

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU



HARVARD LIBRARY
Office for Scholarly Communication

Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time: An Application to Indonesia's Results in PISA Mathematics

Citation
Barrera-Osorio, Felipe, Vincente Garcia-Moreno, Harry Anthoney Patrinos, and Emilio Porta. 2011. Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time: An Application to Indonesia's Results in PISA Mathematics. World Bank Policy Research Working Paper 5584.

Published Version
http://go.worldbank.org/SD6EK6KDX0

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:8140107

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Public Disclosure Authorized

Public Disclosure Authorized

WPS5584

Policy Research Working Paper 5584

Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time

An Application to Indonesia's Results in PISA Mathematics

Felipe Barrera-Osorio
Vicente Garcia-Moreno
Harry Anthony Patrinos
Emilio Porta

The World Bank
Human Development Network
Education Team
March 2011



Policy Research Working Paper 5584

Abstract

The Oaxaca-Blinder technique was originally used in labor economics to decompose earnings gaps and to estimate the level of discrimination. It has been applied since in other social issues, including education, where it can be used to assess how much of a gap is due to differences in characteristics (explained variation) and how much is due to policy or system changes (unexplained variation). The authors apply the decomposition technique in an effort to analyze the increase in Indonesia's score in PISA mathematics. Between 2003 and 2006, Indonesia's score increased by 30 points, or 0.3 of a standard deviation. The test score increase is assessed in relation to family, student, school and institutional characteristics. The gap over time is decomposed into its constituent components based on the estimation of cognitive achievement production functions. The decomposition results suggest that almost the entire test score increase is explained by the returns to characteristics, mostly related to student age. However, the authors find that the adequate supply of teachers also plays a role in test score changes.

This paper is a product of the Education Team, Human Development Network. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at hpatrinos@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time:

An Application to Indonesia's Results in PISA Mathematics*

Felipe Barrera-Osorio[1]

Vicente Garcia-Moreno[2]

Harry Anthony Patrinos[3]

Emilio Porta[4]

JEL codes: J15, I20

Keywords: PISA, education, test scores, Indonesia

* Address correspondence to Harry Anthony Patrinos (hpatrinos@worldbank.org). The comments of Dandan Chen and Menno Pradhan are gratefully appreciated. The views expressed here are those of the authors and should not be addressed to the World Bank Group.

[1] World Bank
[2] World Bank
[3] Teachers College, Columbia University
[4] World Bank

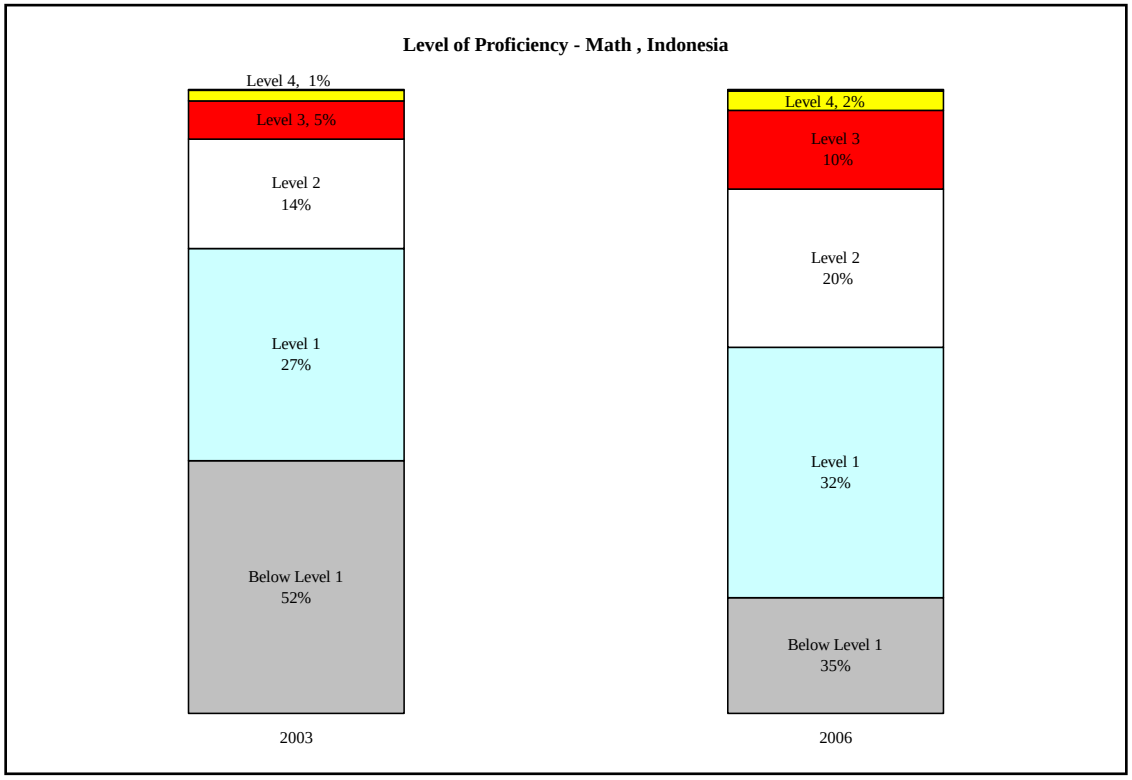
Introduction

The Oaxaca-Blinder technique was originally used in labor economics to decompose earnings gaps and to estimate the level of discrimination. For earnings differentials, the use of multivariate regression analysis allows for the simulation of alternative outcomes and the decomposition of gross differentials. The decomposition method, the technique used for analyzing earnings differentials, was popularized in the economics literature by Oaxaca (1973) and Blinder (1973). It was used earlier in sociology (Siegel 1965; Duncan 1968), and before that in demography (Kitagawa 1955). Although in the economics literature it was first used to analyze the determinants of male/female earnings differentials, the decomposition technique has been used since to analyze ethnic earnings differentials, public/private sector earnings differentials, earnings differentials by socioeconomic background, to test the screening hypothesis, and to test the effectiveness of a job training program, among other uses. It has been applied since in other social issues, including education, where it can be used to assess how much of a gap is due to differences in characteristics (explained variation) and how much is due to policy or system changes (unexplained variation).

We apply the decomposition technique in an effort to analyze the increase in Indonesia's score in PISA mathematics. The test score increase is assessed in relation to family, student, school and institutional characteristics. The gap over time is decomposed into its constituent components based on the estimation of cognitive achievement production functions. The decomposition results suggest that almost the entire test score increase is explained by the returns to characteristics, mostly related to student age. However, we find that the adequate supply of teachers also plays a role in test score changes.

Indonesia has participated in the PISA – the OECD's Programme for International Student Assessment, an internationally standardized assessment administered to 15 year olds in schools – since its first round in 2000. There have been two subsequent rounds since then in 2003 and 2006. Over time, Indonesia has maintained a steady score in science with 393, 395, and 393 points in 2000, 2003 and 2006. The average score among OECD countries is 500 points and the standard deviation is 100 points. Indonesian students have steadily improved their score in reading over time, from 371 in 2000 to 382 in 2003 and 393 points in 2006, an increase of about 10 points, or a respectable 0.10 of a standard deviation, in each round. In math, there was no improvement between 2000 and 2003 (scores of 367 and 360 points), but there was a dramatic improvement in 2006, to 391 points, an increase of 0.30 of a standard deviation – or almost one full school year equivalent – in just three years. Figure 1 shows how the change occurred. In 2003, 80 percent of Indonesian students scored at the lowest levels, level 1 and -1. These are significantly low achievement levels, effectively denoting functional illiteracy. A typical student at level 1 or -1 may be able to read words but will not be able to decode the information they contain. By 2006, the number of students scoring at level -1 decreased drastically, while the proportions at higher levels went up. Nevertheless, there were very few students at the higher levels from 4 and above (and none at all at levels 5 and 6).

Figure 1:



Most developing countries score at the bottom of the scale in most international achievement tests. Until recently, there were very few if any examples of developing countries that had achieved significant improvements in these tests. Critics argue that the international development community has focused almost exclusively on increasing enrollment in the education sector and has ignored the need for that education to be of adequate quality. However, Indonesia is a rare case of a developing country that has achieved some progress.

In order to find out what lay behind Indonesia's exceptional improvement in 2006, we looked at how family, student, school and institutional inputs may have affected the increase in the test score of 15 year olds in math. We decomposed the increase over time into its constituent components using the traditional Oaxaca-Blinder method, based on the estimation of a cognitive achievement production function. Our decomposition results suggest that almost all of the test score increase was unexplained or, in other words, was due to changes in the *returns* to characteristics rather than to changes in the characteristics themselves. We found that most of the positive change was due to the increased returns over time to the variable representing a student's age, which varied only by months in this case (as the PISA is administered to a randomly selected sample of students who are between the ages of 15 years and 3 months and 16 years and 2 months at the time of the test).

Empirical evidence on education production functions exists for both developed countries (for example, Hanushek, 1986 and 2002) and developing countries (for example, Glewwe 2002). Previous empirical studies do not always agree on which school and family inputs improve children's achievement. For example, there is some disagreement about the role played by schooling inputs such as class size, teacher experience, teacher education, and mother's employment. For a survey of related literature, see Todd and Wolpin (2003).

Nevertheless, although a child's achievement is inherently individual in nature, a large body of evidence points to the existence of persistence effects in educational achievement across generations (Fertig 2003; Fertig and Schmidt 2002; Currie and Thomas 1999). Consequently, it is necessary to control for the characteristics of individual students as well as for their family backgrounds. Similarly, it is necessary to control for the characteristics of the school environment as well as its institutional arrangements. Recent evidence from the literature on early test score differentials suggests that differences in children's cognitive ability among families appear at an early age, tend to persist, and may even widen with age. In general, "good" families promote cognitive, social, and behavioral skills, while "bad" families do not. This is important in determining what policy interventions can be successful (Carneiro and Heckman 2003). Evidence also suggests that socioeconomic and family background variables, such as the education levels of a student's parents and the number of books a child has, are very important determinants of test scores at early ages (Fryer and Levitt 2002).

Methodology and Estimation

Our first step was to specify and estimate cognitive achievement production functions that relate student achievement to individual, family, school, and institutional inputs. We then proceeded to decompose the over-time test score change into an explained component (accounting for student, family, school and institutional characteristics) and an "unexplained" component (the efficiency by which the country is able to convert characteristics into student learning outcomes as measured by test scores), using the traditional Oaxaca-Blinder decomposition method (Oaxaca 1973; Blinder 1973).

The model specification that we used to estimate the production function for cognitive achievement is as follows:

$$T_{ija} = T_a(A_{ija}, F_{ija}, S_{ija}, I_{ija}) + \epsilon_{ija} \qquad (1)$$

where T_{iaj} is the observed test score (from the PISA math test) of student i in household j at time a (the time of the test), A_{ija} is a vector of individual student characteristics, F_{ija} is a vector of parent inputs, S_{ija} is a vector of school-related inputs, I_{ija} is a vector of the school's institutional characteristics, and ϵ_{ija} is an additive error, which includes all the omitted variables including those that relate to the history of past inputs, endowed mental capacity, and measurement error. Todd and Wolpin (2003) discuss in detail the assumptions that would satisfy the application of this specification, in which the achievement test score depends solely on contemporaneous measures of family, school, and other inputs. These assumptions state that: (i) current input measures capture the entire history of inputs or, alternatively, only contemporaneous inputs matter and (ii) contemporaneous inputs are unrelated to endowed mental capacity. Its linear specification (after dropping subscript a) is given by:

$$T_{ij} = \beta_0 + \beta_1 A_{ij} + \beta_2 F_{ij} + \beta_3 S_{ij} + \beta_4 I_{ij} + \epsilon_{ij} \qquad (2)$$

where β_0 to β_4 are the coefficients to be estimated. The standard procedure for analyzing the determinants of the test score differences over time is to fit equations between test scores and observed characteristics. The observed test score differential can be decomposed as:

$$T_{2006} - T_{2003} = (X_{2006} - X_{2003})\beta_{2006} + X_{2003}(\beta_{2006} - \beta_{2006}) \qquad (3)$$

where T is the standardized test score, X_i is a vector of student, family, school, and institutional characteristics for the ith individual, β is a vector of coefficients, and 2006 and 2003 subscripts are identifiers of the PISA test score in math in years 2003 and 2006 evaluated at 2006 prices. The overall increase in test score can, therefore, be decomposed into two components. One is the portion attributable to differences in characteristics ($X_{2006} - X_{2003}$) evaluated at 2006 prices or to the performance of the 2006 group of students (β_{2006}), while the other portion is attributable to differences in the effects on performance ($\beta_{2006} - \beta_{2003}$) of 2003 and 2006 students derived from the same characteristics. This second (unexplained) component, while more difficult to interpret in the present context than an earnings gap decomposition framework, may have had more than one explanation. The first and most obvious explanation is that the unexplained portion of the test score increase may reflect certain unobserved family characteristics that are correlated with achievement over time, possibly related to household wealth. The second possible explanation may be that, given that enrollments are rising over time in Indonesia and more students from disadvantaged backgrounds are entering the school system, teachers may pre-judge these students as underachievers and, therefore, use different teaching standards with them than with other students (Ferguson, 1998). A third explanation may be that different cohorts of students do not reap the same benefits from equivalent school and classroom resources. Finally, the differences in the returns may reflect the impact over time of past reforms that both increased school enrollments and helped to improve the quality of school inputs in Indonesia.

Some of these coefficient estimates may be subject to biases. For example, if a school characteristic is correlated with unobserved family characteristics that influence achievement (such as family wealth and parents' motivation), then the effect of attending a school with such characteristics may be biased.

Modified Decomposition

An alternative decomposition is possible using a modified Oaxaca-Blinder method, in which the unexplained part of the test score differential is captured by a year indicator (2006) taking the value of 1 for 2006 and 0 otherwise (2003). Consider a production function for cognitive achievement:

$$T_{ija} = T_a(2006_{ij}, A_{ija}, F_{ija}, S_{ija,} I_{ija}) + \epsilon_{ija} \qquad (4)$$

where 2006_{ija} is a dummy variable equal to 1 if the test was taken in 2006 and 0 otherwise.

In implementing a modified Oaxaca decomposition of the test score gap and assuming a linear specification, the differences of mean test scores for 2006 and 2003 students is given by:

$$(T^{2006} - T^{2003}) = \beta_1 + \beta_2(A^{2006} - A^{2003}) + \beta_3(F^{2006} - F^{2003}) + \beta_4(S^{2006} - S^{2003}) + \beta_5(I^{2006} - I^{2003}) \qquad (5)$$

where coefficient β_1 is an estimate of the portion of the change that remains after accounting for the differences in mean characteristics.

To arrive at the proportions that are explained and unexplained:

$$\beta_1 / (T^{2006} - T^{2003}) = \textit{unexplained} \text{ and}$$

$$\frac{\beta_2(A^{2006} - A^{2003}) + \beta_3(F^{2006} - F^{2003}) + \beta_4(S^{2006} - S^{2003}) + \beta_5(I^{2006} - I^{2003})}{(T^{2006} - T^{2003})} = \text{explained}$$

and the components of the explained portion are:

$\beta_2(A^{2006} - A^{2003}) = \textit{individual characteristics}$

$\beta_3(F^{2006} - F^{2003}) = \textit{family}$

$\beta_4(S^{2006} - S^{2003}) = \textit{school}$

$\beta_5(I^{2006} - I^{2003}) = \textit{institutional factors}$.

While test scores and individual and family information are at the individual level, school resources and other school-related inputs are at the school level. In choosing the estimation method, we recognized that observed test scores can be expected to be correlated at the school level due to clustering effects. Therefore, the assumption that disturbances are independently and identically distributed with fixed conditional variance did not hold. As a result**,** we used the estimation method of OLS by cluster at the school level.

Data

The PISA is an international assessments initiated by the OECD. It assesses 15 year olds in each participating country in three main subject areas – reading, mathematics, and scientific literacy. We focused on the results for Indonesia in mathematics in the assessments for 2003 and 2006. We did not include information for 2000 even though it was available, because the sample was very different. For instance, dataset for the 2000 survey has much fewer observations regarding parents' education than the 2003 and 2006 surveys; while there were 8,828 and 9,292 observations in 2003 and 2006, in 2000 the sample contained only 2,777 observations. In short, we do not believe that the 2000 sample is comparable with subsequent rounds.

Instead of testing the knowledge and skills specified in the national curricula of the participating countries, the PISA aims to test the ability of students to apply their acquired knowledge in the three subject areas in real-life situations. The targeted student population falls between the ages of 15 years and 3 months and 16 years and 2 months who are enrolled in the seventh grade or higher. Indonesia uses a two-stage sampling frame with a cluster design. We applied weights to the data at the student level. The PISA standardizes the data for OECD countries with the mean at 500 points and the standard deviation set to 100. Thus, it is the OECD means and standard deviation that are the benchmarks for the other participating countries.

Table 1: Sample Means, PISA 2003 and 2006, Indonesia

	Pisa 2003				Pisa 2006			
Variable	Mean	S.D.	Missing	N	Mean	S.D.	Missing	N
Institutions								
School determines pedagogy	0.99	0.09	0%	10,761	0.95	0.22	0%	10,647
Adequate supply of teachers	0.46	0.50	1%	10,691	0.71	0.45	1%	10,493
Schools								
Public	0.54	0.50	1%	10,704	0.60	0.49	1%	10,493
Students repeating (%)	0.01	0.04	6%	10,133	0.01	0.03	7%	9,867
Rural	0.32	0.47	1%	10,669	0.26	0.44	2%	10,414
Students								
Grade *8*	0.15	0.36	0%	10,761	0.12	0.33	0%	10,647
9	0.49	0.50	0%	10,761	0.40	0.49	0%	10,647
10	0.35	0.48	0%	10,761	0.44	0.50	0%	10,647
11	0.02	0.13	0%	10,761	0.04	0.21	0%	10,647
Age	15.71	0.27	0%	10,761	15.78	0.29	0%	10,647
Female	0.50	0.50	0%	10,756	0.49	0.50	0%	10,647
Family								
No education	0.15	0.35	2%	10,545	0.14	0.34	1%	10,503
Mother schooling: *Primary*	0.35	0.48	2%	10,545	0.35	0.48	1%	10,503
Lower secondary	0.15	0.36	2%	10,545	0.19	0.39	1%	10,503
Upper secondary	0.17	0.37	2%	10,545	0.22	0.41	1%	10,503
University	0.18	0.38	2%	10,545	0.11	0.31	1%	10,503
Books at home *11-100*	0.58	0.49	10%	9,639	0.69	0.46	4%	10,241
101-500	0.08	0.27	10%	9,639	0.10	0.30	4%	10,241
Home computer (1 or more)	0.16	0.37	0%	10,743	0.15	0.35	4%	10,245
Home language same as test	0.32	0.47	4%	10,364	0.34	0.47	1%	10,517
PISA score	360	74.9	100%	10,761	391	75.3	100%	10,647

Description of the Sample

The means for the variables that we used to analyze the determinants of learning are presented in Table 1. The PISA data contain missing values among the family background characteristic variables, and in. Table 1, we show where those missing values occur. While some might choose to impute the missing values, we decided not to do so in this case. Therefore, if a variable had any missing values, we dropped the observation in its entirety from our analysis.

We realize that deleting cases with missing values can have dangers, as demonstrated by Little and Rubin (1987). Deleting cases is based on the assumption that the deleted cases occur at random and are a relatively small representative proportion of the entire dataset. However, this may not necessarily be the case. The missing data may be indicative of some pattern and cannot safely be assumed to reflect randomness. In such circumstances, deletion can introduce substantial bias into the study. Moreover, the loss in sample size can appreciably diminish the statistical power of the analysis.

As a rule of thumb, if a variable has more than 5 percent missing values, it is advisable not to delete cases, and many researchers are much more stringent than this (Little and Rubin, 1987). Deleting incomplete cases has its attractions, mostly the virtue of simplicity, but one loses information in doing so. This approach also ignores the possible systematic difference between the complete cases and incomplete cases, and the resulting inference may not be applicable to the population associated with all cases, especially with a smaller number of complete cases to take into account. Some techniques exist to impute missing values, ranging from correlations, single imputation, and a multiple imputation procedure (Rubin, 1987). However, very few of our variables had missing values that made up more than 5 percent of the total. Overall, the sample for 2003 dropped from 10,761 students to 8,828, and in 2006, the sample went down from 10,647 to 9,293.

The mean scores associated with each characteristic increased over time (Table 2). The scores for students whose mothers had a university education were much higher in 2006 than in 2003, at more than half a standard deviation. Speaking the same language at home that is used in school increased scores by more in 2006 than in 2003. The largest increase was for children with at least one computer at home – a 66 point increase or the equivalent of two years of learning.

Another important change is the score associated with the school autonomy variable titled "the school determines pedagogy." In 2003, those schools that did not determine pedagogy scored higher than those that had autonomy over their own pedagogy, but by 2006, the opposite was true. Also, the association between gender and math scores changed over time. In 2003, there was little difference in overall scores between boys and girls, but by 2006, boys scored 17 points higher girls in math. Given that we did not impute, we knew there was a possibility that our analysis would be biased. To minimize this risk**,** we examined mean scores by variable for two samples in each year (see Annex Table 1). One was the regression sample, which did not include observations with any missing value, and the other was the full PISA sample. The regression sample, despite its (small) number of missing values, was not very different from the full PISA sample in terms of outcomes. The differences in math scores by characteristic did not vary appreciably, by as little as 1 point in some cases and by no more than 10 points in others. Overall, the scores differed by an average of only 4 points. On a scale with a mean of 500 and a

standard deviation of 100, these are not very large numbers. Also, when we examined the differences in means between the two years, it became apparent that the regression sample was more urban and public school-oriented in both years but particularly in 2003. However, we found that the overall mean test score of the regression sample was very similar to the whole sample mean. Therefore, we concluded that the regression sample was not biased.

Table 2: PISA 2003-2006, Mean Math Scores by Selected Characteristics

	Pisa 2003		Pisa 2006	
	Yes	*No*	*Yes*	*No*
Institutions				
School determines pedagogy	360	371	393	360
Adequate supply of teachers	357	363	396	371
School				
Public	374	344	404	372
Rural	335	371	364	401
Student				
8th grade	313		342	
9th grade	348		366	
10th grade	395		424	
11th grade	413		423	
Female	358	362	382	399
Family				
Mother - no education	347		369	
Mother – primary	350		377	
Mother - lower secondary	360		389	
Mother - upper secondary	398		417	
Mother – university	359		417	
0 - 10 books	358		383	
11 - 100 books	363		393	
101- 500 books	391		415	
Home computer > 1	387	355	453	382
Home language same as test	362	361	402	385

Source: PISA, 2003 and 2006

Regression Results

There are significant premiums associated with attending a public school and with attending a school that was able to determine its own pedagogy or, in other words, had been granted school autonomy (Table 3). There is some controversy about private and public schools in Indonesia. James et al (1996) found that private schools were better managed in Indonesia than public schools, and they argued that private management is more efficient than public management in achieving academic quality. There is also some evidence that private funding also increases efficiency whether the schools are publicly or privately managed. Bedi and Garg (2000) examined the effectiveness of public and private schools in Indonesia using the labor market earnings of their graduates as the measure. Controlling for observable personal characteristics and school selection, they found that graduates of private secondary schools performed better in the labor market than their peers from public secondary schools, contrary to the widely held belief in Indonesia that public secondary schools are superior. Suryadarma et al (2006) compared public primary schools with the smaller sample of private primary schools. They found that, on average, students in the private schools performed marginally better academically than their counterparts in public schools, but the only statistically significant difference was in mathematics performance. The mean differences were slight—less than three points on a 0-100 scale, or 0.11 standard deviations. This suggests that the differences in performance between public and private schools may not be very large. Newhouse and Beegle (2005) evaluated the impact of school type on the academic achievement of junior secondary school students (in grades 7 to 9). They found, after controlling for a variety of other characteristics, that students who graduated from public junior secondary schools scored 0.15 to 0.30 standard deviations higher on the national exit exam than their comparable privately schooled peers. This finding was robust to OLS, fixed-effects, and instrumental variable estimation strategies. The authors also found that students attending Muslim private schools, including Madrassahs, fared no worse academically on average than students attending secular private schools. The authors argued that the results provided indirect evidence that higher quality inputs at public junior secondary schools than at private schools of the same level promote higher test scores.

In our samples, the adequate supply of teachers was associated with higher test scores in 2006. The coefficient for 2003 was statistically no different from zero, whereas the coefficient in 2006 was significant. We also found that the higher the percentage of students who repeated a grade in the school, the greater the significant and negative effect on scores. Living in a rural area had a negative effect, although fewer people lived in rural areas in the 2006 sample than in the 2003 sample and the coefficient was slightly less negative. The negative effect of being female actually increased in 2006. The effect of parental education had some unexpected effects in 2003. In the case of the mother's education, only having a mother with upper secondary schooling had a positive effect. By 2006, all of the signs had become positive, with having a mother with secondary schooling having had the largest effect. Having a large number of books at home used to be associated with a large positive coefficient, but by 2006, this variable was no longer significant. However, having a computer at home had a large and significant positive effect, an effect which grew in 2006. Overall, for both years, the samples were large (8,391 students in 2003 and 8,660 in 2006), representing 1.5 million and 1.8 million students in 2003 and 2006. The 2006 model seems to be more robust, with an R-square of 0.35, compared with an R-square of 0.26 for 2003.

Decomposition Results

The purpose of doing these decompositions was to investigate what changes may have occurred over time that would help us to explain the 30-point increase in math scores in Indonesia between 2003 and 2006. It seems clear that the 2006 score was partly the result of reforms, policies, strategies, and interventions that were put in place years ago, even a generation ago. For example, between 1973 and 1978, the Indonesian government engaged in one of the largest school construction programs on record (the INPRES program). Duflo (2001) studied the effects of this program by combining differences among regions in the terms of the number of schools that were built with differences among different cohorts of students induced by the timing of the program. Her research suggested that each primary school constructed per 1,000 children led to an average increase of 0.12 to 0.19 years of education as well as to a 1.5 to 2.7 percent increase in wages for that cohort. This implies total returns to education ranging from 6.8 to 10.6 percent. This huge increase in school places no doubt had a positive effect on the schooling outcomes of successive generations, including the 2006 class. Figure 2 shows that the change over time represented a shift of students towards higher levels of education and less inequality between the highest and the lowest achieving students.

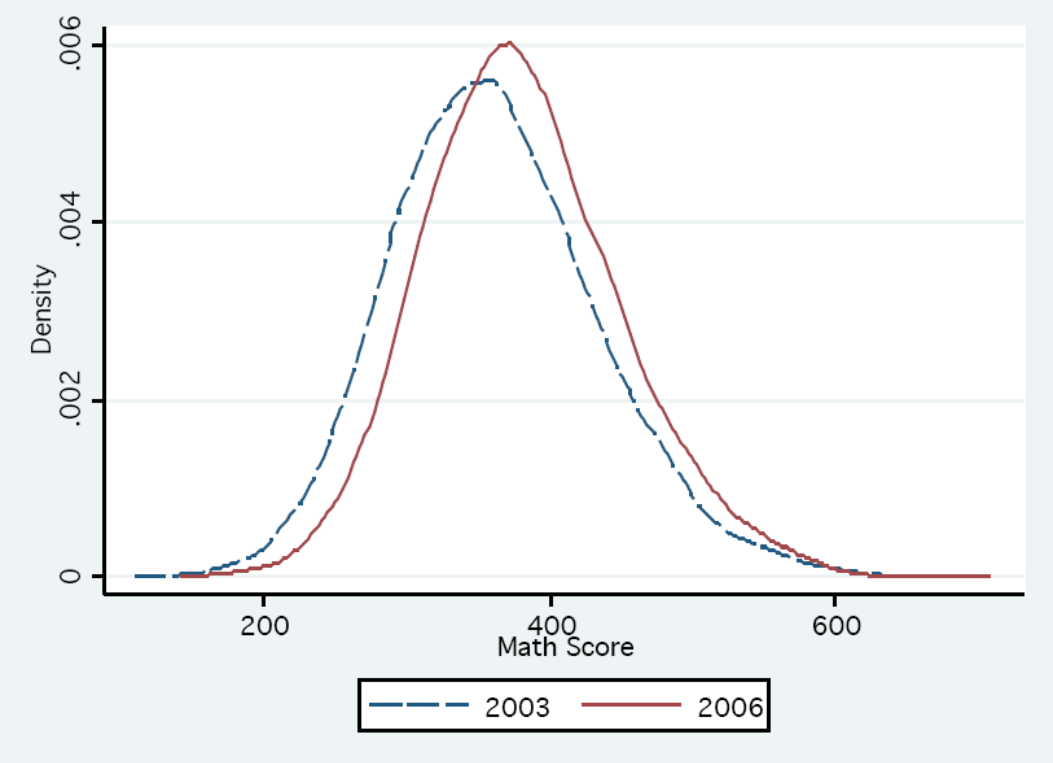
Table 3: Determinants of Learning, Indonesia, PISA

	Pisa 2003				Pisa 2006			
	Coef.	S.E.	Regression mean	Sample mean	Coef.	S.E.	Regression mean	Sample mean
Institutions								
School determines pedagogy	38.55	(4.75)*	0.99	0.99	34.20	(8.24)*	0.96	0.95
Adequate supply of teachers	0.55	(3.71)	0.47	0.46	8.74	(3.81)*	0.71	0.74
Schools								
Public	36.15	(3.59)*	0.70	0.55	33.26	(3.89)*	0.67	0.60
% students repeating grade	-0.90	(0.28)***	1.27	0.90	-1.35	(0.30)*	0.81	0.74
Rural area (<3,000)	-15.72	(0.28)*	0.23	0.31	-14.32	(3.59)*	0.20	0.26
Urban (3,000 and above)			0.77	0.69			0.80	0.74
Student characteristics								
Grade								
8th			0.13	0.15			0.12	0.12
9th	25.79	(1.78)*	0.46	0.49	25.03	(2.58)*	0.46	0.40
10th	74.14	(3.87)*	0.39	0.35	75.63	(4.50)*	0.38	0.44
11th	87.35	(5.60)*	0.03	0.02	78.19	(6.09)*	0.05	0.04
Age	-11.66	(1.87)*	15.71	15.71	-9.45	(2.00)*	15.76	15.78
Female	-7.57	(1.30)*	0.51	0.51	-18.60	(2.51)*	0.51	0.49
Family background								
Mother's education - none			0.14	0.15			0.14	0.14
Mother -Primary	-3.50	(1.86)***	0.31	0.35	7.14	(2.35)*	0.32	0.35
Mother's education - lower sec	-3.66	(2.19)	0.17	0.15	10.38	(2.66)*	0.20	0.19
Mother's education - upper sec	21.37	(2.79)*	0.20	0.17	15.06	(3.24)*	0.25	0.22
Mother's education -university	-2.45	(3.44)	0.19	0.18	7.94	(4.38)***	0.11	0.11
Books at home								
None–10 books			0.33	0.34			0.21	0.21
11–100 books	4.56	(1.22)*	0.59	0.58	1.76	(1.28)	0.69	0.69
101-500 books	22.36	(2.07)*	0.09	0.08	1.30	(1.77)	0.10	0.10
Computers at home								
None			0.82	0.84			0.85	0.85
One or more than one	16.00	(2.21)*	0.18	0.16	41.42	(3.78)*	0.15	0.15
Language speak at home								
Test language speak at home	-12.25	(2.18)*	0.32	0.32	-9.66	(3.34)*	0.34	0.34
Constant	453.94	(30.79)*			437.60	(32.33)*		
Observation	8,549				8,688			
R^2	0.26				0.34			
Total Sample	10,761	0.79			10,647	0.82		

Source: Program for International Student Assessment (PISA) 2003 and 2006

Notes: *** 90%; **95%; *99%

Figure 2: Distribution of Test Scores over Time



Before looking at the results of the decomposition, we examined over-time changes in characteristics and returns. Overall, there were not very many changes in characteristics. The adequate supply of teachers increased considerably, by 24 percentage points, and the returns associated with it increased significantly as well. The percentage of grade repeaters in a school declined significantly; the penalty associated with repeating fell also. More importantly, we could see that there had been a change in the schooling profile of the parents. More parents had primary, lower, and upper secondary schooling in 2006 than in 2003, and the proportion of parents with a university education had gone down. This was probably the result of two trends: first, the level of education of adult Indonesians has been rising steadily over time – thus increasing the proportion of parents with primary and secondary (instead of none) – and, second, student access to secondary schooling has been going up – thus reducing the proportion of parents with a university education in the sample. The returns to mothers' education went up at most levels of education except upper secondary, which was already very high in 2003. Meanwhile, there had been small declines in fathers' educational level of attainment at all levels.

The detailed decomposition is presented in Annex Table 2, while Table 4 shows the results of the decomposition. Almost all of the difference is "unexplained," which, in terms of an over-time decomposition of changes in test scores, means that most of the over-time increase is due to higher returns to all characteristics. Simply put, Indonesia in 2006 was able to convert the characteristics in question into higher levels of learning.

Table 4: Decomposition of Math Scores over Time (as percentage of total test score differential)

	Endowments (Characteristics)	Unexplained (Returns)
Constant	0.0	-99.4
Institutions	0.2	19.2
Schools	10.9	6.6
Family	-8.8	37.1
Student	5.5	128.6
Total	7.8	92.2
Overall	100.0	

The bulk of the overall difference resulting from changes in the returns to characteristics was due to student characteristics. That is, for a given set of student characteristics, Indonesian schools were more able to convert those factors into higher levels of learning in 2006 than in 2003. This is a significant finding since more and more children enter the lower secondary school system with every passing year. As we have shown above, most of the new entrants come from poorer backgrounds and from homes with parents who have received less schooling. According to the 2006 math PISA scores, Indonesia was better able to educate students regardless of their age. Math scores for 2003 and 2006 by mean age are presented in Table 5. In 2003 and 2006, the average-aged student achieved the mean score for the country overall. In 2003, there was not much variation in the ages of students who were one standard deviation above the mean or of those who were one standard deviation below the mean age. With a mean age of 15.71 years and a standard deviation of 0.27 years in 2003, students were between 15½ and almost 16 years of age. In 2006, the mean age was 15.78 years and the standard deviation was 0.29 years, thus the range was 15½ to just over 16 years of age. Scores were higher for all age groups in 2006 but also varied more than in 2003, with 16 year olds averaging 400 points.

Table 5: Average Math Scores by Age

	Below (mean – 1 sd)	Between (mean - 1 sd) and (mean + 1 sd)	Above (mean + 1 sd)
2003	357	360	363
2006	383	390	400

Source: PISA 2003 and 2006

As Annex Table 2 shows, two other characteristics are important in explaining the differences in test scores between 2003 and 2006. An adequate supply of teachers, both in terms of endowments and coefficients, played a positive effect in increasing test scores between 2003 and 2006. Also, in terms of the unexplained part of the decomposition, the coefficient associated with being a female changed from -7.57 in 2003 to -18.6 in 2006.

The results of the alternative decomposition (equation 4) are presented in Annex Table 3, and the overall results are presented in Table 6. The results are in line with the results that we got from the more traditional decomposition.

Table 6: Alternative Decomposition - Determinants of PISA Differentials

PISA Scores	*as % of total test score diff*	
	$b_{2006}(X_{2006}-X_{2003})$	$b_{2006}(X_{2006}-X_{2003})/(T2006-T2003)$
Difference $T_{2006}-T_{2003}$	30.8	
Time	19.5	0.6
Institutions	0.2	0.0
Schools	-1.6	-0.1
Family	-0.2	0.0
Student	1.0	0.0

Sources: PISA 2003 and 2006; authors' calculations

The main explanation behind the change in test scores between 2006 and 2003 is a fixed-time effect that yielded a 19.5 incremental increase in the score. The observable characteristics contributed only marginally to the change in test scores between 2003 and 2006. It is noteworthy that the characteristics of institutions and students made a positive contribution to the positive change in test scores, whereas schools and family played a negative role.

Discussion

It is very impressive that Indonesia was able to achieve such a gain in test scores in math given the increased enrollment of disadvantaged children in the school system (see Figure 3). As enrollments in lower secondary schooling continue to increase, more and more students from families with less well-educated parents are entering the school system. For example, in 2003, the average level of education attained by the fathers of the 15-year-old students was 9.26 years, and for their mothers it was 8.30 years. By 2006, the average level of the fathers' schooling had fallen to 9.09 years, while the mothers' level had fallen to 8.16 years.

Figure 3: Changes in Parental Education and Students' Scores over Time

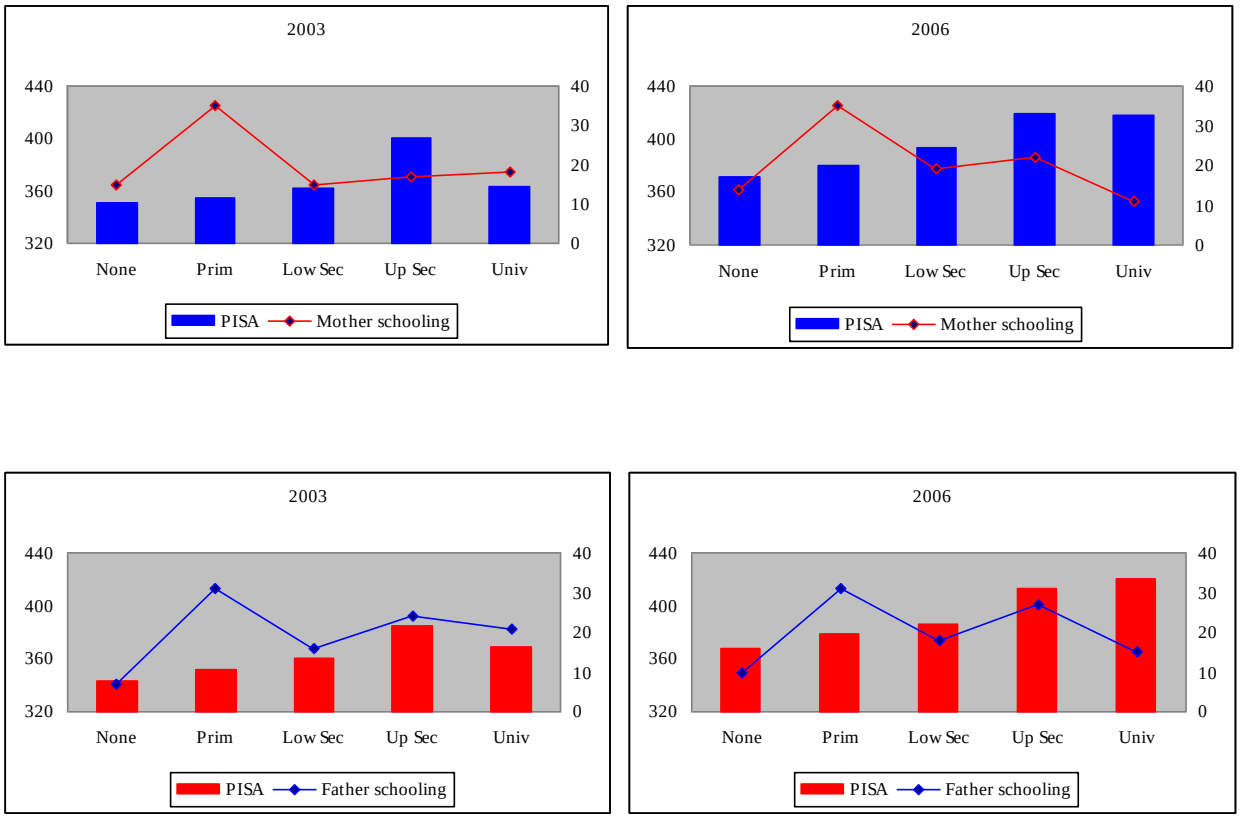


Table 7 presents the variables that are listed in the 2006 dataset. The more institutional variables that we included in our analysis, the more interesting were the findings that emerged. Among other things, firing teachers, which is an indicator of school autonomy, was significant. Also, if a school was having to compete with others in the vicinity, then the effect was large, positive, and significant. Parental involvement in formulating the school budget was also positive and significant. Public schools retained their large advantage. An adequate supply of math teachers played a positive role in the determination of test scores, while grade repetition had a small negative effect. The level of the mother's schooling was significant. Doing math work in class was also important.

Table 7: Determinants of Learning, PISA 2006 - Extended Model

	Coef.	S.E.
School determines pedagogy	25.3	(7.99)*
School competes for students*	35.8	(6.72)*
School can fire teachers	14.1	(6.74)**
Achievement data used to evaluated teachers	-4.9	(4.91)
Public	40.5	(5.64)*
Student-teacher ratio	-0.3	(0.13)*
Adequate supply of teachers	11.3	(3.75)*
Math hours	12.5	(0.37)*
Teachers with certificate	-4.5	(4.11)
% of students repeating grade	-1.7	(0.39)*
Rural (<3,000)	-10.8	(3.45)**
9^{th} Grade	15.6	(2.51)*
10^{th} Grade	65.7	(4.03*
11th Grade	65.9	(6.25)*
Age	-8.2	(1.81)*
Female	-19.3	(2.30)*
Mother's education - Primary	5.6	(2.37)*
Mother's education - Lower secondary	9.1	(2.55)*
Mother's education - Upper secondary	13.8	(3.48)*
Mother's education - University	9.7	(3.94)*
11–100 books (regressor: none-10 books)	1.0	(1.30)
101-500 books (regressor: none-10 books)	1.5	(1.54)
One or more computer at home	35.3	(4.26)*
Language speak at home (language of test)	-4.5	(2.79)
Constant	353.9	(32.13)*
Observation	7,746	
R^2	0.43	
Total Sample	10,761	0.72

Source: Program for International Student Assessment (PISA) 2006
Notes: *** 90%; ** 95%; *99

Despite the impressive gains that were made by Indonesian students in math in the 2006 PISA, Indonesia still has a long way to go to improve its academic standing. In 2006, almost three-quarters of 15-year-olds scored at level 1 and below. Too few students scored at levels 2 and 3, and an insignificant number scored at levels 4 or above. Understanding the reasons why the scores increased in 2006 should help the Government of Indonesia to build on its strengths and make further improvements in the future.

Conclusions

In the 2006 PISA, Indonesia's score in math increased by 30 points, or 0.3 of a standard deviation, in just three years. We explored the reasons behind this increase by Indonesia's 15-year-old students in relation to various family, student, school, and institutional inputs. We decomposed the change over time into its constituent components using the traditional Oaxaca-Blinder method, based on the estimation of a cognitive achievement production function. Our decomposition results suggest that almost all of the test score increase was unexplained, or, in

other words, was due to changes in the returns to the characteristics rather than due to changes in the characteristics themselves. To put it another way, Indonesia was able to better educate its students in 2006 than in 2003 regardless of the characteristics of those students.

References

Bedi A.S. and A. Garg (2000). "The effectiveness of private versus public schools: the case of Indonesia." *Journal of Development Economics* 61(2): 463-494.

Blinder, A. (1973). "Wage discrimination: Reduced form and structural estimates." *Journal of Human Resources 8*(4): 436–455.

Carneiro, P. and J. Heckman (2003). "Human capital theory." NBER Working Paper 9495.

Currie, J. and D. Thomas (1999). "Early test scores, socioeconomic status, and future outcomes." NBER Working Paper no. 6943.

Duflo, E. (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91(4): 795-813.

Ferguson, R. F. (1998). "Teachers' perceptions and expectations and the Black-White test score gap," in C. Jencks and M. Phillips (eds.), *The Black White Test Score Gap*. Washington, DC: Brookings Institution Press.

Fertig, M. (2003). "Who is to blame? The determinants of German students' achievement in the PISA 2000 study." IZA Discussion Paper 739.

Fertig, M. and C. M. Schmidt (2002). "The role of background factors for reading literacy: Straight national scores in the PISA 2000 study." IZA Discussion Paper no. 545.

Fryer, R. and S. Levitt (2002). "Understanding the Black-White test-score gap in the first two years of school." NBER Working Paper no. 8975.

Glewwe, P. (2002). "Schools and skills in developing countries: Education policies and socioeconomic outcomes." *Journal of Economic Literature* 40(2): 436-82.

Green, W.A. (2000). *Econometric Analysis* (4th edition). Prentice Hall.

Hanushek, E. (1986). "The economics of schooling: Production and efficiency in public schools." *Journal of Economic Literature* 24(3): 1141-1177.

Hanushek, E. (2002). "Publicly provided education," in A.J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics* (vol. 4). Amsterdam: Elsevier.

Hernandez-Zavala, M., H. Patrinos, C. Sakellariou and J. Shapiro (2006). "Quality of schooling and quality of schools for indigenous students in Guatemala, Mexico, and Peru." Policy Research Working Paper 3982, World Bank, Washington, D.C.

James, E., E. King and A. Suryadi (1996). "Finance, management, and costs of public and private schools in Indonesia." *Economics of Education Review* 15(4): 387-398.

Little, R.J.A. and D.B. Rubin (1987). *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

McEwan, P.J. (2004). "The indigenous test score gap in Bolivia and Chile." Department of Economics, Wellesley College, Wellesley, MA (January).

Newhouse, D. and K. Beegle (2005). "The effect of school type on academic achievement: evidence from Indonesia." World Bank Policy Research Working Paper Series 3604.

Oaxaca, R. (1973). "Male-female wages differentials in urban labor markets." *International Economic Review 14*(3): 693–709.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Suryadarma, D., A. Suryahadi, S. Sumarto and F. H. Rogers (2006). "Improving Student Performance in Public Primary Schools in Developing Countries: Evidence from Indonesia." *Education Economics* 14(4): 401-429.

Todd, P.E. and I. Wolpin (2003). "On the specification and estimation of the production function for cognitive achievement." *Economic Journal* 113: F3-F33.

World Bank (2005). *Mexico: Determinants of Learning Policy Note* (Report No. 31842-MX) Latin America and the Caribbean, Human Development.

Annex Table 1: PISA 2003-2006, Mean Math Scores by Selected Characteristics

	2003 regression		2003 sample		2006 regression		2006 sample	
	Yes	No	Yes	No	Yes	No	Yes	No
School determines pedagogy	365	373	360	371	397	363	393	360
School can fire teachers	351	381	347	374	383	408	373	402
Achievement data used*	366	355	361	352	396	386	391	380
Public	379	348	374	344	410	375	404	372
Rural	340	376	335	371	367	405	364	401
8th grade	319		313		343		342	
9th grade	352		348		368		366	
10th grade	398		395		428		424	
11th grade	416		413		427		423	
Female	363	367	358	362	386	405	382	399
Mother - no education	351		347		371		369	
Mother - primary	355		350		380		377	
Mother - lower secondary	363		360		394		389	
Mother - upper secondary	401		398		421		417	
Mother - university	364		359		427		417	
0 - 10 books	358		358		385		383	
11 - 100 books	365		363		396		393	
101 - 500 books	393		391		419		415	
Home computer > 1	393	360	387	355	455	384	453	382
Home language same as test	365	365	362	361	409	388	402	385
Scores	**365**		**360**		**394**		**391**	

Source: PISA 2003 and 2006
Note: * to evaluate teacher and principal performance

Annex Table 2: Decomposition of PISA Scores for Indonesia, 2003-2006

	b_{2003}	b_{2006}	X_{2003}	X_{2006}	Determinants of Test scores Differentials			
Test Scores							*as % of total test score diff*	
					Endowments	*Unexplained*	*Endowments*	*Unexplained*
					$b_{2006}(X_{2006}-X_{2003})$	$X_{2003}(b_{2006}-b_{2003})$		
Constant	453.94	437.60	1.00	1.00	0.00	-16.34	0.0	-83.4
Institutions								
School determines pedagogy	38.55	34.20	0.99	0.96	-0.85	-4.29	-4.4	-21.9
Adequate supply of teachers	0.55	8.74	0.47	0.71	2.08	3.85	10.6	19.6
Schools								
Public operation	36.15	33.26	0.70	0.67	-1.06	-2.02	-5.4	-10.3
% of students repeating grade	-0.90	-1.35	0.01	0.01	0.01	-0.01	0.0	0.0
Rural area (<3,000)	-15.72	-14.32	0.23	0.20	0.43	0.33	2.2	1.7
Student characteristics								
9^{th}	25.79	25.03	0.46	0.46	0.03	-0.35	0.1	-1.8
10^{th}	74.14	75.63	0.39	0.39	-0.30	0.58	-1.5	3.0
11th	87.35	78.19	0.03	0.05	1.88	-0.23	9.6	-1.2
Age	-11.66	-9.45	15.70	15.76	-0.57	34.79	-2.9	177.5
Female	-7.57	-18.60	0.51	0.51	0.13	-5.65	0.7	-28.8
Family background								
Mother -Primary	-3.50	7.14	0.30	0.31	0.07	3.22	0.4	16.4
Mother - Lower secondary	-3.66	10.38	0.17	0.20	0.28	2.36	1.4	12.0
Mother - Upper secondary	21.37	15.06	0.20	0.25	0.74	-1.28	3.8	-6.5
Mother -University	-2.45	7.94	0.19	0.11	-0.64	1.96	-3.2	10.0
11–100 books	4.56	1.76	0.59	0.69	0.17	-1.64	0.9	-8.4
101-500 books	22.36	1.30	0.09	0.10	0.02	-1.79	0.1	-9.1
One or more than one	16.00	41.42	0.18	0.14	-1.74	4.58	-8.9	23.3
Language speak at home (language of test)	-12.25	-9.66	0.39	0.40	-0.14	1.00	-0.7	5.1
Total					0.54	19.06	2.7	97.3
Overall					19.60		100	

Source: Program for International Student Assessment (PISA) 2003 and 2006

Annex Table 3: Alternative Decomposition of PISA Scores for Indonesia, 2003-2006

	Pisa 2003-2006				
			Means		
	Coef.	S.E.	2003	2006	$b_{2006}(X_{2006}-X_{2003})$
Time					
PISA 2006	19.51	(2.6-)*	0	1	***19.5***
Institutions					
School determines pedagogy	37.29	(7.15)*	0.99	0.96	-0.93
Adequate supply of teachers	4.79	(2.37)**	0.47	0.71	1.14
Schools					
Public operation	34.97	(2.59)*	0.70	0.67	-1.15
% of students repeating grade	-1.11	(0.21)*	0.01	0.81	-0.89
Rural area (<3,000)	-15.47	(2.49)*	0.23	0.20	0.45
Urban (3,000 and above)			0.77	0.80	
Students					
Grade					
8th			0.15	0.12	0.00
9th	25.15	(1.66)*	0.46	0.46	0.08
10th	76.00	(2.94)*	0.39	0.38	-0.53
11th	81.76	(4.44)*	0.03	0.05	1.96
Age	-10.47	(1.39)*	15.71	15.76	-0.62
Female	-13.69	(1.53)*	0.51	0.51	0.11
Family background					
Mother - No schooling			0.15	0.14	0.00
Mother -Primary	2.34	(1.64)	0.31	0.32	0.02
Mother - Lower secondary	4.17	(1.84)**	0.17	0.20	0.12
Mother - Upper secondary	17.90	(2.54)*	0.20	0.25	0.84
Mother -University	2.82	(3.06)	0.19	0.11	-0.23
Books at home					0.00
None–10 books			0.34	0.21	0.00
11–100 books	3.34	(0.94)*	0.59	0.69	0.33
101-500 books	10.92	(1.31)*	0.09	0.10	0.16
Computers at home					
None			0.82	0.87	0.00
One or more than one	28.86	(1.76)*	0.18	0.13	-1.27
Language at home same as test	-10.35	(2.10)*	0.38	0.40	-0.16
Constant	433.33	(22.97)*			
Observation	17,237				-0.57
R^2	0.32				
Total Sample	10,761	1.60			

Source: Program for International Student Assessment (PISA) 2003 and 2006
Notes: ** 90%; ** 95%; * 99%