diversity) arising from school overall differences, may influence individual student scores.

Students belonging to the same schools are expected to be more alike in their socio economic characteristics than students from different schools, including with respect to unobservable characteristics. Similarly, students in the same school are expected to share the same teachers for some subjects, and benefit from the same facilities and environment. Because of these similarities, we would expect the score of students belonging to the same school to be more alike than those of students in different schools. The implication of these similarities is that treating each student observation as independent from one another (when indeed, some are linked because they are going to the same school) could lead to potentially large biases in the estimation of the determinants of PISA scores. In order to address the issue of such biases, we use hierarchical regression which models explicitly and estimates the multilevel structure of the data by distinguishing between the school effects and the student effects on scores (see Rabe-Hesketh & Skondral, 2008). Adding the school dimension to the investigation gives our estimation results more scope to identify and explain those policy problems and their implications that are more closely related to social disadvantage at the individual student level as well as at the more broadly shared school level.

The paper presents results based on fairly general definitions of social disadvantage and investigates a range of social issues and their impact on student scores. A potential limitation of the LSAY data is the fairly high attrition rate whereby one loses more than 48% of the student sample between the age of 15 in 2006 and 18 in 2009. This high attrition rate requires an appropriate econometric treatment in order to avoid selection bias that may be due to the possibility that the 52% of the students who stayed in the sample may be systematically different from the 48% who did not. Indeed, it is likely that the proportion of students who drop out of school will be higher among those who drop out of the survey than among those who stay in the survey, as the two dropout rates (from the survey and from the school) will be higher among discouraged students. Ignoring this possibility would lead us to underestimate the school dropout propensity (as our sample consists of those – better motivated students – who have remained in the survey) and get biased estimates of the relationship between school dropout and PISA scores/ school ‘quality’.

The paper provides two sets of estimations of school dropout, which are directly testing whether the non-respondents to the survey can be considered as randomly drawn from the student population, or whether they are a self-selected group. The first set of estimations assumes that there is attrition selection bias, and corrects it using a first stage selection equation (with attrition being the dependent dichotomous variable), followed by a second stage school dropout estimation of the conditional probabilities to drop out of school for those who remained in the survey till the age of 18. The second stage uses the LSAY information from 2006 to 2009. We use a selection model adapted to the dichotomous nature of the variable of interest in the second stage equation (Van de Ven & Van Praag, 1981), with ‘early school dropout’, taking the value of 1 if the student drops out of school before 2009 and 0 otherwise. We estimate the conditional probability to drop out of school using full information maximum likelihood methods and correcting for the correlation between the unobserved component of the selection (first-stage) equation and the unobserved component of the school dropout (second-stage) equation. The statistical significance of the estimated correlation coefficient between the two unobserved components provides us with an estimate of whether selection into attrition is based on unobservables or not. We treat the question of whether there is *systematic* survey attrition or not as a largely empirical matter. Where the two-stage estimation informs us that attrition based on unobservable characteristics is statistically significant, we continue with the use of this method to avoid biased estimates. Where the two-stage estimation indicates that selection is only based on observables, we explicitly control for these variables in a simpler estimation framework. In those cases further estimation does not utilise the two-stage estimation method (as we have evidence that it will not yield an improvement) and we revert to using a single-stage estimation of school dropout, as we know this to be statistically the most efficient estimation method.

After we have decided on the appropriate estimation method regarding attrition bias, we employ these two estimation methodologies (that is the one-stage and the two-stage, depending on the absence or presence of attrition, respectively) in order to estimate two types of models, the first model with a basic specification and the second model which augments the basic model with measures of student and school quality. The first model of school dropout is described as follows:

$$y_{ij}^* = X_{ijD}\beta_D + Z_{jD}\gamma_D + \varepsilon_{ijD} \text{ if } \Lambda=0 \text{ (no attrition)} \quad (1)$$

$$\text{With } \begin{cases} y_{ijD} = 1 & \text{if } y_{ijD}^* > 0 \\ y_{ijD} = 0 & \text{otherwise} \end{cases}$$

Where attrition and thus selection on both observables and unobservables is tested for and explicitly modelled as required using two alternative specifications of the basic model (two stages and bivariate probit estimation). The specification, explicitly modeling attrition in the LSAY data, can be written as:

$$\begin{aligned} y_{ij}^* &= X_{ijD} \beta_D + Z_{jD} \gamma_D + \varepsilon_{ijD} \quad \text{if } \Lambda=0 \\ A_i^* &= W_i \beta_A + \varepsilon_{iA} \end{aligned} \tag{2}$$

$$\text{With } \begin{cases} A_i = 1 & \text{if } A_i^* > 0 \\ A_i = 0 & \text{otherwise} \end{cases} \quad \text{and assuming } (\varepsilon_{ijD}, \varepsilon_{iA}) \sim BVN(0, 0, 1, 1, \rho)$$

Vector W of individual characteristics includes variables that belong to X and/or Z , but also includes an exclusion restriction in order to allow for identification. X represents a set of individual characteristics of the student (including background and environmental) deemed to be related to the probability to drop out of school. Z is a vector of school characteristics. y_{ij}^* represents the ‘latent’ individual propensity to drop out of school (unobserved), for which the observable counterpart is whether or not the student drops out of school y_{ij} . Likewise, the attrition probability depends on a set of individual characteristics (which also include some school characteristics) and an unobservable component ε_{iA} , which is assumed to have a bivariate normal joint distribution with the unobservables of the probability of school dropout ε_{ijD} , with a correlation coefficient ρ estimated through the model.

The basic model contains a broad set of covariates which we expect to be relevant in the estimation of school dropout propensity. These are variables that have been shown in the literature to play a role in the determination of dropout probabilities. We begin with variables that control for student motivation and psychological and attitudinal factors (Dewey, Husted & Kenny, 2000). Students who fail to perceive or underestimate the longer-term rewards associated with post school education (as opposed to the immediate smaller rewards from entering the labour force sooner), will be less motivated to stay on and will be more likely to opt out of school before completing Year 12 (see Curtis & McMillan, 2008; Eckstein & Wolpin, 1999; Haveman et al., 2005). We add variables associated with student disadvantage as they have already been identified to influence student scores and are highlighted in the literature as responsible for student early dropout (Goldstein, 2002; Rivkin et al., 2005; Miller & Voon 2011). We also add information available in the PISA 2006 data on parental occupation type. Unfortunately, the 2006 PISA data does not record parental labour force status, so we use the more limited information on whether parents are blue collar workers or not which is readily available. We also include the ESCS index as a broader measure of neighbouring disadvantage, alongside with information on indigenous status and migrant status of the parents and the students, and measures of the relative remoteness of their place of residence. Finally we control for a number of school characteristics and for State of residence. The model is estimated using both specifications (with and without

correction for attrition bias) and the relevant estimation results, here with correction for attrition, are kept for comparison with the extended model containing measures of student and school quality.

The second model, contains the same set of covariates in the right hand side as in the basic model, but crucially adds to it the individual PISA score and an estimate of school quality obtained through a multilevel estimation of the standardized test scores of students at age 15. These two additions are crucial. The possibility is that the PISA score may be endogenous in the dropout process. This could be because lower PISA scores may be the cause of dropouts as students may become discouraged, and because an anticipated future dropout may be the cause of low PISA scores, as those students who plan to drop out are less motivated to make the effort for higher scores. Moreover, the literature identifies past achievement of students as a determinant of their later outcomes (Dewey et al., 2000; Lamb et al., 2001). To ameliorate potential bias we estimate the predictors of the PISA scores using multilevel regression, which distinguishes between first level (the students) and the second level (the school) variation. We use the results of this multilevel regression to derive the predicted values of the PISA scores (an estimate of the student's past achievements) and the estimated random intercept (an estimate of the school quality). The inclusion of these two estimates in the original model allows us to test the impact of student quality and school quality on dropout rates, after we have controlled for all those variables that were included in the original model². It also allows us to compute the total effect of student and school disadvantage on the probability to drop out as the sum between an indirect effect acting through the student's past achievement (PISA score) and a further direct effect on the probability extending beyond the actual PISA scores.

The second (augmented) model is given as:

$$y_{ij}^* = X_{ij}\beta_D + Z_j\gamma_D + \lambda P_{ij} + \varepsilon_{ijD} \quad (3)$$

$$\text{With } \begin{cases} y_{ij} = 1 & \text{if } y_{ij}^* > 0 \\ y_{ij} = 0 & \text{otherwise} \end{cases}$$

Where P_{ij} represents student i 's PISA score in school j .

We assume that P_{ij} is endogenous in (3) and thus we explicitly express the PISA scores of students as depending on a vector of students, family background, schools and environmental factors where we take into account the hierarchical nature of the data through the definition of variance components as follows:

$$P_{ij} = X_{ijs}\beta_S + Z_{js}\gamma_S + u_{js} + \varepsilon_{ijs} \quad (4)$$

where u_{js} is a random intercept associated to the school j student i attends. We use maximum likelihood methods (Rabe-Hesketh & Skondral, 2008) to compute the school specific component u_{js} which gives an estimate of the systematic effect of belonging to school j on scores, over and

² We used an alternative specification based on the simultaneous estimation of both equations (PISA scores and Probability of dropout). We obtained similar results compared to the method described above. As a result we opted for presenting this latter, simpler, method. The results of the simultaneous equations specification are available upon request.

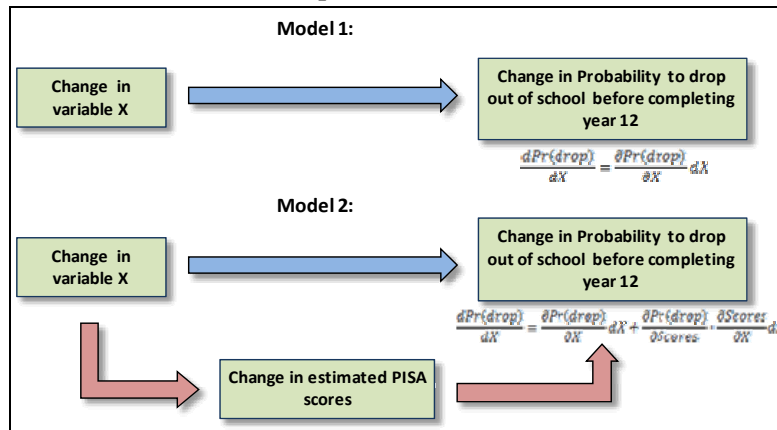
above the effect of school resources. This estimated random intercept is commonly used as an indicator of school quality in the literature which looks at students' outcomes based on an educational production function.

We estimated both models with and without accounting for the presence of attrition in the LSAY data. We found that the correction for selection bias arising from attrition was necessary for the basic model without PISA scores but not for the second model including estimated scores and school 'quality' as covariate. Hence the results are displayed accordingly.

Our analysis was extended to account for the fact that the PISA score Program tests students on three different subjects, namely literacy (reading), maths and science, and provides three different scores. To account for the presence of three different sources of information, we estimated three versions of the second model, including respectively the estimated PISA scores on each subject along with the estimated school 'quality' component associated with each subject. We found that the results of all three models were consistent and remained stable to the change of subject, hence we only display the model using the Reading scores. The results of other models are available upon request.

The comparison between the basic and enhanced models also allows us to distinguish between the direct effect of student disadvantage on school dropout probability and the indirect effect of student disadvantage on school dropout probability (mediated through the effect of student disadvantage on the PISA standardised scores). The interpretation of these estimations is crucial for the understanding of the policy relevance and importance of this paper. Put simply, the enhanced model allows us to establish whether the association between disadvantage and higher dropout rates remains after we have controlled for the association between disadvantage and lower PISA scores. If a disadvantage variable is significant in the first set of estimations (and is significant in the estimation on scores) and remains significant in the second set of estimations which contain predicted scores, we can conclude that this variable is not only associated with the probability to drop out directly, but also via its relationship with predicted PISA scores. In practical terms, such a result would suggest that some types of socio economic disadvantage may have a compounding effect on the drop out probabilities, which extends beyond its impact on PISA scores. One implication of this finding would be that early interventions aimed at removing this specific type of disadvantage would only partly alleviate its negative impact on drop out. Alternatively, we may find that the overall effect of some other types of socio economic disadvantage on the probability to drop out does not extend beyond their indirect effect on drop out probability through PISA scores. The policy implication of this finding would indicate that earlier interventions would help reduce the impact of that disadvantage on the probability to drop out. Altogether, the specification adopted in this paper allows us to identify and distinguish between these socio economic disadvantages that have a compounding effect on the probability to dropout from those that act only at an earlier stage through the PISA scores. Figure 5 portrays a graphical representation of the two models and shows how the probabilities to drop out of school change with corresponding changes in the value of variables in the model.

Figure 5: Effect of a change in the value of a variable on the probability to drop out of school, comparison of the two models



4. Estimation Results

The paper finds that a number of factors systematically influence the likelihood that a student will leave school before completing Year 12. These factors have direct effects, and also interact with each other. The main ones that we have identified are the ESCS of the student, the PISA scores of the student, the quality of the school, the size of the town/city, and the state of residence. For all but the last of these, higher values lead to significantly lower probabilities of dropping out of school before Year 12. The estimated marginal effect for the ESCS of students in the model with PISA scores is -1.3%, which suggests that a one point difference in ESCS compared to the population mean leads to a lower probability of dropout of 1.3 percentage points, independent of the effect of a change in socioeconomic status on PISA scores. Note that a one point difference compared to the population mean (0.21) would place a student in the top quartile of socioeconomic status. In addition, ESCS impacts PISA scores in a positive way. A similar change of ESCS at the individual level as suggested above leads to an increase of reading scores by 7.1 marks for females students and a further 12.6 marks for males. In turn, a higher PISA score has a further effect on the probability of dropout in the expected direction with a marginal effect estimated at - 0.0008. Hence, the individual socioeconomic status of the students produces both a direct and indirect effect on the probability of dropout.

Our results in Table A2 corroborate previous findings in the literature which suggest that the socioeconomic status of the student body in a given school is more important in determining standardised test scores than individual ESCS (see Miller & Voon, 2011 for an illustration on Australia), even though the latter also has a significant positive impact. Indeed, the results of the estimation on PISA scores highlight a coefficient of 56 for the mean school ESCS. Hence, shifting from a school whose student enrolment corresponds to the population mean ESCS to a school in the top ESCS quartile would raise individual student PISA scores by 56 points, and, by extension, would decrease the probability of early dropout.

The following table gives an extract of the estimation results on both PISA scores and dropout probabilities for student and school level ESCS along with the estimated impact of PISA scores onto the dropout probabilities. We only display the results of the preferred specification, including PISA scores.

Table 2: Dropout probability and PISA scores estimation: ESCS

VARIABLES	PISA Reading Score (Coef.)	Dropout (Marginal Effect)
Student ESCS (deviation from pop mean)	7.107*** (1.281)	-0.0130** (0.00639)
Male interacted with School ESCS	12.55*** (3.876)	--
Average School ESCS	55.99*** (17.51)	--
Fixed part estimated PISA score (reading)	--	-0.00079*** (0.000117)

Note: Standard errors in brackets. Detailed results in Table A1 and A2.

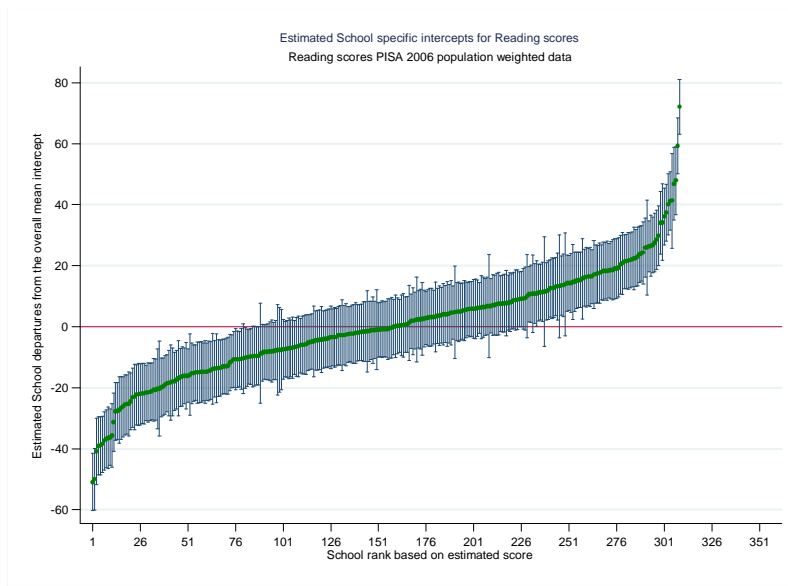
While the marginal effect associated with PISA scores in the dropout equation seems rather small, it is nonetheless significant, and must be seen in the context that low PISA achievers are also likely to cumulate other types of disadvantage that would further affect their probability of dropout. To illustrate this, we computed the expected probability of dropout for four categories of students based on their PISA score achievement. These four groups correspond to as many quartiles of the estimated PISA scores. In order to compute these probabilities, we assume that students also take on the mean characteristics of their respective quartiles for all other variables of the model. Table 3 illustrates the gap between low achievers and high achievers with respect to dropout probabilities.

Table 3: Dropout probability by PISA score quartile

Quartile of PISA scores	Estimated probability of dropout
Quartile 1	11.3%
Quartile 2	6.7%
Quartile 3	4.4%
Quartile 4	2.9%

The multilevel estimation of PISA scores enables us to compute estimates of school ‘quality’ represented by the estimated random intercept. These estimates sum up, for each school, the extent to which student scores systematically differ from those of identical counterparts attending a different school. In other words, for given individual characteristics and school resources, some schools over- and under-perform compared to the overall average school, and the extent of their over- and under-performance is estimated by their random intercept. Figure 6 reports the estimated random intercepts obtained for each school along with their corresponding 95% confidence intervals. Any interval (horizontal line) which does not include zero in it suggests that the corresponding school significantly over-performs (on the right, roughly the 70 highest ranking schools), or under-performs (on the left, roughly the 80 lowest ranking schools), compared to the overall average school.

Figure 6: Distribution of the estimated school quality



Many factors are shown to be of influence. Low ESCS students in low quality schools in smaller towns who achieve lower PISA scores are much more likely to drop out than is the average student. These conclusions are not surprising. Our analysis validates them and provides a clear quantification of the strength of the different sources of increased probability to drop out and how they interact with each other. Students from low ESCS backgrounds are more vulnerable to the adverse effects of smaller towns, and lower quality schools, than are students from more advantaged backgrounds: they have at least double the probability of dropping out.

In order to provide an intuitively appealing and useable quantification of the combined effect of school quality and socioeconomic status on student dropout probabilities, we conducted a counterfactual analysis, whereby we estimated the PISA scores and subsequent probability of dropout for students according to three values of socioeconomic status (corresponding to the three thresholds of ESCS quartiles) while allocating the students to three types of school quality (under-performing, over-performing, and neither under- nor over-performing). The values retained for the school quality correspond to the three thresholds used to define the school quality quartiles. Throughout, the value of all other individual and school characteristics are assumed to correspond to the sample mean. Table 4 contains the results of the counterfactual analysis, providing a range of estimated dropout predictions from 9.58% for the lowest ESCS student in the lowest school quality (School 1), to 4.89% for the highest ESCS student in the highest school quality.

Table 4 shows that the risk of dropout would be reduced by a third if students at the lowest ESCS threshold were to attend a high quality school (highest threshold of school quality) instead of an under-performing school (School 1). Compared to attending an average school (with a zero random intercept), these students would shift from an estimated probability of 8.07% to 6.86%, that is a 15% lower probability of dropout. This counterfactual analysis illustrates the large disparity in terms of dropout probabilities between different ESCS students, whereby the highest ESCS student group would experience a comparable probability of dropout if they attended an under-performing school (school 1) to the probability that would be experienced by the lowest

ESCS students attending an over-performing school (school 3). It is worth noting that in the calculations of Table 4 we have assumed that all students have the same sample mean characteristics, which is not likely to be the case. We would expect that students and schools found in the lowest quartiles will also possess more of the characteristics that are associated with higher dropout probabilities. This means that the estimated dropout gap between ESCS groups could be higher in reality than what Table 4 suggests, as these other characteristics will be unevenly distributed between ESCS and school quality quartiles.

Table 4: Probability of dropout by ESCS and school quality

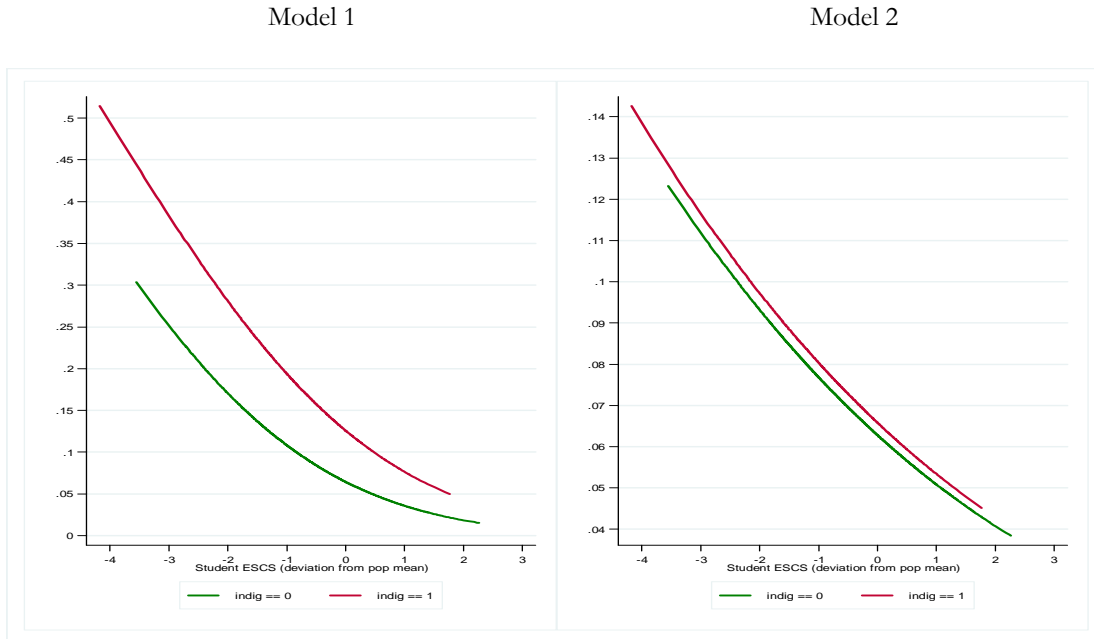
Quartile threshold of ESCS	Quartile threshold of School estimated quality			
	School 1	School 2	School 3	Mean school quality
1	9.58%	8.24%	6.86%	8.07%
2	8.23%	7.03%	5.82%	6.89%
3	7.01%	5.96%	4.89%	5.83%

The results of the model show that some types of disadvantage are associated with both a direct and indirect effect (through PISA scores) on the dropout probability, while others are associated with only an indirect effect. The latter cases would be of added interest because our results suggest that earlier policy intervention to improve the PISA score by targeting these types of disadvantage would have lasting effects in terms of reducing the probability of dropout.

The distinction we are able to make between indirect and direct effects of individual characteristics on the probability of dropout is also informative in questioning certain important and influential observations that are often made (especially in policy circles and in the broader public debate about schools) based on the evidence of simple descriptive statistics. For instance, if one does not account for the effect of PISA scores on the dropout probabilities (as we do in our basic model) one could simply conclude that indigenous students experience a higher probability of dropout. While this appears to be true, the second specification of the model tells us that most of the effect of indigenous status is indirect, and is mediated through lower PISA scores: the effect of indigenous status is no longer significant, although it is positive. This result suggests that *most of the difference in terms of dropout probabilities between indigenous and non-indigenous students comes from the effect of the original differences they have in terms of scores*. As we have seen above, indigenous students have significantly lower scores than non-indigenous students. Since indigenous students obtained on average an estimated 38 points lower score in reading in 2006 than non-indigenous students, the consequences in terms of probability to drop out, if we evaluate it around the mean PISA score is 11.32% for an indigenous student and 7.36% for a non-indigenous student. As our second model specification shows, if we sample repeatedly many pairs students having the same estimated PISA score in reading, one of them being an indigenous student and one not, the probability to drop out of school would not be significantly different between the indigenous and non-indigenous students. The results from model 2 suggest that there is no further disadvantage from being an indigenous student on the probability to drop out, after we have adequately controlled for the negative gap in PISA scores. Figure 7 shows that when we use model 1 (which does not control for PISA score differences) there is a considerable gap between indigenous and non-indigenous students by ESCS. However, once we control for differences in PISA scores in model 2, the differences in probabilities all but disappear. Given the

enormous education and general disadvantage suffered by indigenous Australians, the policy implication of this finding is as clear as it is powerful.

Figure 7: Dropout probability by indigenous status



Similarly, we can focus on migrant status and observe significant differences between native and non-native students. Estimation results show that students speaking a different language than English at home are about 5.24% less likely to drop out of school than those who do not with a 95% confidence interval ranging from 3.8% and 6.6% (the estimate is 6.5 % in the basic model which does not control for PISA score differences). In addition, students whose mother was not born in Australia are a further 1.87% less likely to drop out of school before completing Year 12. An additional important finding from the analysis in this paper is that indigenous students are no more likely to drop out of school before completing Year 12 than are other students of the same socio-economic status. Also interesting is that the children of migrants, especially the non-English speaking ones, are more likely to complete school. A higher ratio of students to teachers increases the likelihood of dropping out, but there is no evidence that private schools perform differently from public ones (once all other factors have been controlled for).

5. Conclusion

This paper estimated school dropout propensity for students between the ages of 15 and 18 focussing on the role of various forms of social and economic disadvantage in the dropout process. The economic motivation of the paper has been that school dropout is a well-recognised predictor of considerably worse short and long-term labour market outcomes. The problem with school dropout events is that they happen at the very end of school life, so they obviously concern those cases where earlier preventive policies were less successful. From the policy maker's angle, they are cases of failure to prevent an event that we know will go on to generate considerable future disadvantage, hence the desire to find ways to reduce the extent of their incidence. To do this we need early, clearly recognisable and statistically reliable warnings, about

the characteristics of the students and the schools where dropout events are more likely to occur. The paper argued that using the PISA data on individual students and their schools (in the form of the PISA standardised individual scores and the estimated school quality indicator, respectively) can serve as an earlier warning indicator for the likelihood of dropout. The paper's main objective was to establish if the evidence supports this argument. The paper's policy motivation is that if we find that the PISA score is a good predictor of dropout, it could be used broadly for policy purposes using the PISA data collection. The paper argued that the PISA data collection has the advantage of containing regular, uniform, and consistent information on students and schools, and that this advantage, combined with the LSAY longitudinal information, results in a powerful policy tool in the design of educational policies.

The paper estimated several models of school dropout probability, with and without the standardised PISA score in the list of explanatory variables, to establish if the PISA score contains such useful information. In order to achieve this, the paper had to first handle several methodological problems. The first methodological hurdle was the problem of possibly non-random attrition in the LSAY data. The paper found attrition to not be a problem in the school dropout estimations that utilised the PISA scores. The second methodological hurdle was the potential endogeneity of the PISA score in the school dropout equation. The paper found that treating endogeneity is necessary. The third methodological advance was the recognition that, for the results to be useful for policy design, we need to distinguish between the variation that is due to difference between students and the variation that is due to differences between schools. To identify and estimate the different student and school information in the data the paper used multilevel regression which estimated student quality (based on the standardised PISA scores) and school quality (based on school data surrounding the PISA scores). The resulting estimated measures were incorporated in the form of predictions in the estimation of school dropout rates. Several measures of economic, demographic, and social disadvantage were also included in the analysis in order to measure their direct and indirect relationship with school dropout propensity.

The main result of the paper was that PISA scores can predict dropout propensities in a statistically significant manner. They do so more for the bottom achievers, a finding which amplifies the policy message and value of this paper. We found that the distinction between a *direct* association between of disadvantage and dropout propensity and an *indirect* association (mediated through the association of the disadvantage and the PISA score) can be traced empirically, and is statistically significant. We found that the measure of socioeconomic status (the ESCS) is associated with dropout rates both directly and indirectly. Several other important factors (such as being male, of indigenous status, maternal occupation and school estimated value added) were found to be associated with dropout rates indirectly through the PISA scores but not directly. Several other factors were found to have both a direct and an indirect association, where the two associations work in an opposite direction, resulting in a lower net value (such as students with English as their second language, who have lower scores, but who also have lower dropout probabilities). Some factors were found to have both a direct and an indirect association, both working in the same direction, resulting in a cumulated association. In some cases the association can be very high. For example the combination of low PISA scores, low socioeconomic status and low school value added, produces a dramatic increase in the probability of school dropout. Several demographic groups were shown to be different from others in informative ways. Gender matters, but only in the indirect relationship mediated through PISA scores. For indigenous students the observed higher dropout rate is entirely due to the indirect

association between lower PISA scores and higher dropout rates, and holds for all socioeconomic levels. Finally, although location seems to matter, there are no discernible patterns we could find.

In conclusion, the paper has shown that the PISA scores can be used as predictors of dropout, and could therefore be useful tools as a warning device for interventions after the age of 15 when the PISA data is collected. The clear association with several identifiable socio-demographics and the capacity of our model to distinguish between direct and indirect associations adds to the value of the PISA scores as a warning measure. Above all, the strong role of student disadvantage in the relationship between PISA scores and resulting dropouts indicates convincingly that the policy message is stronger where it is needed most, namely for those students with the most severe disadvantage. By showing that those with the most intense disadvantage are most likely to take an educational pathway which is known to generate further disadvantage, the results of this paper reveal and quantify an educational pathway which cumulates economic and social disadvantage.

References

- Curtis D., McMillan J., (2008), School Non-completers: Profiles and Initial Destinations, LSAY Research Reports, Australian Council for Educational Research, no.54.
- Dewey J., Husted T., Kenny L.W., (2000), “The ineffectiveness of School inputs: A product of misspecification”, *Economics of Education Review*, Vol 19, pp. 27-45.
- Eckstein Z. ,Wolpin K. I., (1999), “Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities”, *Econometrica*, vol.67, pp.1295–1339.
- Goldstein H., 2002, *Multilevel Statistical Models*, Hodder Arnold.
- Haveman R., Wilson K., Wolfe, B. (2005), “The Role of Expectations in Adolescent Schooling Choices: Do Youths Respond to Economic Incentives?”, *Economic Inquiry*, vol.43, iss.3, pp.467-492.
- Lamb S., McKenzie P., (2001), *Patterns of success and failure in the transition from school to work in Australia*, LSAY Research Reports, Longitudinal surveys of Australian youth research report, no. 18.
- Miller P.W., Voon D., (2011), “Lessons from My School”, *The Australian Economic Review*, vol 44 no 4, pp. 366-86.
- Oreopoulos, P. (2007), “Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling”, *Journal of Public Economics*, vol.91, iss.11-12, pp.2213-2229.
- Rabe-Hesketh S., Skrondal A., 2008, *Multilevel and Longitudinal Modeling using Stata*, Second Edition, Stata Press, College Station, Texas.
- Rivkin S.G., Hanushek E.A., Kain J.F., (2005), “Teachers, Schools, and Academic Achievement”, *Econometrica*, vol.73, no.2, pp.417-458.
- Rumberger R.W., Lamb S.P., (2003), “The early employment and further education experiences of high school dropouts: a comparative study of the United States and Australia”, *Economics of Education Review*, Vol 22, Issue 4, August 2003, Pages 353–366.
- Van de Ven W.P.M.M, Van Praag B.M.S, (1981), The demand for deductibles in private health insurance: A probit model with sample selection, *Journal of Econometrics*, Vol 17, Issue 2, November 1981, Pages 229–252.

Appendices

Table A1: Estimations of the probabilities to drop out of school, models with and without estimated PISA scores (PISA/LSAY)

VARIABLES	Model without PISA scores			Model with PISA scores	
	attrition	Dropout	Marginal effects	Dropout	Marginal effects
Student doesn't think it is important to do well in science Q36_1		0.350*** (0.0566)	0.0306*** (0.00653)	0.300*** (0.0579)	0.0418*** (0.00921)
Student doesn't think it is important to do well in reading Q36_3		0.0611 (0.0807)	0.00535 (0.00709)	0.0608 (0.102)	0.00781 (0.0136)
Student ESCS (deviation from pop mean)	0.207*** (0.0201)	-0.183*** (0.0501)	-0.0160*** (0.00529)	-0.105** (0.0518)	-0.0130** (0.00639)
Fixed part estimated PISA score (reading)				-0.00639*** (0.000945)	-0.00079*** (0.000117)
Estimated residual school quality				-0.00207 (0.00152)	-0.000255 (0.000188)
Male	-0.160*** (0.0273)	0.0958* (0.0516)	0.00839* (0.00482)	-0.0756 (0.0602)	-0.00929 (0.00742)
Indigenous status	-0.242*** (0.0577)	0.215** (0.102)	0.0220* (0.0125)	0.0239 (0.116)	0.00300 (0.0148)
Father is a blue collar	0.0234 (0.0292)	0.105** (0.0515)	0.00936** (0.00475)	0.0953 (0.0584)	0.0120 (0.00746)
Mother is a blue collar	0.0509 (0.0390)	0.143* (0.0759)	0.0137* (0.00817)	0.106 (0.0835)	0.0139 (0.0117)
Student non Australian born	-0.202*** (0.0452)	0.0902 (0.101)	0.00840 (0.0101)	0.0101 (0.115)	0.00126 (0.0144)
Mother non Australian born	0.0501 (0.0321)	-0.167*** (0.0649)	-0.0138*** (0.00522)	-0.160** (0.0746)	-0.0187** (0.00821)
Student speaks language other than English at home		-0.493*** (0.121)	-0.0301*** (0.00568)	-0.653*** (0.136)	-0.0524*** (0.00703)
NSW	-0.137* (0.0779)	-0.0691 (0.0943)	-0.00586 (0.00772)	-0.147 (0.102)	-0.0171 (0.0112)
VIC	-0.0852 (0.0786)	-0.289** (0.114)	-0.0215*** (0.00743)	-0.432*** (0.124)	-0.0433*** (0.00982)
QLD	-0.220*** (0.0844)	-0.315*** (0.104)	-0.0230*** (0.00635)	-0.370*** (0.110)	-0.0374*** (0.00904)
SA	-0.0164 (0.0841)	-0.178 (0.111)	-0.0138* (0.00773)	-0.289** (0.120)	-0.0302*** (0.0106)
WA	-0.149* (0.0786)	-0.0336 (0.125)	-0.00287 (0.0104)	-0.0722 (0.136)	-0.00851 (0.0153)
TAS	-0.140 (0.0992)	0.0232 (0.104)	0.00206 (0.00939)	-0.0708 (0.109)	-0.00834 (0.0123)
NT	-0.431*** (0.0888)	0.0935 (0.123)	0.00879 (0.0127)	-0.0282 (0.126)	-0.00340 (0.0149)
Student teacher ratio (deviation from pop mean)		0.0229*** (0.00787)	0.00200*** (0.000728)	0.0319*** (0.00898)	0.00393*** (0.00107)
School's percentage of funding from government		0.000703 (0.00394)	6.15e-05 (0.000344)	0.000325 (0.00397)	4.01e-05 (0.000490)
School's percentage of funding from student fees		-0.00597 (0.00443)	-0.000522 (0.000398)	-0.00452 (0.00457)	-0.000557 (0.000562)
School is in a village		-0.0293 (0.122)	-0.00251 (0.0102)	-0.1000 (0.144)	-0.0115 (0.0154)
School is in a small town (300 to 15000 people)		0.372*** (0.0889)	0.0422*** (0.0137)	0.369*** (0.0908)	0.0573*** (0.0168)
School is in a town (15000 to 100000 people)		0.117 (0.0834)	0.0108 (0.00809)	0.111 (0.0906)	0.0143 (0.0121)
School is in a city (100000 to 1 million people)		0.159** (0.0750)	0.0148* (0.00763)	0.160** (0.0810)	0.0209* (0.0109)
Average School ESCS	0.0958				

	(0.0591)		
students are not grouped by ability within their classes	-0.0644*		
	(0.0348)		
residence in a particular area is not considered for student admission	0.0876**		
	(0.0360)		
student records are not considered for student admission	0.0152		
	(0.0336)		
regional or national educ authorities do not influence instructional content	-0.153***		
	(0.0460)		
Constant	0.230***	-1.806***	2.014***
	(0.0816)	(0.403)	(0.672)
Athrho	0.700***		
	(0.254)		
Observations		11,399	5,474
Number of clusters (schools)		321	308
Likelihood		-9149	-1434
Restricted likelihood			-1664
chi2		114.8	401.7
Rho		0.604	
Pseudo R2			0.138

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A2: Multilevel estimations of the PISA scores of students

VARIABLES	Dependent variable: PISA Scores		
	Reading	Maths	Science
Age (deviation from pop mean)	22.24*** (2.284)	23.80*** (2.160)	27.15*** (2.573)
Male	-36.79*** (1.461)	14.19*** (1.383)	0.172 (1.641)
Indigenous	-37.51*** (2.684)	-33.84*** (2.539)	-41.02*** (3.020)
Mother did not complete Year 12	-3.074** (1.468)	-3.003** (1.388)	-6.439*** (1.656)
Father did not complete Year 12	-12.38*** (1.503)	-10.39*** (1.421)	-10.17*** (1.695)
Dad is a blue collar	-3.751** (1.478)	-4.636*** (1.398)	-7.082*** (1.667)
Dad is a blue collar	-7.935*** (2.104)	-6.602*** (1.990)	-7.304*** (2.373)
Student non Australian born	-20.39*** (2.545)	-16.23*** (2.407)	-17.91*** (2.870)
Mother non Australian born	6.624*** (1.840)	6.202*** (1.740)	3.783* (2.073)
Father non Australian born	-0.0484 (1.783)	-0.982 (1.686)	-2.719 (2.007)
Student speaks language other than English at home	-19.09*** (2.913)	0.470 (2.756)	-15.68*** (3.279)
Student does not have a desk to study	-5.390** (2.625)	-6.017** (2.482)	-6.123** (2.961)
Home education resources (reference, nb of books at home >100)			
books 0 to25	-38.91*** (1.970)	-39.19*** (1.863)	-49.77*** (2.220)
books 26 to 100	-19.66*** (1.545)	-19.25*** (1.461)	-25.39*** (1.742)
No quiet place to study	-8.231*** (2.221)	-7.646*** (2.101)	-7.943*** (2.506)
No link to internet	-18.18*** (2.408)	-13.39*** (2.277)	-17.78*** (2.715)
Minutes of class time reading per week (pop mean: 296.57)	0.00623 (0.00786)	-0.0401*** (0.00743)	-0.0323*** (0.00888)
Minutes of class time maths per week (pop mean: 298.83)	0.0436*** (0.00793)	0.0699*** (0.00750)	0.0577*** (0.00896)
Minutes of class time science per week (pop mean: 231.36)	0.0869*** (0.00556)	0.0908*** (0.00526)	0.125*** (0.00628)
School average minutes of reading class	-0.141** (0.0555)	-0.0383 (0.0535)	-0.0764 (0.0531)
School average minutes of maths class	0.103* (0.0597)	0.0331 (0.0576)	0.0411 (0.0572)
School average minutes of science class	-0.00440 (0.0398)	0.0598 (0.0384)	0.0536 (0.0377)
Student ESCS (deviation from pop mean)	7.107*** (1.281)	6.076*** (1.211)	5.665*** (1.446)
Average School ESCS	55.99*** (17.51)	62.67*** (16.90)	68.33*** (16.64)

Male interacted with School ESCS	12.55***	7.272**	10.79**
	(3.876)	(3.672)	(4.296)
Computers _per_student (pop mean 2006: 0.265)	1.062	14.15	15.17
	(11.26)	(10.88)	(10.62)
school experiences shortage of qualified teachers	4.563	4.532	3.628
	(3.261)	(3.151)	(3.066)
Student teacher ratio (deviation from pop mean: 13.03)	0.0372	0.0323	0.230
	(0.568)	(0.548)	(0.538)
constant pressure from many parents about academic performance	5.324*	3.494	7.280**
	(3.225)	(3.117)	(3.025)
Percent government funding (pop mean: 70.09)	-0.327*	-0.191	-0.0358
	(0.171)	(0.165)	(0.161)
Percent fee funding (pop mean: 25.16)	-0.262	-0.150	-0.0741
	(0.206)	(0.199)	(0.194)
students are not grouped by ability within their classes	-1.308	-3.716	-2.555
	(2.855)	(2.760)	(2.683)
residence in a particular area is not considered for student admission	5.667*	6.341**	8.793***
	(3.317)	(3.205)	(3.122)
student records are not considered for student admission	-5.410*	-1.534	0.421
	(2.884)	(2.787)	(2.710)
school competes with two or more schools for its students	3.205	4.925	2.232
	(3.915)	(3.781)	(3.700)
regional or national educ authorities do not influence staffing	-8.731**	-6.765*	-8.655**
	(3.720)	(3.595)	(3.498)
regional or national educ authorities do not influence budgeting	2.904	3.338	1.314
	(3.503)	(3.385)	(3.298)
regional or national educ authorities do not influence instructional content	-4.068	-4.116	-2.994
	(3.525)	(3.406)	(3.317)
States (reference ACT: Australian Capital Territory)			
New South Wales (NSW)	8.563	13.80	23.70**
	(12.28)	(11.86)	(11.63)
Victoria (VIC)	14.10	21.49*	21.50*
	(12.22)	(11.80)	(11.56)
Queensland (QLD)	25.78**	42.99***	44.88***
	(12.24)	(11.82)	(11.59)
South Australia (SA)	16.10	24.00**	29.44**
	(12.35)	(11.93)	(11.70)
Western Australia (WA)	23.08*	28.90**	39.05***
	(12.70)	(12.26)	(12.03)
Tasmania (TAS)	23.82*	33.29***	33.65***
	(12.52)	(12.09)	(11.84)
Northern Territory (NT)	-17.83	-5.031	3.422
	(12.94)	(12.49)	(12.26)
States interacted with Socioeconomic status (ESCS)			
nsw_sescs	-17.38	-30.09*	-36.22**
	(18.34)	(17.71)	(17.35)
vic_sescs	-42.51**	-42.59**	-44.21**
	(18.33)	(17.70)	(17.31)
qld_sescs	-28.82	-48.23***	-50.43***
	(19.10)	(18.45)	(18.06)
sa_sescs	-49.11**	-40.37**	-46.84***
	(19.12)	(18.47)	(18.12)
wa_sescs	-33.25	-21.38	-42.71**

	(22.49)	(21.71)	(21.34)
tas_sescs	-14.24	-42.75**	-31.44*
	(19.64)	(18.96)	(18.58)
nt_escs	9.535**	6.757*	7.246*
	(3.814)	(3.607)	(4.299)
Constant	547.5***	491.1***	519.6***
	(15.82)	(15.27)	(15.03)
Random effects: Var(const)	404.5***	382.9***	307.8***
	(44.2)	(41.1)	(38.7)
var (residuals)	4819.2	4307.9	6151
	(63.38)	(56.70)	(80.9)
Variance partition coef	0.077	0.082	0.048
Observations	11,880	11,880	11,880
Number of groups	309	309	309

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1