

Working Paper
in Economics and
Development Studies



Department of Economics
Padjadjaran University

No. 201107

**Teacher Certification in Indonesia:
A Confusion of Means and Ends**

Mohamad Fahmi,
Achmad Maulana, and
Arief Anshory Yusuf

Center for Economics and Development
Studies (CEDS)

Padjadjaran University

July, 2011

Center for Economics and Development Studies,
Department of Economics, Padjadjaran University
Jalan Cimandiri no. 6, Bandung, Indonesia.

Phone/Fax: +62-22-4204510

<http://www.ceds.fe.unpad.ac.id>

For more titles on this series, visit:

<http://econpapers.repec.org/paper/unpwpaper/>

Teacher Certification in Indonesia: A Confusion of Means and Ends¹

Mohamad Fahmi, Achmad Maulana, and Arief Anshory Yusuf

Center for Economics and Development Studies (CEDS)
Padjadjaran University

October, 2011

Abstract

In 2006, Indonesia started implementing a nation-wide program of teacher certification with the aim to certify as many as 2.3 million teachers by 2015 with the budgetary cost of as much as US\$5,600 million. Using data from a teacher survey we applied two different impact evaluation techniques namely Propensity Score Matching (PSM) and Difference-in-Difference (DD) to evaluate the impact of certification. These techniques can be used to estimate the difference in student's performance (in this case national exam score) attributed to the certification. Both methods suggest that teacher certification has no impact on student's achievement. The certification program may have improved teacher's living standard as remuneration increase is an elemental part of it, yet its formally-stated goal to improve the quality of education as should be indicated in better students' performance may not have been achieved. This program, being the largest in the nation's history, may have confused means and ends. We propose some policy recommendations. Two of them are: first, the government should implement a reward and punishment scheme to motivate teachers to continuously perform well; second, the government should introduced a teacher performance indicator are as close as possible to student's performance as key evaluation criteria and the reward-punishment scheme must be based on these criteria.

Keywords: teacher certification, propensity score matching, impact evaluation, Indonesia

1. INTRODUCTION

Teachers have an important role in pupil academic achievement. Studies in different countries find that qualified teachers are a major determinant of student achievement (Darling-Hammond, 2000, OECD, 2001). OECD study (2001), for example, concluded that the ability of education and training systems to respond to growing expectation from the society for a better education for their people depends on whether teachers have the ability to deliver the educational content in ways that meet this growing expectation. It is quite common to find that the focus of educational policy makers is to increase teachers' quality. This will ensure that teachers' qualification is adequate while

¹ This study is part of a Global Development Network (GDN) and Center for Economics and Development Studies (CEDS) project on Strengthening Institutions to Improve Public Sector Accountability. Usual disclaimer applies.

at the same time improving the teachers' salaries and working conditions. This in turn will attract best people into the profession. Teacher certification is an attempt to reach these ends.

Many studies have tried to shed light on the issue whether certification program affects student achievement (Hanushek et. al. 1999; Goldhaber and Brewer 2000; Harris and Sass 2007; Darling-Hammond 2001; Jepsen and Rivkin 2002). The general findings of those literatures, however, are mixed. Moreover, studies that attempt to look at the impact of teacher certification in developing countries on student's performance hardly exist in the literature.

In Indonesia, a nation-wide program of teacher certification was started in 2006 with a target of certifying around 2.3 million elementary and secondary teachers in 2015. With this large-scaled certification program, all teachers in Indonesia will eventually be certified by 2015. The budgetary cost of this program is estimated to be about US\$460 million. To the best of our knowledge, with this program's magnitude, this could be the biggest teacher certification program in the developing world, if not in the world.

The objective of this study is to analyze the impact of the teacher certification program on students' achievement. To this end, we carried out a survey of both certified and non-certified elementary school teachers, recorded the national-standard exam score of their pupils, as well as the teacher's relevant characteristics. Considering that the teacher's likelihood of being certified is endogenously determined by their characteristics, such as their qualification. Which will make a simple mean comparison of student's exam score biased, therefore we employed the propensity matching score to minimize such bias. We also used the Difference-in-Difference method, another alternative of impact evaluation technique commonly used in the literature. Using both methods, we found no evidence that the teacher certification has an impact on student's performance, as measured by national standard students exam score.

The certification, as formally stated in the law that governs it, has the objective to increase the quality of education. One elemental part of the program is improving the remuneration of certified teachers as an incentive. However, as our finding suggests no impact of the certification on student's performance, it may confirm some concerns that the certification's objective is not oriented to teacher's performance, but more to their living standard, as reflected by their student's achievement that does not make any progress. This large-scaled and expensive certification program may have confused ends and means.

The paper is organized into six sections. A summary of the motivations is highlighted in the introduction section. In Section 2, we describe in greater length the teacher certification program in Indonesia. Section 3 summarizes previous literature on the effectiveness of teacher certification. The methodology of the study including the data collection and analytical method will be discussed in Section 4. Section 5 discusses the findings, followed by concluding remarks in Section 6.

2. TEACHER CERTIFICATION PROGRAM IN INDONESIA

Teacher certification program in Indonesia was mandated by the Law Number 14, year 2005 on Teachers and Lecturers (or the so called “Teachers Law”). The law is an effort by Indonesian government to improve the quality of education in Indonesia. On the other hand, there has been a changing concern from accessibility to quality issue in the educational sector in developing countries. The objective of the Teacher Law is to create good quality national teachers as they should have good competencies in pedagogy, teaching professionalism, personal character and social issues.

The teacher certification program is not the first attempt to improve quality of teachers and the overall quality of education sector (MONE, 2009). During the period 1951-1960, Indonesian government had attempted to eliminate the high illiteracy rate by implementing some crash teacher training programs. After 1960, the teacher training school was transformed to *Sekolah Pendidikan Guru* (SPG or School of Teacher Education). The main objective of SPG is to prepare primary school teachers as huge number of junior secondary school graduates enrolled to SPG and create a surplus in primary school teachers. Yet, beginning from 1989, teacher recruitment became less selective as there was an excessive shortage of primary and secondary school teachers. Under the Education Law of 1989, the basic level of teacher education was increased from secondary education to higher education level. In 1950, the government established teacher training institutions (*Lembaga Pendidikan Tenaga Kependidikan* or LPTK) to improve teacher qualification to higher education level. Now, LPTK has been transformed to university (for example: Yogyakarta State University in Yogyakarta and Universitas Pendidikan Indonesia in Bandung).

Learning from the past experiences, Indonesian government designs a teacher certification program to improve all aspects of teacher quality including competency, academic qualification, certification, welfare, and status and reward systems for teachers. The government believes that this program is the most comprehensive strategy for teacher quality improvement (MONE, 2009).

2.1. Teacher Certification Process Mandated by the Teacher Law²

Teacher certification program, mandated by the Teacher Law, is one of the programs that the government of Indonesia (GOI) has implemented to reform national education system. With it, GOI expects to boost teacher competencies, pedagogy, personality, social and professionalism.

Basically, there are two types of teachers in Indonesia: in-service and pre-service teachers. The process for the former to get the certificate is relatively more convoluted than the latter. In this section we will only describe the process for the in-service teachers, since the government stated in 2005 that all in-service teachers should have teacher certificate within 10 years period.

²The main reference for this section is UU. No. 14/2005 on Teachers and Lectures, PP No.74/2008 on Teachers, and Kepmendiknas No. 16/2007 on Standard of Academic Qualification and Teacher Competence.

A teacher in Indonesia is classified as in-service if they meet one of the following criteria (i). S1 or D-IV graduate; (ii). Having teaching experience;; (iii). Having accumulated professional credits equivalent to grade IV-a; and, (IV) Acting as a supervisor (pengawas satuan pendidikan) in the current application. These in-service teachers need to write a portfolio which later must be submitted to *Dinas Pendidikan* (local technical agency) who will submit the dossier to LPTK. In LPTK, two evaluators are selected to review the teacher's portfolio. If the evaluators agree that the minimum standard has been met, then the LPTK grants the teacher the certificate. On the other hand, if they think the teacher has met the minimum standard but has some documents to complete then LPTK will ask the teacher to complete all the requirements. If the teacher has not met the minimum standard then LPTK offers two options either (1) teacher can enroll in Portfolio and Education training for Educational profession (PLPG) or (2) they have to revise their portfolio to be submitted later for next evaluation. After submitting the revised version, if the evaluators from LPTK still think the teacher's achievement is below the standard then the teacher has to enroll in the PLPG program.

Upon the completion of the PLPG program, teachers will be evaluated by means of the competence test. If they pass the test then they will be certified. If they fail, then they are allowed to sit for the re-take competence test. Once they pass, they will get certified. However, if they do not pass the test, they will be transferred to the local education technical office for further training.

In practice, however, the procedure of teacher certification has been far from perfect. According to Hastuti et. al. (2009), who gathered teacher certification data from six regencies/municipalities (*kabupaten/kota*) in Indonesia, the implementation of teacher certification had several weaknesses. Horizontal coordination between institutions, varying degree of socialization of the program, informational discrepancies are some of the weaknesses that they had identified.

Increased remuneration for certified teachers is an important element in the program. This, particularly, has been warmly welcome by many elements of Indonesian society as being a teacher has been commonly considered analogous to low-paid profession. However, actually there are four types of payment in teacher certification program: (1) remuneration or cost of professional allowances; (2) cost of pre-certification; (3) cost of certification process; and, (4) cost of upgrading after certification. The largest cost will be the professional allowance or about 91 percent of total certification related cost. The detail cost structures are provided in Table 1, Table 2 and Table 3.

Table 1. Sharing the Costs Associated with Certification

Costs associated with Certification	Who bears the cost?	
	Government	Teachers
Remuneration		
- Professional allowance	Yes	No
Certification Process		
- Portfolio review	Yes	No
- Remedial training for teachers who fail the certification process	Yes	No
- Re-undertaking the certification process for teachers who fail	Yes	No
- Pre-certification induction for new teachers (one semester for ECD, and primary; two semesters for JSS, SSS)	Yes	Yes
- Administrative costs of running certification	Yes	No
Upgrading		
- Upgrading through distance learning	Yes	Yes
- Upgrading through university courses	Yes	Yes
- Recognition of Prior learning (process)	Yes	No
- Upgrading through KKG-MGMP (Kelompok Kerja Guru - Musyawarah Guru Mata Pelajaran) (some Upgrading credits to be available)	Yes	Yes
- Opportunity cost for undergoing the Upgrading process	No	Yes

Table 2. Estimated cost of the Teacher Certification Process by Year

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014
Quota of teachers (000)	20.0	180.5	200,0	346.5	396.5	396.5	396.5	258.9	111.5
Cost (billion of Rp in 2006 prices)	40.0	360.9	400.0	693.0	793.0	793.0	793.0	516.1	223.0
Cost (millions of US\$)	4	36	40	69	79	79	79	52	22

Table 3. Comparison of Costs Associated with Certification

Item	2007	2008	2009	2010
Professional allowance (remuneration)	158,742	3,608,100	8,649,720	16,134,120
Assessment & certification	360,900	400,000	693,000	793,008
In-service upgrading	1,323,300	1,466,667	2,541,000	2,907,696
Total real cost (2006 prices)	1,842,942	5,474,767	11,883,720	19,834,824
Professional allowance as % of total	9%	66%	73%	81%
Certification as % of total	20%	7%	6%	4%
Upgrade as % of total	72%	27%	21%	15%

3. A BRIEF REVIEW OF LITERATURE

The general findings of previous literatures on the impact of teacher certification on student's performance are mixed. A study on matched panels of students and teachers in the US found that significant relationship between teacher salaries and pupil achievements hold only for experienced teachers but not for the new hires (Hanushek et. al. (1999)). Moreover, the certification test initiated to ensure high quality teaching is not significant in explaining student achievement.

Nevertheless, there is a significant impact of teacher certification on Mathematics and Reading scores, even though the positive and statistically significant effect of the certification status hold only for student achievement in Mathematics subject (Goldhaber and Brewer 2000).

Another study by Darling-Hammond et.al. (2001), evaluates whether certified teachers are more effective than those who have not met the requirements for certification. In addition, they also evaluate whether Teach for America (TFA) candidates are as effective as experienced certified teachers. Reviewing 4th and 5th grade student achievements scores on six different reading and mathematics tests over 1995 to 2002; they conclude that certified teachers produce significantly stronger student achievement gains than uncertified teachers. The same findings were also found when the certified teacher is compared with TFA recruits and teachers with non-education diploma. Even after controlling for teaching experience, degrees and student characteristics, uncertified TFA recruits are still less effective than certified teachers. These finding is congruent with Darling-Hammond (2000) study, which concludes that teachers preparation and certifications have the strongest correlation with student achievement in reading and mathematics (Darling-Hammond, 2000).

However, one must be careful on the size of the effect of teacher certification. According to Jepsen and Rivkin (2002), teacher certification only account for small effect on student achievement with the model that has taken into account nonlinearity in return to experience. In this case, they used multi-period data that combines student demographic and test performance, and class size as well as teacher certification information.

4. RESEARCH METHODS

4.1 Estimating the Impact of Teacher Certification

4.1.1 Propensity Score Matching

A simple measure of estimating the impact of teacher certification on student's achievement, such as the exam score, is by comparing the mean of the exam score of students taught by two different groups of randomly selected teachers: the certified and non-certified teachers. However, this 'naive' comparison will be biased when we know that the likelihood of one teacher to belong to the certified group is not a random process. Table 4 will help illustrate the problem.

Table 4. Potential Bias in Simple Mean Comparison

	Before certification (ex-ante)	After certification (ex-post)
Certified teachers	A	C
Non-certified teachers	B	D

Suppose we randomly survey certified teachers and non-certified teachers after the certification (ex-post). We then calculate the mean of their student's score, C for the

certified teachers, and D for the non-certified teachers. We then conclude that the impact of the certification is simply the difference between the two, namely, C – D.

However, it is not difficult to see that better estimate of the impact of certification is in fact C – A, where A is the mean of the student’s score of certified teachers before they are certified (ex-ante). C – D and C – A will be different when the student’s exam score of the certified and non-certified teachers are already significantly different even before certification takes place. In fact, the process of the certification in Indonesia will be in such a way that teachers with better qualification, hence better student’s exam score, is more likely to be certified. C – D then is not only capturing the impact of certification, but capturing other characteristics or qualification unrelated to the certification. In many cases, however, researchers find the ex-ante situation, in this case A and B, is unobservable, whereas C – A is actually a counter-factual measure.

Being aware of this problem, Rosenbaum and Rubin (1983) introduce the Propensity Score Matching (PSM) method to tackle the problem. In this case, each teacher in the sample, belonging to both certified and non-certified groups, will be assigned a score (it is called propensity score) that measures the likelihood or probability of being certified. There can be cases that some teachers have similar likelihood of being certified even though they belong to different groups, i.e., certified and non-certified groups. By comparing the student’s exam score of only a subset of teachers in both certified and non-certified groups that have similar likelihood of being certified, we can eliminate other factors, such as qualification and other characteristics that may explain their difference in the exam score, other than certification. The way how to find this subset of teachers is called matching. This is how the name of Propensity Score Matching is originated.

The likelihood of being certificated is estimated using a logistic regression, where the probability is a function of teacher’s characteristics including qualification. We use the principle of parsimony with regard to the evaluation criteria formally adopted in the teacher certification process to consider variables to be included in the model. More formally the model can be written as follows:

$$Pr(CER_i = 1) = G \left(\beta_0 + \beta_1 EDU_i + \beta_2 EXP_i + \sum_m \gamma_m IND_{mi} + \sum_n \delta_n SCH_{ni} \right)$$

Where:

CER_i is whether teacher i is certified, 1 if certified and 0 otherwise.

EDU_i is the years of education of teacher i .

EXP_i is the years of teaching experience of teacher i .

IND_{mi} is others individual characteristics of teacher i that may constitute the portfolio evaluated for the certification process.

SCH_{ni} is school characteristics where teacher i works that may affect the likelihood of teacher i being certified.

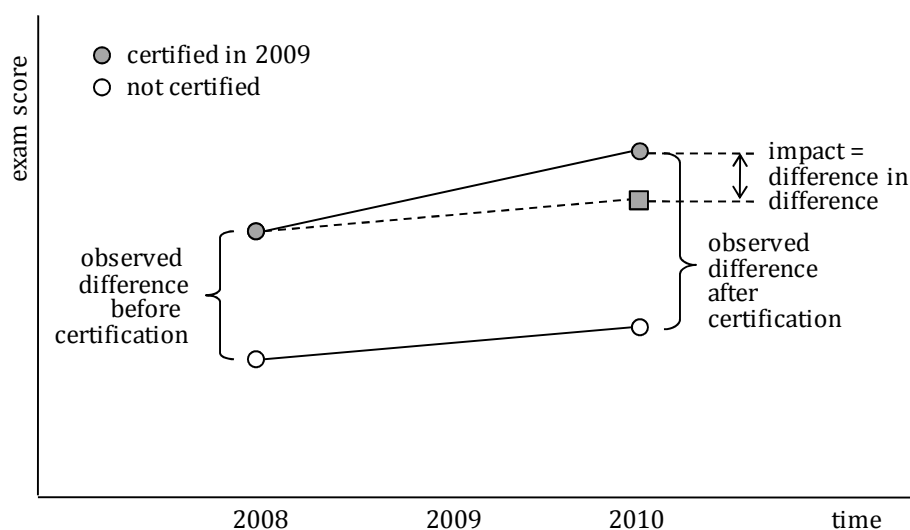
$G(Z)$ is a logistic function of $G(Z) = \exp(Z)/[1 + \exp(Z)]$.

After we estimated the parameter of the *logit* model, we then utilize the model to calculate teacher’s propensity scores as the predicted probability estimated with the model. By doing this, regardless their program status, each teacher will be assigned an estimated predicted probability to receives teacher certification. This predicted probability is called propensity score. By finding teachers from both certified and non-certified groups that have similar score or similar predicted probability of being certified and comparing their student’s exam score, we can conclude that such difference in the scores, if any, is only attributed to the certification, not other factors or characteristics. We apply the matching by pairing the propensity score between the certified and non-certified groups using various matching algorithms.

4.1.2 Difference-in-Difference

As the alternative to the PSM method, we used another method to estimate the impact of program or intervention, i.e., the Difference-in-Difference method. Figure 1 below may illustrate how this method works in estimating the impact of the teacher certification program.

Figure 1. Illustration of Difference-in-Difference Method



From a teacher survey data conducted in 2010 we managed to collect the information of the 2008 student’s exam score and the 2010 student’s exam score. We could also identify who among the teachers surveyed were certified in 2009. As illustrated in Figure 1, we can see that the student’s exam score of both certified and non-certified teachers has improved. However, the rate of the improvement is higher for the teachers who were certified in 2009. The difference in the rate of improvement can be interpreted as the impact of the certification in 2009. This impact is called Difference-in-Difference. It should be noted that the Difference-in-Difference method as illustrated above will truly reflect the impact of the certification relying on the assumption that had the teachers who were certified in 2009 not been certified, the rate of the improvement in their exam score would have been the same as those who were never certified.

Technically speaking, the Difference-in-Difference of the teacher certification program will be estimated using the following equation:

$$SCORE_{it} = \beta_0 + \beta_1 CER_i + \beta_2 P_t + \beta_3 CER_i \cdot P_t + \epsilon_{it}$$

where:

$SCORE_{it}$ is the student's exam score of teacher i in period t (2008 and 2010)

CER_i is whether teacher i was certified in 2009 (1 if certified and 0 otherwise)

P_t is the time period, 1 if 2010 and 0 if 2008.

ϵ_{it} is the error term.

The estimated parameter β_3 is the estimated impact of the certification, or the Difference-in-Difference.

4.2 Data Collection

To estimate the impact of teacher certification we conducted a teacher survey with the aim to collect information that includes teacher characteristics, professional affiliation, innovation in teaching, and most importantly their students average national exam scores on two subjects: Indonesian Language (*Bahasa*) and Mathematics. We conducted a survey to two groups of teachers: teachers who have been certified and those who have not yet been certified.

Given cost consideration, we conducted the teacher surveys in two regencies of Greater Bandung comprising Bandung Municipality, Cimahi Municipality, Bandung Regency, and West Bandung Regency. In designing the sampling, first we collected the teacher individual data from education agency of local government. We gathered a complete list of teachers who have been already certified and those who have not. The list was then used to randomly select teachers from both groups. Out of four regencies in Greater Bandung areas, only two handed the teacher list to our team. Therefore, we decided to limit our samples to two regencies; City of Cimahi and West Bandung Regency. Since City of Cimahi consists only of four districts, we decided to census all the schools in Cimahi. In the case of West Bandung Regency we selected five rural districts as urban areas have been represented by Cimahi Municipality

We purposively choose teachers from both certified and non-certified groups based on the following conditions: (1) the teachers must teach final year student in 2009 or earlier so that we can collect their national exam score; (2) they have to be the class primary teacher not a sport or art teacher which mean they are responsible to teach Math and Language, the subject we will use to see student's performance. It should be noted that for the certified group, we only include teachers whose application for the certificate had been approved prior to 2010 to make sure that the time is adequate to see the impact, if any.

For the student's performance we use the nationally-standard exam score averaged over students whom the teacher is responsible to teach. For elementary school, the subject is Indonesian Language (*Bahasa*) and Mathematics. These exams are standardized nationally so we could use it as means of comparison between teachers in different groups and areas.

The survey took three weeks, from the first to the third week of July 2010. In total, we have 202 teachers as the treatment group and 97 teachers as the control. The questionnaires contained questions about teachers' individual characteristics, the detailed cost for applying the teacher certification if they are already certified, teaching activities, their participation in training and organization outside schools, list of awards, and their current school characteristics.

Table 5. Summary Statistics of Teachers in the Sample

Variables	Obs.	mean	s.d.	Min.	Max.
<i>Already certified in 2009</i>	290	0.321	0.468	0	1
<i>Education and experience</i>					
Years of education	294	15.480	1.021	12	18
Teaching experience (years)	294	24.014	6.681	2	37
<i>Teachers portfolio</i>					
Training experience (dummy)					
on school management	294	0.323	0.468	0	1
on teaching	294	0.867	0.340	0	1
on specific subjects	294	0.820	0.385	0	1
Active in social organizations (dummy)	294	0.493	0.501	0	1
Ever received awards in teaching (dummy)	294	0.180	0.385	0	1
<i>Other characteristics</i>					
Gender (female = 1, 0 otherwise)	294	0.571	0.496	0	1
School size (number of classrooms)	294	9.500	7.614	2	46
School area (urban=1, 0=rural)	294	0.531	0.500	0	1

Source: Teacher survey

5. RESULTS AND DISCUSSIONS

Results from Propensity Score Matching

Table 6 shows the result of the logistic model estimation. As a reminder, the dependent variable of this model is the probability of teachers already certified in 2009. The result suggests that length of education and teaching experience are the strongest predictors of certification. Every additional one year of education increases the probability of teacher being certified by 0.167, whereas every additional one year of teaching experience increases the probability of being certified by 0.038. The effects are strongly significant at 1% level. The effect of education is a lot stronger than that of experience (more than 4 times).

Another significant variable is school size. This may reflect other variables reflecting school quality that has an impact on teachers being certified. Bigger school is normally better than smaller school as it reflects the school's ability to attract students.

Another interesting finding is that teachers' portfolios other than education and experience are not good predictors of certification. These portfolios are formally factors to be considered in the certification process. Variables, such as training experience, activity in social or professional organization and awards in teaching are not statistically significant. This is in strong contrast with years of education and teaching experience in which their influence on the probability of certification are quite large and strongly significant.

Table 6. Logistic Model of Certification

Variables	Coef- ficient	standard error	Marginal effects	standard error
<i>Education and experience</i>				
Years of education	0.975	0.222***	0.167	0.034***
Teaching experience (years)	0.220	0.043***	0.038	0.006***
<i>Teachers portfolio</i>				
Training experience (dummy)				
on school management	0.256	0.316	0.045	0.057
on teaching	0.377	0.458	0.060	0.067
on specific subjects	0.028	0.404	0.005	0.069
Active in social organizations (dummy)	0.043	0.306	0.007	0.053
Ever received awards in teaching (dummy)	0.561	0.369	0.106	0.075
<i>Other characteristics</i>				
Gender (female = 1, 0 otherwise)	0.017	0.319	0.003	0.055
School size (number of classrooms)	0.037	0.020*	0.006	0.003*
School area (urban=1, 0=rural)	-0.331	0.343	-0.057	0.059
Constant	-22.395	3.966		
Likelihood ratio	77.79***			
Log likelihood	-143.05			
Pseudo R2	0.2138			
Number of observation	290			

Note: (***) is significant at 1%, (**) is significant at 5%, (*) is significant at 10%

Table 7 lists our estimate of the impact of teacher certification using the Propensity Matching Score method. We report the results using various different matching algorithms in order to check for robustness.

Simple mean comparison (or unmatched comparison) suggests that the difference in the student's exam score between certified and non-certified teachers is negligible. Therefore, if the student's national exam score represents the teacher quality, it may imply that being certified does not necessarily mean they belong to teachers with better qualification. If there is such difference, they are not statistically significant. This is in contrast with all the procedure of the certification which clearly state that qualifications are important consideration in the decision to certify teachers.

As can be seen from Table 7, there seems to be only a slight or negligible difference in the exam score between certified and non-certified teachers. Without the matching, the exam score, for both Indonesian Language (Bahasa) and Mathematics, is slightly higher, yet statistically insignificant, for certified teachers. However, as expected, the matching eliminates those differences. Although, small in magnitude, the propensity matching score may work in removing the bias due to endogeneity in certification.

The results suggest that no-impact of certification is quite robust to various different matching algorithm. All matching algorithm produces very low t-statistics, suggesting no-difference attributed to the certification. Moreover, with the exception of the radius matching, all 4 matching algorithm attenuated the difference in the exam score between certified teachers and the non-certified.

Table 7. The Matched and Unmatched Difference in The Student Exam Score in 2010 by Various Matching Algorithm

<i>Math</i>	<i>Sample</i>	<i>Treated</i>	<i>Controls</i>	<i>Difference</i>	<i>S.E.</i>	<i>T-stat</i>
Nearest-neighbor	Unmatched	7.438	7.386	0.052	0.137	0.380
	ATT	7.449	7.571	-0.122	0.222	-0.550
Caliper	Unmatched	7.438	7.386	0.052	0.137	0.380
	ATT	7.454	7.623	-0.169	0.233	-0.720
Kernel	Unmatched	7.438	7.386	0.052	0.137	0.380
	ATT	7.449	7.446	0.003	0.169	0.020
Radius	Unmatched	7.438	7.386	0.052	0.137	0.380
	ATT	7.449	7.386	0.063	0.136	0.460
Ties	Unmatched	7.438	7.386	0.052	0.137	0.380
	ATT	7.449	7.565	-0.116	0.222	-0.520
<i>Bahasa</i>	<i>Sample</i>	<i>Treated</i>	<i>Controls</i>	<i>Difference</i>	<i>S.E.</i>	<i>T-stat</i>
Nearest-neighbor	Unmatched	6.917	6.854	0.063	0.078	0.810
	ATT	6.918	6.938	-0.020	0.137	-0.140
Caliper	Unmatched	6.917	6.854	0.063	0.078	0.810
	ATT	6.931	6.951	-0.019	0.142	-0.140
Kernel	Unmatched	6.917	6.854	0.063	0.078	0.810
	ATT	6.918	6.916	0.003	0.095	0.030
Radius	Unmatched	6.917	6.854	0.063	0.078	0.810
	ATT	6.918	6.854	0.064	0.074	0.860
Ties	Unmatched	6.917	6.854	0.063	0.078	0.810
	ATT	6.918	6.933	-0.015	0.137	-0.110
<i>Math & Bahasa</i>	<i>Sample</i>	<i>Treated</i>	<i>Controls</i>	<i>Difference</i>	<i>S.E.</i>	<i>T-stat</i>
Nearest-neighbor	Unmatched	7.177	7.127	0.050	0.100	0.510
	ATT	7.184	7.254	-0.071	0.161	-0.440
Caliper	Unmatched	7.177	7.127	0.050	0.100	0.510
	ATT	7.193	7.287	-0.094	0.168	-0.560
Kernel	Unmatched	7.177	7.127	0.050	0.100	0.510
	ATT	7.184	7.184	-0.001	0.123	0.000
Radius	Unmatched	7.177	7.127	0.050	0.100	0.510
	ATT	7.184	7.127	0.057	0.099	0.570
Ties	Unmatched	7.177	7.127	0.050	0.100	0.510
	ATT	7.184	7.249	-0.065	0.160	-0.410

Note: ATT stands for Average Treatment Effect of the Treated. It is the estimated difference due to treatment, in this case, due to certification. Source: Author's calculation

Results from Difference-in-Difference

The estimated model to calculate the impact of certification using the Difference-in-Difference method is shown in Table 8 below:

Table 8. Difference-in-Difference Estimates

	Math	Bahasa	Math & Bahasa
Constant	6.516*** (0.115)	6.828*** (0.077)	6.672*** (0.089)
Certified in 2009 (1 if yes, 0 otherwise)	0.088 (0.336)	-0.072 (0.225)	0.008 (0.261)
Period (1 if 2010, 0 if 2008)	0.871*** (0.146)	0.027 (0.098)	0.456*** (0.114)
Certified × Period	-0.063 (0.428)	0.117 (0.286)	0.020 (0.333)

Note: ***) is significant at 1%, **) is significant at 5%, *) is significant at 10%
Number in parentheses is standard error.

From Table 8, we can see that for the non-certified teacher, their student's score in Math and Bahasa in 2008 is around 6.672 in average. Between the periods of 2008 to 2010 there was an increase of 0.871 point in the student's Math score and it is statistically significant at 1% level. The 0.027 increase in Bahasa score, however, is not statistically significant. We can also see from the coefficient of the certification, that there is no significant difference in the student's exam score of certified and not-certified teachers.

The impact of the certification can be found from the interaction variables (Certified × Period). From the estimated coefficient of this variable, we can conclude, for example, that the certification in 2009 has increased the student's score of Bahasa by 0.117 point. However, this is not statistically significant. In fact, the coefficients of the interaction variables in all three models are not statistically significant. We cannot conclude that the impact of the certification is statistically different from zero. Hence, our Difference-in-Difference method found similar conclusion as the PSM method that the teacher certification has no impact on student's performance as measured by the nationally-standard score of Math and Bahasa exams.

There are some possible explanations on why the teacher certification does not have the expected impact on student's performance. In general, it can be divided into two factors. First is the weakness in its design, and second is its obstacle in its implementation. On the design issue, if the certification needs to have impact on such objective indicator as national exam score, then this needs to be explicitly reflected in the incentive system. This does not happen to be the case. Student's performance, as measured by their national exam score is not part of the parameter to be evaluated regularly.

One may argue that certification may have impact on teacher's performance and eventually student's performance because certified teachers are given more financial

incentives. More financial incentive means more financial security and teachers do not need to find extra teaching jobs, so that they can more focus on their main teaching jobs. However, this is not generally true. Hastuti et. al. (2009), for example, in their study on impact of certification program in Indonesia, found that most of the respondents they interviewed in their studies believed that teacher certification program will not increase the teachers' quality, even though they were aware that the additional income for certified teachers may increase teachers' welfare and at the end, teachers could more focus on their task and have more preparation to increase their teaching technique. They believe that the important factor of teacher performance is more of the integrity, such as the commitment to do their best.

On the issue of implementation, Hastuti et. al. (2009) found at least two factors that contribute to the ineffectiveness of teacher certification program. First, the concerns that the selection is not designed to identify best teachers. In three provinces of their study--Jambi, West Jawa, and West Kalimantan—Hastuti et. al. (2009) find there is an indication of manipulation in teacher selection process. Second, the respondents knew that many of their colleagues had manipulated their portfolio documents. They believe that portfolio method in certification process is an incorrect method to determine a good teacher as it creates incentives for teacher to cheat. Furthermore, Hastuti et. al. (2009) argue that teacher certification process by portfolio method does not have any clear paradigm and will not increase the teachers' quality as it only assesses documents not the real performance of the teachers. They believe that intensive training and education program could be a better method to increase teacher ability than portfolio method. In short, the certification, due to its drawback in its implementation, did not really manage to pick 'oranges' from 'lemons'.

Overall, this study provides a finding of a quantitative analysis which suggests that teacher certification in Indonesia may have no impact on student's performance. The recent teacher certification program in Indonesia may well be useful in improving the living standard of teachers, but whether or not it can translate into teacher's performance and in turns the student's performance remains questionable.

6. CONCLUDING REMARKS AND POLICY RECOMMENDATIONS

Indonesia had just recently started a large-scaled teacher certification program with the target that all teachers will be certified by the year 2015. With around 2.3 million teachers involved as well as its associated high cost, this program is by far among the most ambitious government-supported certification program in developing countries. Nonetheless, there has never been any rigorous attempt to evaluate its impact on teacher's performance, especially their students' achievement. This paper intends to fill this gap. We conducted survey of teachers in greater Bandung area. This survey collected information on teachers' certification status, their individual characteristics and their pupils' achievement.

Considering that the teacher's likelihood of being certified is endogenously determined by the teacher's characteristics, such as their qualification which will make a simple mean comparison of student's exam score biased, we employed the propensity matching score to minimize such bias. We also use the Difference-in-Difference method

as alternative evaluation technique to check for the robustness of the analysis. The result supports some concerns that the teacher certification has no impact on student's performance.

The certification, as formally stated in the law that governs it, has the objective to increase the quality of education. One elemental part of the program is improving the remuneration of certified teachers as an incentive. In fact, the largest cost will be the professional allowance or about 91 percent of total certification related cost. However, as our finding suggests no impact of the certification on student's performance, it may support some concerns that the certification's objective is not oriented to teacher's performance but only to their living standard as reflected by their student's achievement.

The main problem with the current design of the teacher certification is that it has very limited characteristics of a performance-based incentive system. As it uses a portfolio assessment, some teachers can be certified and pay rise earlier but eventually (in the next two or three years) they will get certified. It is very hard to expect improvement in performance when you know that eventually everyone will get reward disregarding of your improvement.

Second, after some teachers are certified and get pay rise, there is hardly any system of penalties in place that may credibly threaten them of losing the pay rise when their performance is not better than the uncertified teachers. When we expect that the teacher certification should improve teacher's performance, such as reflected by their student's achievement, then the improvement in the system needs to work around these issues. Otherwise, there is no need to mention that this certification is aimed to improve teacher's performance. Its sole objective is just to increase teacher's welfare. But again, it can be such a waste of resource, given the nature and the size of the initiative.

Therefore, we need to create a better solution on how to improve this teacher certification program. Such improvement in the system can be developed by experts in greater detail but in any case, they need to have characteristics, at the fullest extent, of a performance-based system. Some elements of those characteristics, but not limited to, among others are: (a) it should reward better teachers (as reflected by student's performance, as final goal, or other efforts as intermediary goals) and penalize less-performing teachers using the same criteria; (b) it should reward teachers when their performance improved over time and penalize them when they perform consistently worse than before; (c) the emphasis of the performance-based system should be stated very explicitly and clearly in the rule of the game; and, (d) it should be credible.

The example of practical version of the amendment to the system can be as follows: (a) Stating and emphasizing very explicitly that the increase in remuneration can be cancelled when teachers do not perform a minimum standard of services and performance and show this as a credible rule. Minimum standard of services can be a minimum time to spend at school. This will regulate teacher's other side-jobs, such as teaching in other schools so they can concentrate more on preparing classes or even concentrate on giving more attention to the least performing students; (b) Using indicators that are as close as possible to student's performance as key evaluation criteria and the additional incentive system must be based on these criteria. For

example, teacher who can improve their student's national exam score will be rewarded financially as well as non-financially (awards is among the example). It should be noted that there is no need to just use solely national exam score, as it can only apply to certain subjects, for example, but also use other innovative evaluation indicators that can be tailored according to different needs; (c) Complementing the fixed amount remuneration (as already reflected in the current system) with variable financial incentives, based on performance. Other than national exam score, nationally standardized student's evaluation of teachers can also be attempted. This can monitor teacher's performance at least overtime. When they get consistently poorer and poorer evaluation from students over time then the teacher's should get warning and penalized. Another example of alternative basis for additional compensation is additional roles and responsibilities to be taken by teachers that are aimed to improve student's performance; and, (d) Eliminating some requirements of portfolio on professional development that are loosely associated with student's performance.

There could be longer list of rooms for improvement when all stakeholders and experts can think again and improve the certification programs. There could be even more options when we learn more about what other countries are doing in their attempt to improve the quality of education process and at the same time improving the living conditions of teachers. The problem with the current certification program in Indonesia is that despite its relevant and much needed role to improve teacher's welfare, its impact on the quality of education process is unclear. This is an urgent call for revisions in its design and better governance in its implementation.

REFERENCES

- Cameron, A. and Trivedi, P, K. (2005). *"Microeconometrics: Methods and Applications."* Cambridge University Press
- Darling-Hammond, L. (2000). "Teacher Quality and Student Achievement: A Review of State Policy Evidence". *Educational Policy Analysis Archives*, 8(1).
- Darling-Hammond, L, Berry, B and Thoreson, A. (2001). "Does Teacher Certification Matter? Evaluating the Evidence. *Educational Evaluation and Policy Analysis*, 23(1), pp.57-77
- Goldhaber, D, and Brewer, D. (2000). "Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement." *Educational Evaluation and Policy Analysis*, 22(2), pp.129-45
- Hanushek, E, and Rivkin, S. (2006). "Teacher Quality." *Chap 18 in Handbook of the Economics of Education*, Volume 2, Edited by Eric A Hanushek and Finish Welch, Elsevier BV.
- Hastuti, Sulaksono B, Akhmadi, Syukri, M, Sabainingrum,U and Ruhmaniyati. (2009). "Implementation of the 2007 Certification Program for Practicing Teachers: A Case Study of Jambi, West Java, and West Kalimantan Provinces." SMERU Research Report, June.

- Johnstone, J dan Jiyono. (1983). "Out-of-school Factors and Educational Achievement in Indonesia." *Comparative Education Reviws*, 27(2), pp. 278-95.
- Keputusan Menteri Pendidikan Nasional No. 16/2007 on Standard of Academic Qualification and Teacher Competence.
- Leuven, E, and Sianesi, B. (2003). "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing."
- Mohandas, R. (2000). "Report on the Third International Mathematics and Science Study (TIMSS)." National Institute for Research and Development of the Ministry of National Education, Jakarta. Mimeo.
- OECD. (2001). *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000*
- Peraturan Pemerintah (PP) No. 74/2008 on Teachers
- Rosenbaum, P and Rubin, D. (1983). "The central role of the propensity score in observational studies for causal effects." *Biometrika* vol 70(1), p41.
- Suryadarma, D. (2010). "The Quality of Education in Indonesia." *Presentation in Indonesia Update 2009 Conference, Canberra*.
- Suryadarma, D, Suryahadi, A, Sumarto, S and Halsey, R. (2004). "The Determinants of students performance in Indonesian public primary schools: the role of teachers, and schools." *SMERU Research Institute Working Paper*, December.
- Undang-Undang (UU) No. 14/ 2005 on Teachers and Lecturers