Improving education services: District governance and student learning in Indonesia

Pradhan, M.; de Ree, J.

DOI
10.4337/9781784715571.00018

Publication date
2015

Document Version
Final published version

Published in
Governance in Developing Asia

Governance in Developing Asia

Public Service Delivery and Empowerment

Edited by

Anil B. Deolalikar
University of California, Riverside, USA

Shikha Jha
Asian Development Bank, Philippines

Pilipinas F. Quising
Asian Development Bank, Philippines

CO-PUBLICATION OF THE ASIAN DEVELOPMENT BANK AND EDWARD ELGAR PUBLISHING

Edward Elgar PUBLISHING

Cheltenham, UK • Northampton, MA, USA

ADB

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

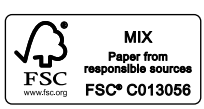
Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

Asian Development Bank
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
Tel +63 2 632 4444
Fax +63 2 636 2444
www.adb.org

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2014954962

This book is available electronically in the Elgaronline
Social and Political Science subject collection
DOI 10.4337/9781784715571



ISBN 978 1 78471 556 4 (cased)
ISBN 978 1 78471 557 1 (eBook)

Typeset by Servis Filmsetting Ltd, Stockport, Cheshire
Printed and bound in Great Britain by T.J. International Ltd, Padstow

9. Improving education services: district governance and student learning in Indonesia

Menno Pradhan and Joppe de Ree

1. INTRODUCTION

Increasing evidence suggests that low education outcomes in developing countries partly result from the low efficiency of public education spending. Recorded teacher absenteeism rates in developing countries in the range of 16–27 percent starkly illustrate that education spending does not always lead to better education services (Chaudhury et al. 2006).[1] The World Bank's 2004 *World Development Report*, likewise, notes a lack of correlation between changes in education spending and outcomes at the country level; it argues that improving accountability for public spending is crucial to improving education outcomes (World Bank 2004).

Comparing countries, empirical evidence suggests that bad governance hampers the efficiency of public spending in education. The two studies most relevant for this chapter use corruption perception as an indicator of governance. Comparing public education spending in 57 countries, Rajkumar and Swaroop (2008) find that whereas public spending on education as a share of gross domestic product does not have a significant relationship to levels of educational attainment, the interaction between public spending and governance does. Countries with lower corruption are better able to ensure that public education spending translates into higher educational attainment for their citizens.

In Indonesia, Suryadarma (2012), comparing across districts, finds a relationship between enrollment levels and corruption, although the relationship vanishes when national test score results are used as the outcome variable. It is difficult to pinpoint the mechanisms behind such observed patterns, although one obvious one is that public funds are used for private gains in corrupt states and, thus, dilute the funds available for education services, resulting in lower education outcomes. Yet, corruption perceptions could also indicate other failures in education systems.

Rajkumar and Swaroop (2008) find very similar results when a perception of bureaucracy is used instead of corruption. Perceptions could also be influenced by more general indications that the system is failing, as exemplified by poor education services or high dropout rates. For this reason, policy conclusions have to remain very general about the importance of good governance.

This chapter focuses on Indonesia, which presents a case study of the effects of governance on education outcomes. Indonesia decentralized most government functions, including basic education, to the district level in 2001 (Kaiser et al. 2006). Districts now receive most of their revenues as block grants from the central government and have considerable leverage in how to spend them. District-level elections for parliaments and direct elections for district heads are now the main mechanisms to ensure local accountability, and the availability of cross-district data allows the study of governance using comparable data for districts that inherited a common institutional setting.

The Governance and Decentralization Surveys, a collaboration between the World Bank and the Centre of Public and Policy Studies at Gajah Mada University, are the most systematic attempt so far to measure governance at the district level in Indonesia. The surveys, done in 2002 and 2006 in a large sample of districts, collected information on governance focusing on thematic areas such as participation, effectiveness and efficiency; transparency; social equity; rule of law; responsiveness; accountability and conflict management (Widyanti and Suryahadi 2008).

Two studies have used data from these surveys to assess the relationship between governance and education service delivery. Lewis and Pattinasarany (2009) use the second round of the survey in 2006 to analyze the relationship between satisfaction with education services, objective measures of service delivery and governance. They find generally high levels of satisfaction and a positive relationship with objective measures of service delivery and governance perceptions. As expected, satisfaction increases with better classrooms, more accessible schools, lower student–teacher ratios and, perhaps surprisingly, with a higher share of teachers who are civil servants.

Alatas and Filmer (2004), in an unpublished paper, relate the results of the first round of the Governance and Decentralization Survey of 2002 to educational outcomes. They conducted their analysis at the district level, comparing averages across about 140 districts,[2] and took data from various sources. They took test scores—the dependent variable—from junior secondary school examination results, school quality from central government registers, district spending data from the Ministry of Finance, household background data from the National Socioeconomic Household

Survey and governance data from the Governance and Decentralization Survey. The results are mixed. Focusing on the governance variables, the authors find that higher test scores for both junior and senior secondary schools are associated with more interaction between local nongovernment organizations and government, an increased role for principals to set goals and missions and the greater involvement of teachers in recruitment. Most of the time, however, the effects are insignificant or even contradict a positive contribution of governance.

This chapter aims to contribute to the empirical evidence on the relationship between education governance and outcomes. It casts the analysis in a school-level production function, rather than a district-level production function of education, thereby focusing on what happens at school. Our study includes a measure of school financial and human assets, and governance. By focusing on changes in test scores rather than levels, the empirical strategy effectively controls for unobserved time invariant characteristics, which inevitably vary with learning outcomes in levels.

As explanatory variables, the study includes indicators of the physical and human capital of schools and indicators of district teacher management. It also controls for learning levels at the beginning of the school year. Learning outcomes are the product of inputs received throughout life, even before entering school and, here, parents obviously play a key role (one of the main reasons children from wealthier, better-educated parents do better in school). But the problem is that these parents tend to live in areas with better service delivery altogether, which generates a correlation between learning achievements and district governance. To value the importance of governance systems, one needs to control for this.

District governments not only provide resources, but can also play an important role in teacher management. Differences in governance across districts should therefore translate into differences in learning across districts. This study provides evidence in support of this claim. It documents substantial variability in student-learning trajectories across districts, with district effects accounting for about 15 percent of the variation in test score gains between schools where gains are measured over a two-year period.

Our study explains this finding. It includes variables in the regression model that are, to a greater or lesser extent, related to or influenced by district leadership. Two of the key dimensions considered in the study relate to efforts in teacher management: the average number of teacher evaluations and the participation rate in teacher working groups. The first refers to the more traditional style of teacher management, in which a superior evaluates the performance of a subordinate. In this study we focus on whether evaluations have been conducted in addition to the legally required minimum. The second dimension focuses on improving

teacher skill and motivation, with teacher working groups enabling educators to interact with peers at a professional level. Additional dimensions considered include student–teacher ratios, school budgets and student socioeconomic background.

The results indicate that schools in districts with more active teacher management policies achieve higher gains in learning, indicating, for example, that support to teacher working groups pays off. But whereas some indicators appear significant in these regressions, they account for only a fraction of the variation in student-learning outcomes. This suggests that while there is substantial structural variation in achievement gains across districts, these differences cannot be explained by policy relevant measurable characteristics.

The international evidence on teacher management is that the link between in-service teacher training and student learning is weak. A review of the effects of school resources in developing countries (Glewwe et al. 2011) finds that in six higher-quality studies, three show the positive effects of teacher training on student learning, but the others show no effect. Several high-quality studies from the United States (US) found no effect of in-service teacher training on student learning (Jacob and Lefgren 2004; Harris and Sass 2011). The findings resonate with a key conclusion in the literature on teacher effectiveness, that teachers are the main factor in education production, but the literature is so far inconclusive in pinpointing why (Rivkin et al. 2005). This study finds evidence that district governance matters, but the regressors in the model can explain only a fraction of this variation. These results make intuitive sense, because it is likely that it is the quality of spending of school budgets or the support of teacher working groups that matters, rather than just the amounts. Further research could shed light on this issue when better measures on governance are combined with gains in student-learning achievement.

The chapter presents the context for the study, describes the data, provides descriptive statistics and discusses the models and estimation results.

2. CONTEXT OF THE STUDY

For Indonesia, the challenge lies in turning increased public spending on basic education into better quality education. Public spending has historically been low, but has been rapidly increasing since 2001's decentralization of several government functions, including basic education. By 2010, 21 percent of public spending was on education, up from 11 percent in 2001 (World Bank 2013a). A constitutional amendment in 2002, requiring both local and central governments to spend at least 20 percent of their

budgets on education, is driving this trend. In many other countries that decentralized, public spending on social sectors subsequently increased (del Granado et al. 2005).

Indonesia's enrollment levels in basic education are high, with gross enrollment in primary school at 118 percent,[3] junior secondary at 92 percent and senior secondary at 39 percent, with only small differences between boys and girls (United Nations Educational, Scientific and Cultural Organization 2012). That said, learning achievements have been disappointing.

The country scored 386 on a scale benchmarked at 500 in the 2011 Trends in International Mathematics and Science Study (TIMSS) assessment of the mathematical skills of children in grade 8 aged at least 13 (Mullis et al. 2011). This score represents the median student in Indonesia. In the US, the 5th percentile in the test distribution scores stood at 410. The median student in Indonesia thus scores below the 5th percentile student in the US. The 75th percentile in Indonesia roughly corresponds to the level of the 10th percentile in the US. And the trend is negative. In comparison with the 2007 TIMSS results, Indonesia scores 11 points lower. Comparing Indonesia with Malaysia presents a more positive picture. Malaysia scores roughly in the middle between the US and Indonesia with a median student scoring 440.

Indonesia does not lack teachers. The average student–teacher ratio is 17 for primary and 12 for junior secondary education, which is very low compared with the global average of 31 students per teacher. But absenteeism of full-time teachers is considerable. This was at 20.1 percent in 2002, although it dropped to 14.8 percent in 2007, based on estimates from the findings of unannounced visits in a random set of public primary schools during school hours (Toyamah et al. 2010). Teacher absenteeism is higher in remote areas, at 23 percent in 2007 for schools that received a "remote area allowance" and 33 percent in Papua, according to a recent United Nations Children's Fund survey (Universitas Cenderawasih et al. 2012).

Even when teachers are at their best, there is scope for improving teaching methods. A video study (World Bank 2010a) investigated the teaching methods of mathematics teachers included in the TIMSS assessment. Compared to mathematics teachers in developed economies (Australia; Hong Kong, China; the Czech Republic; the Netherlands; Switzerland; and the US) the most striking difference was that Indonesian teachers spoke much less in class and their students even less. Indonesian teachers spoke for less than half the time of teachers in the comparator countries and for every 28 words the teacher spoke, their students spoke just one. In these economies, this ratio ranged from 8:1 to 16:1.

Correlations with the TIMSS scores indicate that classes with a higher degree of student–teacher engagement also had higher scores.

A more effective teacher appraisal process could therefore improve learning. Currently, teacher appraisal is the same as that for other civil servants; that is, teachers are evaluated yearly using an evaluation list of tasks performed. This assesses allegiance, performance, responsibility, compliance, honesty, cooperation, initiative and leadership. But because these are the same performance criteria for all civil servants, the evaluation method takes no account of teaching performance, making it an inappropriate instrument for the task. Principals conduct the evaluation and the results feed into salary increases, promotions and transfers (Turner et al. 2009). The appraisal system still bears the characteristics of structures in place before decentralization. Back then, loyalty to nationally designed policies was considered all-important, even to the extent that it hindered centrally promoted local decision making in schools (Bjork 2004).

Other hindrances are a cultural trait in Indonesia that considers exposing other people's weaknesses inappropriate, and a mismatch between formal in-service training opportunities and the needs of teachers for professional upgrading. Teachers often attend training simply to gain credit points for promotion and learn little (Listiawati 2003).

However, district offices and head teachers can organize their own evaluations in addition to the nationally mandated evaluation list. The empirical analysis of this study focuses on these additional teacher evaluations.

The national law on teachers and lecturers, Law No. 14 passed in 2005, introduced new elements into the appraisal process.[4] It defined teacher competencies in pedagogic, personal, social and professional settings, and incorporated these into national teacher standards. Teacher certification was introduced, entitling certified teachers to an additional allowance one times their base salary.[5] The main requirement for certification is a four-year bachelor's degree and a workload of at least 24 teaching hours per week. Until recently, the evaluation was done on the basis of a portfolio that demonstrated experience and qualifications, but with no formal test. The evaluation is now done before teachers enter the certification process. However, much of the effectiveness of the certification program depends on the effectiveness of the testing tools in providing the right incentives to teachers to improve their skills. The certification program also introduced a strong incentive to meet the criteria of certification, resulting in big demand for additional qualifications, particularly the bachelor's degree, as this is a requirement for entry into the certification process (World Bank 2010b).

Teacher working groups are perhaps the most accessible form of teacher upgrading. These have existed for over 20 years in Indonesia and allow

teachers to discuss problems and work with educators from other schools on common tasks, such as curriculum development, teaching aids and the design of tests, as well as more advanced activities, such as lesson study and classroom action research (World Bank 2013b). For primary schools, working groups are organized around classroom teachers, while for secondary schools subject teachers are grouped together (Musyawarah Guru Mata Pelajaran). Teachers see the benefits of these groups as improved knowledge, skills, competencies and professionalism. Two studies on the topic, in 2007 and 2010, indicated a greater emphasis on training and development of lesson plans in these groups and less emphasis on developing tests or the dissemination of local government initiatives (World Bank 2013b).

District governments can support teacher working groups by providing travel funds and training activities. In a review of the teacher appraisal processes in Indonesia, Listiawati (2003) argues that teacher study meetings provide the most promising instrument for teacher appraisal, as they already contain an element of intercollegial review of teaching aspects.

3. DATA USED IN THE STUDY

The study presented here uses two datasets that include test scores from a cohort of students, reports on teacher evaluations and participation in teacher working groups, teacher characteristics, school infrastructure and school budgets. Both datasets were collected for different research purposes, but lend themselves well to this analysis.

The school committee dataset was collected during 2007–2008 for testing various strategies to strengthen these bodies. See Pradhan et al. (2014) for more details of this study and the data.[6] The study was conducted in 517 public primary schools in six districts in Central Java and Yogyakarta.[7] Schools were selected at random and excluded subdistricts with very few schools, schools with parallel classes and schools with extremely good or bad average 6th grade examination scores in mathematics or Indonesian. The study followed a cohort of children in grade 4 at the baseline and in grade 6 at the end line. It also interviewed the teachers of this cohort and collected school-level data from principals in both rounds.

The teacher certification dataset was collected during November 2009–April 2012 in 360 public primary and junior secondary schools in 20 districts to evaluate the effect of the teacher certification program. See World Bank (2013b) and de Ree et al. (2013) for more details on this study and the data.[8] A stratification procedure was applied to ensure the sample had enough teachers eligible to be certified. This ensures a more

Table 9.1 Variation in education indicators across 26 districts

	10th percentile	Median	90th percentile
Student–teacher ratio	8.13	21.86	30.49
School budget per year (million Rp)	37.00	240.00	399.00
Number of students	80.00	165.00	248.00
Share of female teachers	0.47	0.67	0.81
Share of teachers with a bachelor's degree	0.15	0.41	0.73
Share of teachers with a civil service status	0.71	0.83	0.92
Working group meetings*	2.58	4.05	9.62
Frequency of evaluation*	1.61	2.97	6.84

Notes:
Rp = rupiah.
* See further details in Table 9.2.

Source: Authors' calculations based on school committee and teacher certification datasets.

homogenous teacher population across districts, which is also beneficial for this analysis. Within these strata, schools were sampled at random.

This study used a sample of 234 primary schools in 20 districts.[9] It interviewed all classroom teachers and collected school-level data from principals. To ensure comparability with the test score data from the school committee dataset, panel data for grade 4 children at baseline and grade 6 children at the end line were used (most of which were the same children tested two years later).

That both datasets were collected to test different research questions should not influence this analysis. In both cases, the interventions (to strengthen school committees or to provide preferred access to the certification program) were balanced across districts. This study mainly focused on cross-district variation and, by design, there was no cross-district variation in the interventions tested in the two studies. School-level averages were constructed for all variables used in the regression models. The study also constructed 26 district dummy variables, which are included in some of the specifications.

Table 9.1 shows the spread of the indicators across the 26 districts, with considerable variability evident among the indicators. At the bottom end, for example, only 15 percent of primary teachers have a bachelor's degree, compared to 70 percent at the top end.[10] School budgets also differ widely across observations. One important reason for these differences is most likely the Bantuan Operasional Sekolah program, which provides block

Table 9.2 Descriptive statistics on teacher management, mean of district averages

	Teacher certification data	School committee data
Number of meetings with teacher working group per year	5.24	. . .
Hours devoted to teacher working group/ training meetings per month	. . .	3.26
Number of teacher evaluations per year	2.57	. . .
Usual frequency of evaluations (rescaled to number of evaluations per year)	. . .	6.62
Number of districts	20.00	6.00

Note: . . . = data not available. The surveys collected information on participation in teacher working groups and teacher evaluations in different ways.

Source: Authors' calculations based on school committee and teacher certification datasets.

grants to schools on the basis of enrollment. The program has intensified, meaning that its impact on the budget should be more pronounced for the teacher certification data, as a study conducted recently shows.

The two teacher management variables—working group meetings and frequency of evaluation—are not perfectly comparable between the two data sources. Table 9.2 shows the means of the two indicators and highlights the differences between the sources of information. The teachers were asked how often their performance is evaluated outside of their yearly evaluations. The phrasing is slightly different between questionnaires. In the school committee dataset, the question was asked in terms of the interval in months between evaluations, while in the certification survey it was asked in terms of the number of evaluations per year. For this analysis, the school committee data are transformed in terms of the number of evaluations per year.[11]

Comparing participation in teacher working groups in the two datasets is not straightforward, as participation is measured in hours per week in the school committee dataset and in frequency and number of visits per year in the teacher certification dataset. Here the variables cannot easily be placed on a uniform scale, which is a problem for the precise comparability of the regression parameters. To address the problem, this study adds 1 to the raw observations and takes logs before including the variables in its regressions. For the same reason, school budgets are included in logs to allow for different price levels in different time periods.

The median school has 165 students, which, with 6 grades in a primary school, implies a class size of about 27. Around 80 percent of teachers are civil servants and over half are female.

4. DISTRICT TEACHER MANAGEMENT POLICIES AND STUDENT LEARNING

This study estimates a school-level, value-added model. The regression equation takes the form of the lagged score value added specification often used by the literature.

$$Y_{sk6} = \alpha + \beta' X_{sk} + \gamma Y_{sk4} + \varepsilon_{sk6}$$

Where $Y_{sk6}(Y_{sk4})$ is the average grade 6 (4) test score of school s in district k. The test scores are standardized, so that the effect parameters β can be interpreted as standardized effects. X_{sk} are school-level characteristics, including teacher- or student-level indicators that are averaged at the school level. The model is estimated on the basis of a cohort of students (balanced panel) tested in grades 4 and 6.

Although in principle it would have been possible to estimate this model at the individual level by adding a subscript for the student, it was decided not to do this because our focus is on measuring and explaining the variation in learning outcomes across schools. We were also concerned about measurement error in the test scores. A large body of research recognizes the problem of measurement error in student test scores used on the right-hand side of a regression equation. This will normally bias the parameter γ downwards and also bias the effect parameter β in the process. Averaging scores within schools alleviates some of these concerns since much of the individual student specific measurement error in test scores is expected to average out when averaged within schools or classrooms.[12]

The model measures student learning over the course of about two years and evaluates whether these learning trajectories are related to some of the school-level characteristics X_{sk}. Such a specification is an improvement over models that use only the level of learning outcomes as the dependent variable, without controlling for the starting values. Learning outcomes are the result of cumulative inputs of parents and school into child development throughout life. It is well known, empirically, that children of wealthier, better-educated parents do better in school. But wealthier families are also more likely to live in areas that provide better services, such as better schools. Moreover, these particular (unobserved) parental inputs

may also drive district policies, because it is the same parents who vote for district governments. Correlations between student learning levels and district-specific characteristics can therefore be hard to interpret. But by controlling for lagged test scores in the regression model, thereby focusing on learning gains rather than levels, these heterogeneities are in some way accounted for.

Table 9.3 shows the results: columns 1–5 report the results based on the teacher certification data and columns 6–10 on the school committee data. Columns 1 and 6 show the results of the simplest value-added specification and only include the lagged grade 4 test score in the model. The results between the two data sources are different. Lagged scores explain about 49 percent of the variation in test scores two years later for the teacher certification data, yet only 13 percent in the school committee data.

Column 2 includes the district dummy variables in the regressions. The fit of the models increases a fair amount: the district dummies explain an additional 20 percent (teacher certification data) or 11 percent (school committee data) of the variation in the outcome scores.[13] A lot of the variation in gain scores in Indonesian primary schools can therefore be explained at the district level. Although this is relevant information in itself, policymakers can be helped by understanding the reasons why some districts do better than others.

Table 9.3, in columns 3–5 and 7–10, includes a set of relevant covariates in the regression models. With the teacher certification data, we find positive results on participation in teacher working groups and on the proportion of teachers with a bachelor's degree. With the school committee dataset, none of the included regressors is statistically significantly different from zero. The school resource variables are insignificant, although the signs of the effects are consistent across both data sources.

Whereas some of these factors appear significantly in the regression, they do not explain much of the variation in learning outcomes. Only 3 percent of additional variation in the outcome score can be explained by these factors in the teacher certification dataset, while none of the variation in the school committee data is explained. One likely explanation for the only moderate increase in the fit of the mode (as indicated by the R^2) is measurement error: R^2s are lower when outcome scores have more measurement error. Especially for the school committee data, this could be a problem. But because the district dummies explain much more of the variation in learning outcomes, this is taken as empirical support for the idea that district governance matters for learning. Still, it is hard to establish the reasons for these differences empirically. School budget and human resource indicators do not explain the differences across districts, nor does the frequency of meetings of teacher working groups. This makes sense

Table 9.3 School production function estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Teacher certification data					School committee data				
Grade 4 student scores	0.697*** (0.000)	0.453*** (0.000)	0.661*** (0.000)	0.667*** (0.000)	0.660*** (0.000)	0.359*** (0.000)	0.244*** (0.000)	0.324*** (0.000)	0.328*** (0.000)	0.323*** (0.000)
Student-teacher ratio			0.006 (0.517)	0.003 (0.728)	0.006 (0.528)			−0.010 (0.362)	−0.010 (0.350)	-0.010 (0.356)
School budget (in logs)			0.036 (0.331)	0.024 (0.496)	0.032 (0.394)			0.011 (0.745)	0.013 (0.702)	0.011 (0.752)
Number of students (in logs)			−0.146 (0.299)	−0.093 (0.522)	−0.140 (0.323)			−0.093 (0.560)	−0.108 (0.495)	−0.103 (0.513)
Share of female teachers			−0.176 (0.434)	−0.094 (0.670)	−0.178 (0.427)			0.253 (0.239)	0.261 (0.223)	0.249 (0.245)
Share of teachers with a bachelor's degree			0.494*** (0.005)	0.541*** (0.003)	0.493*** (0.006)			0.445 (0.134)	0.446 (0.128)	0.440 (0.135)
Share of teachers who are civil servants			−0.092 (0.642)	0.026 (0.896)	−0.078 (0.700)			0.221 (0.499)	0.206 (0.523)	0.225 (0.491)

Table 9.3 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Teacher certification data					School committee data				
Intensity of teacher working groups			0.274*** (0.005)		0.264*** (0.010)			0.057 (0.213)		0.059 (0.195)
Frequency of teacher evaluations				0.120 (0.195)	0.056 (0.555)				−0.051 (0.185)	-0.052 (0.173)
Constant	0.006 (0.892)	−0.607** (0.045)	−0.580 (0.511)	−0.423 (0.617)	−0.592 (0.502)	−0.000 (1.000)	−0.184* (0.093)	−0.162 (0.844)	0.017 (0.984)	-0.032 (0.969)
District dummies included	No	Yes	No	No	No	No	Yes	No	No	No
N	234	234	234	234	234	503	503	503	503	503
R^2	0.488	0.693	0.542	0.520	0.543	0.129	0.239	0.151	0.152	0.155
Adjusted R^2	0.486	0.664	0.525	0.503	0.524	0.127	0.230	0.137	0.138	0.139
Leave-one-out R^2	0.478	0.627	0.502	0.482	0.498	0.120	0.216	0.121	0.121	0.121

Note: Cluster (at school level) robust p-values in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: Authors' calculations based on school committee and teacher certification datasets.

because differences in learning outcomes are more likely due to the quality of budget spending and the quality of the engagement during working group meetings.[14]

Across the two sets of results there are important differences and similarities. For the teacher certification dataset, for example, much stronger correlations are found with the lagged outcome score than with the school committee dataset. In a related point, the R^2 is much higher when the teacher certification data is used, suggesting lower levels of noise in the score variables. The results also differ in the statistical significance of the effect parameters. Some regressors appear statistically significant with the teacher certification, whereas this is not seen with the school committee data.

The results based on the school committee and teacher certification data are similar, however, in the sense that district dummies explain much more of the variation in learning than the substantive variables included in the regressions.

5. CONCLUSIONS

This chapter investigated the importance of district governance and teacher management policies to student learning in Indonesian primary schools. Because Indonesia decentralized responsibility for the delivery of primary education to district governments, variations in teacher management policies and, therefore, learning trajectories could be expected. Indeed, the study documents substantial heterogeneity in learning gains. It also shows that schools with more active teacher working groups and higher qualified teachers achieved better learning gains. However, teacher management policy variables—including school budgets, participation rates in teacher working groups, and student–teacher ratios—explain only a fraction of the differences in learning across districts. It is therefore likely that the *quality* of operation matters.

The analysis was carried out by investigating correlates of student test-score gains across 737 primary schools across 26 districts in Indonesia. Unlike earlier studies in this field, a school production model was used that tries to explain changes in test scores, rather than test scores in levels.

The heterogeneity in learning gains across districts cannot be explained by financial or human resource indicators such as school budgets, student-teacher ratios and so on. In fact, the study concludes that learning in school does not correlate with these measures. It does find that the general education levels of teachers and the frequency of participation in teacher working groups positively correlate with learning in school. However,

these can also only account for a fraction of the differences in learning observed across Indonesian districts.

Documenting differences across districts is important. The differences suggest that districts that manage to improve their system can actually expect improvements in learning outcomes, which will increase their relative competitiveness in the region. Documenting the differences also serves as a starting point to deepen understanding of what works and what does not in Indonesia's education system. Future research could focus on understanding why some districts do better than others.

ACKNOWLEDGEMENTS

The authors thank Joanne Asquith, Emmanuel San Andres, Shikha Jha, Anil Deolalikar and other participants in the workshops sponsored by the Asian Development Bank (ADB) at the University of California, Riverside and ADB headquarters in Manila for useful comments on earlier drafts of this chapter. The authors are also grateful to the World Bank and the Government of Indonesia for sharing the data we used in this study.

NOTES

1. Teachers were recorded as absent if they could not be found in the facility for any reason at the time of a random, unannounced spot check. The range is based on representative studies conducted in Bangladesh, Ecuador, India, Indonesia, Peru and Uganda in 2003.
2. Referred to in the study as the unmatched sample. Results from the matched sample are not discussed, as the authors note that "the sample sizes in the matched sample are small, which makes it hard to identify effects".
3. Gross enrollment measures the number of children enrolled in school as a percentage of the number of children in the corresponding age category. Because children outside this age range can be enrolled, the percentage may end up above 100 percent.
4. Available in Indonesian at http://sa.itb.ac.id/Ketentuan%20Lain/UUNo142005%28Guru%20&%20Dosen%29.pdf.
5. Base salary depends on rank and years of experience. In the certification dataset, civil servant teachers reported an average of 1.6 million rupiah per month as their base salary, equivalent to about $160 at the average exchange rate in 2010.
6. The data were collected by Moores Roland under supervision of the Ministry of Education, with technical assistance of the World Bank and with the financial support of the Japan Social Development Fund grant number TF053814 and the Dutch Government.
7. The districts are Brebes, Gunung Kidul, Pemalang, Sragen, Temanggung and Wonogiri.
8. The data were collected by Survey Meter, the Ministry of Education, and World Bank field teams, with technical assistance of the World Bank, and with financial support of the Dutch Education Support Program. The tests used in the analysis are developed by the center for educational assessment of the Government of Indonesia.
9. The 20 remaining districts are Bantul, Bengkulu, Ciamis, Deli Serdang, Gowa, Hulu

Sungai Selatan, Jakarta Timur, Kudus, Lamongan, Lombok Timur, Lumajang, Maluku Tenggara Barat/Maluku Barat Daya, Ogan Ilir, Probolinggo, Purwakarta, Semarang, Tapanuli Tengah, Tebo, Toli and Tuban.

10. The certification dataset has a high share of teachers with a bachelor's degree. This is likely due to the sampling procedure applied for this study. It gave schools that had many teachers eligible for the certification program (those with a bachelor's degree) a higher chance of inclusion. Furthermore, the data for the certification study were collected more recently, and Indonesia's teacher certification program had a huge impact on motivating teachers to obtain their bachelor's degrees through academic upgrading.
11. The evaluations variable is further transformed by adding 1 and calculating logs before they are included in the regressions.
12. Note that it is easy to think of types of measurement error that do not average over students within schools. For example, whether surveyors come to school at the beginning or toward the end of a school day is likely to matter for the performance of the entire class.
13. When including district dummies in this regression, "overfitting" the variation in test score gains across schools is a potential problem. This problem is bigger for the certification data as the school-to-district ratio is lower. To address this concern, the Leave-One-Out R^2 is calculated. This measure is much less subject to overfitting problems than the standard R^2 or even the adjusted R^2. Indeed, the Leave-One-Out R^2 indicates that overfitting the outcome variable with the district dummies is mainly an issue with the certification data. The real gain in fit due to the district effects is 15% for the certification data and about 10% for the school committee data.
14. Note that when the district dummies are added to the model presented in column 5, they still explain roughly 15% of the variation.

REFERENCES

Alatas, V. and D. Filmer (2004), 'Local level governance and schooling in decentralizing Indonesia', unpublished, available at http://www1.worldbank.org/publicsector/decentralization/Feb2004Course/Background%20materials/Filmer.doc.

Bjork, C. (2004), 'Decentralisation in education, institutional culture and teacher autonomy in Indonesia', *International Review of Education*, **50** (July), 245–262.

Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan and F.H. Rogers (2006), 'Missing in action: teacher and health worker absence in developing countries', *Journal of Economic Perspectives*, **20** (1), 91–116.

de Ree, J., K. Muralidharan, M. Pradhan and H. Rogers (2013), 'Double for what? The impact of unconditional teacher salary increases on performance', unpublished, Washington, DC: World Bank.

del Granado, F.J., J. Martinez-Vazquez and R.M. McNab (2005), 'Fiscal decentralization and the composition of public expenditures', Working Paper 0501, Atlanta: Georgia State University, Andrew Young School of Policy Studies.

Glewwe, P.W., E.A. Hanushek, S.D. Humpage and R. Ravina (2011), 'School resources and educational outcomes in developing countries: a review of the literature from 1990 to 2010', NBER Working Paper No. 17554, Cambridge, MA: National Bureau of Economic Research.

Harris, D.N. and T.R. Sass (2011), 'Teacher training, teacher quality and student achievement', *Journal of Public Economics*, **95** (7), 798–812.

Jacob, B.A. and L. Lefgren (2004), 'The impact of teacher training on student

achievement: quasi-experimental evidence from school reform efforts in Chicago', *Journal of Human Resources*, **39** (1), 50–79.
Kaiser, K., D. Pattinasarany and G.G. Schulze (2006), 'Decentralization, governance and public services in Indonesia', in P. Smoke, E.J. Gómez and G.E. Peterson (eds), *Decentralization in Asia and Latin America: Towards a Comparative Interdisciplinary Perspective*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Lewis, B.D. and D. Pattinasarany (2009), 'Determining citizen satisfaction with local public education in Indonesia: the significance of actual service quality and governance conditions', *Growth and Change*, **40** (1), 85–115.
Listiawati, N. (2003), 'Teacher appraisal in Indonesia', *Journal of Southeast Asia Education*, **4** (2), 45–57.
Mullis, I.V.S., M.O. Martin, P. Foy and A. Arora (2011), 'TIMSS 2011 international results in mathematics', TIMSS and PIRLS International Study Center, Boston College, available at http://timss.bc.edu/timss2011/international-results-mathematics.html.
Pradhan, M., D. Suryadarma, A. Beatty, M. Wong, A. Gaduh, A. Alisjahbana and R.P. Artha (2014), 'Improving educational quality through enhancing community participation: results from a randomized field experiment in Indonesia', *American Economic Journal: Applied Economics*, **6** (2), 105–126.
Rajkumar, A.S. and V. Swaroop (2008), 'Public spending and outcomes: does governance matter?', *Journal of Development Economics*, **86** (1), 96–111.
Rivkin, S., E. Hanushek and J. Kain (2005), 'Teachers, schools and academic achievement', *Econometrica*, **73** (2), 417–458.
Suryadarma, D. (2012), 'How corruption diminishes the effectiveness of public spending on education in Indonesia', *Bulletin of Indonesian Economic Studies*, **48** (1), 85–100.
Toyamah, N., B. Sulaksono, M. Rosfadhila, S. Devina, S. Arif, S.A. Hutagalung, E. Pakpahan and A. Yusrina (2010), 'Teacher absenteeism and remote area allowance baseline survey', SMERU Research Institute, available at http://www.smeru.or.id/report/research/absenteeism/absenteeism_eng.pdf.
Turner, M., A. Imbaruddin and W. Sutiyono (2009), 'Human resource management: the forgotten dimension of decentralisation in Indonesia', *Bulletin of Indonesian Economic Studies*, **45** (2), 231–249.
United Nations Educational, Scientific and Cultural Organization (UNESCO) (2012), *Youth and Skills: Putting Education to Work*, Paris: UNESCO.
Universitas Cenderawasih, Universitas Negeri Papua, Social Monitoring and Early Response Unit, Independent Research Institute of Indonesia, Badan Pusat Statistik and United Nations Children's Fund (2012), 'We like being taught: a study on teacher absenteeism in Papua and West Papua', available at http://www.unicef.org/indonesia/Teacher_Absenteeism_Study_Papua_ENGLISH.pdf.
Widyanti, W. and A. Suryahadi (2008), *The State of Local Governance and Public Services in the Decentralized Indonesia in 2006: Findings from the Governance and Decentralization Survey 2 (GDS2)*, Jakarta: SMERU Research Institute.
World Bank (2004), *World Development Report 2004: Making Services Work for Poor* People, New York: Oxford University Press.
World Bank (2010a), 'Inside Indonesia's mathematics classrooms: a TIMMS video study of teaching practices and student achievement', available at http://documents.worldbank.org/curated/en/2010/01/18127789/inside-indone

sias-mathematics-classrooms-timss-video-study-teaching-practices-student-achievement.

World Bank (2010b), 'Transforming Indonesia's teaching force', Vol. 2 *of Pre-service Training to Retirement: Producing and Maintaining a High-quality, Efficient and Motivated Workforce*, Washington, DC: World Bank.

World Bank (2013a), *Spending More or Spending Better: Improving Education Financing in Indonesia*, Washington DC: World Bank.

World Bank (2013b), *Teacher Reform in Indonesia: The Role of Politics and Evidence in Policy Making*, conference edition, Washington, DC: World Bank.