Journal of Southeast Asian Economies Vol. 39, No. S (2023), pp. S62–S79 *ISSN 2339-5206 electronic*

DOI: 10.1355/ae39-Se

The Impact of Sending Top College Graduates to Rural Primary Schools

Masyhur A. Hilmy

Teacher quality is crucial to delivering good education. However, improving teacher quality in developing countries can be a tough task. This paper investigates the impact of a teacher placement programme that sends college graduates with a strong academic track record to teach in rural primary schools in Indonesia on student test scores. Using a difference-in-difference approach, the study finds that exposure to programme teachers for a semester is associated with a 0.16 standard deviation increase in their students' average mathematics scores. The weakest students benefited more, with an increase in score by 0.20 standard deviation. Students receiving direct instructions from programme teachers during scheduled classroom periods benefited even more. Attracting better talents to teach in rural schools could be an important pathway to improving the academic achievements of the weakest students at rural schools.

Keywords: Education, alternative teacher placement, Indonesia.

1. Introduction

Teacher quality is crucial to delivering good education (Chetty, Friedman, and Rockoff 2014; Glewwe et al. 2013). However, rural schools often struggle to meet this promise (Chaudhury et al. 2006). Selection into teaching is a key issue—education majors in colleges and universities do not attract the brightest talents, and few of them relish the career prospect in rural schools. Teacher absenteeism is rampant. Even when the teachers are present, the students are still often left with teachers who do not master their lessons or do not know how to teach, or both (Bold et al. 2019). To address these problems, governments and NGOs invest significant resources in a variety of interventions, but much remains unknown about their effectiveness (Evans and Popova 2016).

This paper studies a programme that places college graduates with strong academic and leadership backgrounds to teach at schools in rural areas in Indonesia. In particular, the article examines the Indonesia Mengajar programme, which has placed hundreds of teachers in rural schools since 2010.[1] Indonesia

Masyhur A. Hilmy is a PhD candidate in Economics at Boston University, 270 Bay State Rd, Boston, MA, 02215 USA; email: mhilmy@bu.edu

Mengajar recruits graduates of top Indonesian universities, trains them for six to eight weeks, and then sends them as teachers to primary schools across seventeen districts. Very few (<10 per cent) Indonesia Mengajar recruits have studied education majors in college. Most of them majored in engineering, natural and social sciences, or literature and the humanities. This contrasts with the regular teacher force in the programme districts, of whom 90 per cent have an education major. Programme recruits are assigned to specific schools just before deployment, and they take their placements as given.

The Indonesia Mengajar programme shares characteristics with Teach for America (TFA) and similar schemes in other countries, although Indonesia Mengajar is not an official member of its network (Teach for All 2021). Each Indonesia Mengajar teacher is contracted to teach for a year in rural Indonesia, but the school can host a succession of programme teachers for up to five years. Headmasters in the programme schools assign the teachers to either teach students as homeroom teachers (who teach multiple subjects for a particular grade) or as subject teachers (who teach specific subjects such as mathematics across grades). Indonesia Mengajar teachers live near their assigned schools, and the students in a treatment school regularly interact with them. The initiative may improve student outcomes because it exposes students to teachers with stronger academic backgrounds and who are more consistently present.

This paper investigates the impact of the programme using a difference-in-difference strategy between treatment and comparison schools. The treatment schools are schools where Indonesia Mengajar placed their first cohort of teachers in 2010. The comparison schools are schools where Indonesia Mengajar placed subsequent cohorts and other never-treated schools located in close proximity to the treatment school (<3 km). The programme's impact is identified under the assumption that outcome trends would be similar in both treated and comparison schools in the absence of treatment. The study estimates the impact of the Indonesia Mengajar programme on the students' mathematics scores using the Ministry of Education's 2008–11 examination score database. Because the 2011 examination took place before the second cohort of Indonesia Mengajar teachers were deployed, students in comparison schools had not been exposed to the programme during the examination. This allows a comparison to be made to estimate the programme's impact. At the same time, students in the treatment schools had been exposed to Indonesia Mengajar teachers for half a year, which allows the resulting estimates to be interpreted as the programme's short-term effect. The Ministry's dataset records each school's minimum, average, and maximum mathematics scores. This allows us to investigate how the programme teachers may affect students with various ability levels.

The results of this study show that exposure to Indonesia Mengajar teachers is associated with higher average mathematics scores by 0.14 points at the 10 per cent statistical significance level, which is equivalent to a 0.16 standard deviation. Indonesia Mengajar teachers seem to be particularly more effective in teaching the weakest students, and they raise the minimum score by 0.20 points. Meanwhile, the estimated effect on the maximum examination score is positive, but lower than the effect on the average score and not statistically significantly different from zero.

These estimates align with the most recent randomized evaluation of Teach for America (TFA) in the US. Students of TFA teachers in grades 1 and 2 perform significantly better in mathematics by 0.16 standard deviations (Clark and Isenberg 2020). However, the TFA evaluation measured the impact on students after a longer exposure than the Indonesia Mengajar teachers in this study (i.e., a two-year tenure for TFA fellows versus a half-year exposure to Indonesia Mengajar teachers at data collection). Suppose students benefit from more exposure to teachers with stronger academic ability. In this case, the estimated short-term impact of the programme may understate the total learning gains that the students received from the entire duration of the programme.

Classroom instructions from Indonesia Mengajar teachers drive these effects. To separate the effect of direct instruction from other changes (e.g., increased supervision) from the subdistrict superintendents that the programme's high-visibility status may have brought to treatment schools, the study uses Indonesia

Mengajar organizational reports that record the teaching assignments for all of the first cohort teachers. It is found that the mathematics score was higher for students with scheduled classroom instructions from Indonesia Mengajar teachers: their mathematics classes are associated with 0.40 points higher scores.

The weakest students appear to benefit more from Indonesia Mengajar teachers' Indonesian and science classes than the mathematics classes. The estimated effects are 0.74 and 1.04 points for Indonesian and science classes, respectively. These results suggest that the students benefited both from the use of mathematics concepts in science lessons and from more intensive use of the national language. Although school examinations are written in the Indonesian language, most of the population speaks local languages at home. Thus, comprehension problems may underlie the students' poor mathematics scores, and instructions that improve comprehension can boost performance.

The analysis in this paper contributes to several strands of literature. First, it presents new evidence of a Teach for America-style programme from a developing country, where expanded schooling access in recent decades has typically led to universal enrolment with low learning levels. To the best of the author's knowledge, this study is the first evaluation of such a programme outside the US and the UK. Since TFA's inception in 1990 and its first expansion to the UK as Teach First (TF) in 2003, this scheme has spread globally under the Teach for All (TFAll) network with affiliated programmes currently operating in sixty countries, including India, Peru, Nigeria, and many others (Teach for All 2021). This figure excludes programmes that are not officially part of the TFAll network but share similarities, such as the Teach First Norway and the Indonesia Mengajar programme, which adds to the global influence of the TFA idea. Despite the rapid expansion, there is little empirical research on the impact of the TFAll programmes outside of the two original countries (Thomas, Crawford-Garrett, and Rauschenberger 2021; see, e.g., Clark and Isenberg 2020 and the references therein for TFA and Allen and Allnutt 2017 for TF).

This paper also adds evidence to the literature on interventions that send educated individuals to areas with a low level of learning. Two recent studies are related to this paper. Chen et al. (2020) evaluated the impact of the send-down movement in the 1960s People's Republic of China and found that exposure to educated urban youths affected by the mandate to resettle in the countryside increased rural children's educational achievement. In the Gambia, Eble et al. (2021) show that a bundled para-teacher intervention programme modelled from a similar scheme in India (Lakshminarayana et al. 2013) led to a dramatic improvement in children's literacy and numeracy test results. This literature suggests that an effective intervention at a low baseline setting could lead to large gains in educational achievements.

More broadly, this paper also connects to the literature on the personnel economics of the state. This literature connects governance in developing countries with the public employees who perform government functions (Finan, Olken, and Pande 2017). Frontline service providers (e.g., teachers and nurses) play an instrumental part in the development process. The setting of this paper exemplifies the impact that talented individuals with prosocial leanings can have when they provide a public good in remote areas (Ashraf et al. 2020).

The remainder of this paper is organized as follows. The next section describes the context of the programme implementation. This third section outlines the empirical strategy. The results are described in the subsequent section. The fifth section concludes.

2. Context: The Indonesia Mengajar Programme

2.1 Background and Recruitment Process

The Indonesia Mengajar programme (literal translation: Indonesia Teaches) sends top university graduates to teach for a year in rural elementary schools across Indonesia. To become a teacher with the programme, individuals apply through the official website during the recruitment period. Applicants must provide

academic background information, complete essay prompts, and supply references. These initial screening shortlists applicants based on academic strength. Shortlisted applicants are then invited to the interview rounds to participate in individual interviews, group discussions, and classroom simulations. The later-stage screening further selects prosocial motivations and behaviours. Depending on the cohort, Indonesia Mengajar admits between thirty-three and seventy-five individuals to participate in its pre-deployment training camp. With thousands of applicants per cohort, this translates to a highly selective admission rate of under 1 per cent (Gozali 2020).

Indonesia Mengajar regularly attracts college graduates from top Indonesian universities. A college degree is required by Indonesian law to teach in primary schools. However, in practice, 32 per cent of primary school teachers in the seventeen districts where the programme operated did not meet this standard (Table 1). Whereas more than 90 per cent of primary school teachers in these districts majored

TABLE 1
Comparison of Indonesia Mengajar Teachers and Other Teachers by Education Level and College Majors

	IM Cohort 1		*IM 2010–15*		*Other Teachers*	
	N	*%*	*N*	*%*	*N*	*%*
Highest education level						
High school or lower					10,274	21%
Associate degree					5,470	11%
Bachelor's degree or higher	51	100%	614	100%	32,323	67%
College majors (for holders of associate degree or higher)						
Education (primary school)			6	1%	23,251	62%
Education (other than primary school)			64	10%	10,787	29%
Engineering and Computer Science	12	24%	81	13%	44	<1%
Literature and Humanities	10	20%	68	11%	266	1%
Economics, Business, Management	4	8%	63	10%	146	<1%
Communications	2	4%	57	9%	13	<1%
Public Admin, Poli Sci, Intl Relations	6	12%	53	9%	66	<1%
Basic Sciences	3	6%	49	8%	343	1%
Psychology	6	12%	48	8%		
Forest, Marine, Agriculture	3	6%	35	6%	21	<1%
Medicine, Pharmacy, Health			34	6%	3	<1%
Law			19	3%	45	<1%
Architecture, Planning and Development	2	4%	17	3%		
Art and Design	1	2%	11	2%	36	<1%
Other	2	4%	9	1%	694	2%
N/A					2,078	5%
Total	51	100%	614	100%	37,793	100%

NOTES: IM refers to Indonesia Mengajar. "IM 2010–2015" data includes the first ten cohorts of teachers. Statistics for "Other teachers" came from a subsample of primary school teachers who took the 2015 teacher competency test dataset and was teaching in one of the 17 programme districts.
SOURCE: Author's calculations.

in education, Indonesia Mengajar teachers typically did not graduate from an education major. None of the teachers that made up its first cohort had an education degree. Among the teachers it recruited until 2015, one in ten held an education major. The majority of these teachers instead have degrees in various engineering and science fields or literature and the humanities. Meanwhile, the origin universities of Indonesia Mengajar recruits are highly placed in the national ranking, with the top ten universities contributing more than half of its total teachers (Table 2).

Primary school teachers who graduated from the same universities as Indonesia Mengajar teachers scored higher on the nationwide competency test that the Ministry of Education held in 2015 than teachers in districts where the Indonesia Mengajar programme operated. The Ministry's threshold for the pass rate was 55/100, and the national average score was 53. Across Indonesia Mengajar operational districts, teachers score 48.6 on average, lower than the passing threshold and the national average. In contrast, teachers who were educated in top universities, where 75 per cent of the Indonesia Mengajar teachers

TABLE 2
Indonesia Mengajar Teachers by Origin Universities

	IM Teachers		*University Rank*	
University Name	*Cohort 1*	*2010–15*	*Indonesia*	*World*
UI/Universitas Indonesia	13	86	1	694
UGM/Universitas Gadjah Mada	7	78	6	1,496
ITB/Institut Teknologi Bandung	14	62	2	896
UNPAD/Universitas Padjajaran	3	39	27	2,986
IPB/Institut Pertanian Bogor	3	32	13	1,972
UNDIP/Universitas Diponegoro	3	29	9	1,753
UNAIR/Universitas Airlangga	5	23	7	1,551
UNIBRAW/Universitas Brawijaya		21	3	1,178
ITS Surabaya	1	15	4	1,220
UNS/Universitas Sebelas Maret		12	10	1,913
UPI Bandung		11	15	2,178
UM/Universitas Negeri Malang		10	23	2,839
UMM/Univ. Muhammadiyah Malang		10	34	3,298
UNHAS/Universitas Hasanuddin	1	9	17	2,550
USU/Universitas Sumatera Utara		8	8	1,575
Universitas Paramadina	1	8	168	7,816
UNP/Universitas Negeri Padang		7	25	2,919
UNESA/Universitas Negeri Surabaya		7	40	3,494
UNY/Universitas Negeri Yogyakarta		7	22	2,772
Overseas		19		1,606
Other		121	75	4,659
Total IM teachers/average rank	51	614	25	2,290

NOTES: IM refers to Indonesia Mengajar. IM teachers 2010–15 tabulated cohorts 1–10. University rank data from Webometrics, July 2020 ranking. Ranking for "overseas" and "other'' categories are the mean of specific universities, rounded down to the nearest integer. See Table A1 for the full list of overseas and other universities.
SOURCE: Author's calculations.

TABLE 3
Average Score from 2015 Teacher Competency Test, by Origin University and Age

	Mean	*Std Dev*	*N*	*Mean*	*Std Dev*	*N*
	All Nationwide			*<30 Year Olds*		
UI	72.5	12.7	246	74.6	12	20
UGM	77.9	10.2	349	80.4	7.4	13
ITB	80.8	8.7	39	81.8	.	1
UNPAD	72.2	11.7	452	68.3	12.1	37
IPB	77.5	9.8	380	76.1	9.4	26
UNDIP	77.7	9.5	312	77.7	10.3	32
UNAIR	75.9	10.9	144	68.2	12.9	5
UNIBRAW	75.4	10.7	249	73.5	11.1	17
ITS	80.1	10.4	109	70.8	11.4	11
UNS	67.5	13.1	5,645	75	10.7	1,308
UPI	60.3	12.7	19,413	65.7	11.9	4,646
UM	68.9	12.7	3,368	74.2	11.1	983
UMM	64.1	12.1	1,273	67.1	11.9	166
UNHAS	61.9	12.4	180	59.6	15.1	12
USU	64.9	12.2	283	66.9	11.5	37
PARAMADINA	81.8		1			
UNP	58.3	12	8,282	63	12.4	2,057
UNESA	65.2	13.3	3,919	69.7	11.9	823
UNY	70.2	12.6	5,113	73.7	12.3	1,521
Overall	74.4	12.3	49,757	72.8	12	11,715
	All Programme Districts			*<30 Year Olds*		
All education levels	48.6	11.3	48,067	50.3	11.4	7,417
Any college	49.5	11.3	37,793	51.2	11.6	4,511
Bachelor's and up	50.4	11.3	32,323	51.6	11.6	4,085
Open University	51.1	11.3	13,916	54.8	11.1	911
Other universities	49.5	11.2	17,159	50.4	11.5	3,014

NOTES: Statistics from a subsample of all 1.3 million primary school teachers who took the 2015 teacher competency test and graduated from the 19 universities who contributed the most Indonesia Mengajar teachers. Teachers in this summary statistics are located in all 34 provinces. Of the 19 universities here, only UNS, UPI, UM, UNP, UNESA, and UNY are historical teacher colleges. The mean and standard deviation in the bottom row (overall) is an average of the origin university-level observation, weighted by the number of Indonesia Mengajar teachers it contributed to between 2010–15. The national test average was 53/100 and the passing grade was 55/100.
SOURCE: Author's calculations.

graduated from, performed better on the test with a weighted average score of 74.4 (Table 3), even though most did not graduate with an education major.[2]

The Indonesia Mengajar programme shares characteristics with Teach for All affiliate programmes in various countries. It attracts applicants with strong academic leadership backgrounds, runs a highly selective screening process, trains recruits without formal education degrees, and contracts them to teach in low-income schools for a short period. The programme was launched in 2010, a period of rapid expansion

for the Teach for All network (Thomas, Crawford-Garrett, and Rauschenberger 2021). Nevertheless, Indonesia Mengajar is not an official member of the Teach for All network (Teach for All 2021). Instead, recruitment materials and other organizational publications refer to a send-down programme that deployed college students from Java to teach high schools in the outer islands between 1951 and 1962 as its origin.[3]

2.2 School Selection

Between 2010 and 2015, Indonesia Mengajar sent teachers to seventeen districts across Indonesia (Figure 1). These districts agreed to receive Indonesia Mengajar teachers[4] as is typical of less-developed districts that routinely suffer a high rate of teacher absenteeism. These include border districts, areas nearby Java with poor performance, and other remote districts.[5]

To select the target schools within the district, Indonesia Mengajar looked for schools with demonstrable needs. These schools often lack (permanent) teachers due to their location in remote areas (e.g., in a small island or mountain range beyond the electricity grid and cell coverage). Within a district, the programme also considers the geographical spread. A local contact listed prospective schools that programme officers visited from Jakarta before finalizing the school selection. Every year Indonesia Mengajar sends teachers to four to ten schools per district, and each target elementary school receives one Indonesia Mengajar teacher.

Teachers are sent to a school for up to five years. However, because each teacher is only contracted to teach for a year, the school will receive a new Indonesia Mengajar teacher every year for the duration of the programme. The target schools take teachers' placements from Indonesia Mengajar as given, but the headmasters have discretion in assigning duties to the Indonesia Mengajar teachers.

2.3 Teacher Preparation, Assignment, and Deployment

Indonesia Mengajar sends two cohorts per year: one in November–December and another in July. The organization views them as equivalent. The staggered timing happened because the recruitment drive for the first cohort was so unexpectedly successful, with more than 1,300 completed applications for just fifty-one places that the organization saw it fit to expand its operation into two recruitment-deployment cycles per year (Gozali 2020).

Indonesia Mengajar prepares the teachers they recruited with a six- to eight-week intensive preparation camp. During this pre-deployment camp, the teachers receive pedagogy training from education experts, study the national curriculum standards for grades one to six, take part in classroom practicums, and participate in leadership exercises.

The assignment of teachers to programme districts and individual schools is conducted in the latter half of the training camp. The aim is to achieve a balance in the following dimension across districts: gender, religion, and STEM/humanity majors. The majority of teachers come from Java. However, for those who are not from Java, the programme favours teachers from eastern Indonesia for assignments in the western region and vice versa. Indonesia Mengajar does not take the teachers' personal assignment preferences into account, and the teachers take their district and school assignments as given.

Headmasters in the programme schools assign the teachers to teach students either as homeroom teachers or across grades as subject teachers. In the afternoon, many give extra lessons to students, teach at nearby secondary schools, or hold Quran reading classes. During their year-long tenure at the assigned school, the organization also charges individual teachers to provide training to other teachers and engage in education advocacy with local stakeholders.

Table 4 shows that half of the first cohort teachers were homeroom teachers, while the other half were subject teachers. While Indonesia Mengajar teachers had frequent contact with students of all grades, their

FIGURE 1
Districts where Indonesia Mengajar Sent Teachers in 2010–15 (Cohorts 1–10)

SOURCE: Indonesia Mengajar website.

TABLE 4
Indonesia Mengajar Teacher Activities, Cohort 1

Activities	*No. of Teachers*	*%*
Home teacher	26	51%
Grade 2	3	6%
Grade 3	5	10%
Grade 4	5	10%
Grade 5	12	24%
Grade 6	7	14%
Subject teachers any grade	25	49%
Any grade 6 subject	24	47%
Math grade 6	11	22%
Indonesian grade 6	4	8%
Science grade 6	6	12%
After-hours Grade 6 lessons		
Grade 6 home teachers	1	2%
Non-grade 6 home teachers	4	8%
Subject teachers	7	14%
Teachers' capacity-building events		
Within school	14	27%
Subdistrict clusters	20	39%
Teaching hours at non-programme schools		
Other elementary	3	6%
Junior high schools	2	4%
Senior high schools	3	6%
Total cohort 1 IM teachers	51	

NOTES: Tabulation of cohort 1 Indonesia Mengajar teacher activities. Data from Indonesia Mengajar operation records.
SOURCE: Author's calculations.

interactions with sixth-grade students merit further detail. Indonesia Mengajar teachers who taught across grades were often assigned grade six for the specific subjects that they were teaching, while homeroom teachers for grades one to five often taught multiple classes simultaneously (including grade six) because they substituted absent teachers. Beyond regular school hours, many Indonesia Mengajar teachers also provide afternoon lessons for grade six students in preparation for the exit examination. Overall, more than three-fifths of them interacted with students in grade six during scheduled instruction time, but a higher proportion could impact these students in practice.

3. Empirical Strategy

3.1 Regression Specification and Data

This study estimates the impact of the Indonesia Mengajar programme using a difference-in-difference approach. Essentially, it compares treated and control schools before and after programme implementation. The identification in this approach relies on the assumption of parallel trends (i.e., that outcome trends would be similar in both treated and comparison schools in the absence of treatment). The treated group consists of schools receiving the first cohort of Indonesia Mengajar teachers. The control group is a mixture of schools receiving Indonesia Mengajar teachers after the first cohort and other primary schools near the treated school that did not receive such teachers.

The empirical strategy leverages the unsynchronized timing between primary students' grade six exit examination and the programme teacher deployments. Indonesian primary school students sit for an exit examination at the end of their sixth grade, which usually takes place in May. In 2011, this examination took place two months before the second Indonesia Mengajar deployment in July, and grade six students in comparison schools where Indonesia Mengajar was to send the second cohort remained unexposed to programme teachers. Meanwhile, students in the treatment schools had been exposed to the programme since November 2010, which allows us to interpret the resulting estimates of the programme's impact after six months.[6]

The basic regression specification is as follows:

$$Score_{st} = \alpha + \sum_t \beta_t \, \mathrm{IM}_s \times \mathrm{year}_t + \gamma \, \mathrm{IM}_s + \sum_t \delta_t \, \mathrm{year}_t + \varepsilon_{st} \tag{1}$$

where $Score_{st}$ is the school s's examination score in year t, IM_s is a dummy variable for the treatment schools where Indonesia Mengajar sent their first cohort teachers, and year_t is a set of year dummy with 2010 as the omitted year. Our coefficient of interest is β_{2011}, which represents the impact of exposure to Indonesia Mengajar teachers at programme schools.

The dataset for this analysis comes from the Indonesian Ministry of Education's 2008–11 records. Because the dataset has a panel structure, we can estimate an alternate specification with fixed effects, as follows:

$$Score_{st} = \alpha + \sum_t \beta_t \, \mathrm{IM}_s \times \mathrm{year}_t + \mathrm{schoolFE}_s + \delta_t + \varepsilon_{st} \tag{2}$$

The inclusion of school fixed effects allows us to adjust for characteristics that do not vary with time, but which could influence the outcomes, such as location-specific characteristics. The estimates from this equation will be the preferred specification throughout the analysis. The standard errors are clustered two-way at the school level and the year level (Cameron, Gelbach, and Miller 2010).

The dataset recorded the scores for examinations that covered materials from grades four to six. The examinations were not identical across regions because they were prepared by committees at the provincial level. In finalizing the examinations, provincial committees were required to use questions from the national test bank and locally written tests in a twenty-five/seventy-five proportion. Nevertheless, the mathematics examinations were likely to be comparable across regions for two reasons. First, the mathematics curriculum in grades four to six was structured with significant overlaps in topics across grades (e.g., fractions and integer operations are progressively covered every year in the January semester). This consolidates the possible range of topics for the examination into just several core topics. Furthermore, the committees were also bound by a legal guide in the form of a ministerial decree that explicitly stipulates the competencies to include in the examination (see, e.g., Education Ministry Decree No. 2/2011). These provided assurances on the comparability of the mathematics examinations across regions and years.[7]

The Ministry dataset records the minimum, average, and maximum mathematics scores for each school. These scores should reflect the ability of the weakest student in class, the average student, as well as the strongest student. These details allow an investigation of the impact of programme teachers on students with various ability levels.

3.2 Classroom Instructions

If there are other changes to treatment schools concurrent with the programme implementation, then this would undermine the interpretation of the estimated coefficient of interest as the impact due to the Indonesia Mengajar teachers. Here the study examines a possible scenario in which the programme led to existing teachers increasing their efforts after the Indonesia Mengajar teachers arrived. This could be triggered by the programme's high-visibility status, which brought more awareness and supervision from the headmaster to other teachers or even from the subdistrict superintendents. In this case, the estimated effects are still arguably a result of the programme, although these would be indirect effects instead of being directly due to the Indonesia Mengajar teachers.

To separate the effect of direct instruction, this paper uses Indonesia Mengajar organizational reports that recorded the teaching assignments for all first cohort teachers. It estimates the coefficients for an alternate specification where the Indonesia Mengajar exposure dummy variable is interacted with whether the Indonesia Mengajar teachers have a scheduled classroom instruction time on mathematics, Indonesian, or science (other two-way interaction terms that are collinear are collapsed).

$$Score_{st} = \alpha + \sum_t \phi_t \text{IM}_s \times \text{Y6subject}_s \times \text{year}_t + \sum_t \beta_t \text{IM}_s \times \text{year}_t + \text{schoolFE}_s + \delta_t + \varepsilon_{st} \quad (3)$$

In this specification, Y6subject_s is the dummy variable for scheduled instruction time for grade six in one of the three subjects. The variable Y6subject_s takes on a value of one if the Indonesia Mengajar teacher in school s is teaching mathematics either as a homeroom teacher or a subject teacher, and zero otherwise, and is reported in the regression table as Y6Math. Following this definition, about one-third of the treated schools have a scheduled instruction time for mathematics (Table 4). Indonesian and science instruction are constructed in the same way, and are reported as Y6Indonesian and Y6Science, respectively. As before, the 2010 year is the omitted category for the year dummies.

The coefficient ϕ_{2011} allows us to assess the effect of scheduled classroom instructions directly from Indonesia Mengajar teachers beyond the effect of being in a school where an Indonesia Mengajar teacher has been assigned. Specifically for mathematics instruction, this study compares treated schools where the Indonesia Mengajar teacher taught mathematics and treated schools where the Indonesia Mengajar teacher did not teach mathematics. The estimates that we recover will be equivalent to running the specification in equation (2) with the Y6math_s dummy in place of the IM_s for the subsample of treated schools, while avoiding the loss of precision from discarding observations in the study sample. The differential impact of scheduled instruction time is thus identified under the assumption of parallel trends for schools assigned to Indonesia Mengajar teachers who taught mathematics and schools receiving Indonesia Mengajar teachers who did not teach mathematics. The estimation results are discussed in the next section.

4. Results

4.1 Main Results

This study finds that exposure to Indonesia Mengajar teachers is associated with higher average mathematics scores for their students: the coefficient β_{2011} for the mean score is 0.14 points, and is statistically different

TABLE 5
Impact of Exposure to Indonesia Mengajar Programme on
Grade Six Mathematics Exit Examination Score

	(1)	*(2)*	*(3)*
	Avg math	*Min*	*Max*
IM x 2008	0.11	0.08	0.02
	(0.16)	(0.16)	(0.18)
IM x 2009	0.06	0.16	0.04
	(0.09)	(0.12)	(0.07)
IM x 2010	0	0	0
	(.)	(.)	(.)
IM x 2011	0.14*	0.20***	0.08
	(0.05)	(0.02)	(0.13)
control mean	4.8	3.7	6
control SD	0.9	1	1.3
N	825	825	825

NOTES: This table reports the estimates of equation (2) based on exit examination data from the Ministry of Education 2008–11. The outcomes of interest are mean, minimum, and maximum mathematics scores from the exit examination in a given year. Control mean and SD is the average score and its standard deviation among non-treatment schools in 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by school and year.
SOURCE: Author's calculations.

from zero at a 10 per cent significance level (Table 5, column 1). Compared to the 0.9 points standard deviation of mean score among control schools in 2010, the estimated effect for mean mathematics score is equivalent to a 0.16 standard deviation.

Indonesia Mengajar teachers seem to be particularly effective in teaching the weakest students, raising the minimum score by 0.20 points (0.20 standard deviation, column 2). Meanwhile, the estimated effect on the maximum examination score is positive at 0.08 points, but is lower than the effect on the average score and not significantly different from zero (column 3).

The impact on mathematics scores for the Indonesia Mengajar programme lines up with benchmark estimates from TFA, which is the most evaluated programme of its kind (Turner et al. 2018). The most recent randomized evaluation of the programme shows that students of TFA teachers in grades one and two perform significantly better in mathematics by 0.16 standard deviations (Clark and Isenberg 2020). This finding is in line with earlier randomized evaluation results in Decker et al. (2004), which report a better performance of TFA students in mathematics by 0.15 standard deviations. In middle and high school, Clark et al. (2013) report that TFA teachers increased their students' mathematics achievements by 0.07 standard deviations. In England and Wales, a difference-in-difference evaluation of Teach First shows positive and statistically significant improvements in the students' General Certificate of Secondary Education (GCSE) scores by 0.05 and 0.08 standard deviations in years two and three of TF roll-out (Allen and Allnutt 2017).

It is worth noting that the aforementioned TFA and TF evaluations measured the impact on students after a more prolonged exposure than the Indonesia Mengajar teachers in this study. Fellows with the TFA and TF programmes typically teach for two years, while Indonesia Mengajar teachers are only contracted to teach for a year. In practice, for this study, the students were observed just six months after the start of Indonesia Mengajar teachers' deployment to treated schools (November 2010 to May 2011). Suppose students benefit from more exposure to Indonesia Mengajar teachers with stronger academic ability. In this case, the estimates in this study may understate the total learning gains the students in treated schools achieved during the entire duration of the programme.

This was the case for an intervention in India that provided government schools with contract teachers (*balsakhi*) to work with students who were falling behind their peers. An evaluation of this intervention in the cities of Vadodara and Mumbai showed that the remedial education programme increased average test scores in the treatment schools by 0.14 standard deviations in the first year, and 0.28 in the second year (Banerjee et al. 2007). More generally, McEwan's (2015) meta-analysis for education interventions in developing countries highlighted the potential of using contract teachers to improve student achievements. In his review, he identified eight studies with a contract or volunteer teacher intervention, with a mean effect size of 0.10 standard deviations on student achievements. However, he noted that these interventions often implied a reduction in class size, and it is still not clear whether smaller classes are a necessary condition for the effectiveness of contract teachers.

The programme's effects on the average student and the highest scoring student do not attain precision at the conventional statistical significance level of 5 per cent, which may be caused by the Indonesia Mengajar dummy variable recording student exposure with noise. While more than 60 per cent of Indonesia Mengajar teachers had a class schedule with grade six students in any subjects, not all of them did.[8] The next subsection explores the role of scheduled classroom instructions.

4.2 Classroom Instructions

The estimated effects on the average and minimum mathematics examination scores appear to be driven by classroom instructions from Indonesia Mengajar teachers. Table 6 shows the estimated coefficients for the interaction with a dummy variable for mathematics instruction. The magnitude of the interaction terms' coefficients suggests that classroom instructions drove the main result. The mean score increased by 0.25 points (significant at the 10 per cent level), the minimum score by 0.40 points (at the 5 per cent level), and the maximum score by 0.29 points (not statistically significant). For the weakest students, this is a meaningful increase. This increase may bring their score from an average of 3.7 to above a 4.0 mark, which is the guideline threshold for graduation as outlined in the ministry regulation.[9]

The higher impact on the weaker students' test scores is consistent regardless of which subject the Indonesia Mengajar teachers taught them. When the students were exposed to the Indonesia Mengajar teachers through classroom instruction in Indonesian, the minimum mathematics score increased by 0.74 points, which is higher than the estimated effect for the mean score at 0.08 points (not significant, Table 7). For Indonesia Mengajar teachers teaching science (Table 8), the minimum mathematics score has the biggest estimated effect of all, with an increase of 1.04 points, which is again higher than the mean score with an increase of 0.72 points. All the estimated effects for minimum mathematics score are significantly different from zero at the 5 per cent level. None of the estimates for maximum mathematics scores is statistically significant.

These results suggest that the students benefited from the use of mathematics concepts in science lessons and more intensive use of the national language. Nationwide, only one in four individuals uses Indonesian at home, and most of the population speaks local languages at home. Because the examinations

TABLE 6
Impact of Indonesia Mengajar Exposure on Mathematics Score by Classroom Instructions in Mathematics

	(1)	*(2)*	*(3)*
	Avg math	*Min*	*Max*
IM x 2008	0.1	–0.01	0.15
	(0.18)	(0.2)	(0.2)
IM x 2009	0.1	0.08	0.13
	(0.13)	(0.15)	(0.09)
IM x 2010	0	0	0
	(.)	(.)	(.)
IM x 2011	0.05	0.07	–0.01
	(0.07)	(0.04)	(0.17)
IM x Y6 Math x 2008	0.05	0.28	–0.42
	(0.32)	(0.27)	(0.34)
IM x Y6 Math x 2009	–0.11	0.27	–0.28**
	(0.12)	(0.21)	(0.05)
IM x Y6 Math x 2010	0	0	0
	(.)	(.)	(.)
IM x Y6 Math x 2011	0.25*	0.40**	0.29
	(0.1)	(0.12)	(0.27)
2010 control mean	4.8	3.7	6
2010 control std dev	0.9	1	1.3
N	825	825	825

NOTES: This table reports the estimates of equation (3) based on exit examination data from the Ministry of Education 2008–11 and Indonesia Mengajar operational records. The outcomes of interest are mean, minimum, and maximum mathematics scores from the exit examination in a given year. Control mean and SD is the average score and its standard deviation among non-treatment schools in 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by school and year.
SOURCE: Author's calculations.

were conducted in Indonesian, this could suggest that comprehension problems underlie the students' poor mathematics scores, and instructions that improve comprehension can boost performance.

5. Conclusion

Does an alternative teacher placement programme that sends college graduates with strong academic and leadership backgrounds to teach rural primary schools impact student outcomes? This paper compares the mathematics score between programme and control schools using a difference-in-difference strategy using the national exit examination dataset from the Ministry of Education. It finds that teachers deployed by the Indonesia Mengajar programme raised the mean score by a 0.16 standard deviation, which was

TABLE 7
Impact of Indonesia Mengajar Exposure on Mathematics Score by Classroom Instructions in the Indonesian Language

	(1)	*(2)*	*(3)*
	Avg math	*Min*	*Max*
IM x 2008	0.15	0.02	0.18
	(0.16)	(0.17)	(0.18)
IM x 2009	0.03	0.07	0.07
	(0.09)	(0.13)	(0.08)
IM x 2010	0.00	0.00	0
	(.)	(.)	(.)
IM x 2011	0.12	0.05	0.14
	(0.06)	(0.03)	(0.14)
IM x Y6 Indonesian x 2008	–0.17	0.27	–0.78
	(0.45)	(0.40)	(0.44)
IM x Y6 Indonesian x 2009	0.16	0.47	–0.15
	(0.23)	(0.31)	(0.09)
IM x Y6 Indonesian x 2010	0.00	0.00	0
	(.)	(.)	(.)
IM x Y6 Indonesian x 2011	0.08	0.74**	–0.33
	(0.13)	(0.13)	(0.36)
2010 control mean	4.8	3.7	6
2010 control std dev	0.9	1.0	1.3
N	825	825	825

NOTES: This table reports the estimates of equation (3) based on exit examination data from the Ministry of Education 2008–11 and Indonesia Mengajar operational records. The outcomes of interest are mean, minimum, and maximum mathematics scores from the exit examination in a given year. Control mean and SD is the average score and its standard deviation among non-treatment schools in 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by school and year.
SOURCE: Author's calculations.

significant at the 10 per cent level. The weakest students benefited most from exposure to the programme, with an increase of 0.20 standard deviation, which is more precisely estimated at the 5 per cent level. The estimated effects are higher for the weakest students who had classroom time with programme teachers, with bigger gains from Indonesian and science instruction of up to 1.04 points.

This study provides new evidence on programmes that are modelled on a Teach for America programme from a developing country. TFA-style programmes have spread globally based on the idea that they are an effective intervention to address achievement gaps in rural or disadvantaged areas. However, virtually no rigorous evaluation has been done in countries other than the US and the UK. This study presents the first attempt to estimate the causal impact of such programmes outside the original two countries. The findings from this evaluation suggest that, especially for the weakest students in rural schools, improvements in their teacher quality may lead to meaningful academic improvements in their achievements. At the same

TABLE 8
Impact of Indonesia Mengajar Exposure on Mathematics Score by Science Classroom Instruction

	(1)	(2)	(3)
	Avg math	Min	Max
IM x 2008	0.26	0.23	0.05
	(0.16)	(0.19)	(0.15)
IM x 2009	0.09	0.15	0
	(0.10)	(0.12)	(0.06)
IM x 2010	0.00	0.00	0
	(.)	(.)	(.)
IM x 2011	–0.02	–0.03	–0.03
	(0.06)	(0.04)	(0.17)
IM x Y6 Science x 2008	–0.64	–0.67	–0.12
	(0.42)	(0.29)	(0.59)
IM x Y6 Science x 2009	–0.12	0.06	0.19
	(0.19)	(0.36)	(0.18)
IM x Y6 Science x 2010	0.00	0.00	0
	(.)	(.)	(.)
IM x Y6 Science x 2011	0.72**	1.04**	0.5
	(0.16)	(0.18)	(0.27)
2010 control mean	4.8	3.7	6
2010 control std dev	0.9	1.0	1.3
N	825	825	825

NOTES: This table reports the estimates of equation (3) based on exit examination data from the Ministry of Education 2008–11 and Indonesia Mengajar operational records. The outcomes of interest are mean, minimum, and maximum mathematics scores from the exit examination in a given year. Control mean and SD is the average score and its standard deviation among non-treatment schools in 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by school and year.
SOURCE: Author's calculations.

time, the low level of baseline achievements may have been driving the positive results here. Finally, the education policy community would benefit from more empirical studies on similar programmes.

Acknowledgements

This project benefited from discussions with Aisy Ilfiyah, Ajeng Tri Handini, Frina Lin, Marliyanti, and Patrya Pratama, as well as participants of the Development Reading Group at Boston University. I am grateful to Galih Ramadian Nugroho Putra, Hikmat Hardono, Lizara Patriona, Subhanudin Husen, and the leadership in Indonesia Mengajar for facilitating various data access. The views expressed in this paper are those of the author and should not be attributed to the various individuals and organizations acknowledged above.

NOTES

1. Disclosure: The author was a teacher in the Indonesia Mengajar programme, cohort V (November 2012 to January 2014).
2. This unintuitive relationship between low teacher competency score and their education degree could be driven by several characteristics of the higher education system in Indonesia. First, education college degrees are predominantly offered by private institutions, which on average are of lower quality than public universities. Wicaksono and Friawan (2011) noted that about 75 per cent of PhDs in Indonesia are concentrated in just four public universities (UI, ITB, UGM, and IPB, which are all located in Java and are major contributors to Indonesia Mengajar teacher recruits). Another factor is student sorting. High school graduates with a high ability sort into top universities and lower quality students sort into education majors, which have a less strict screening process. The sorting effect may also be exacerbated by the differential survival rates of education majors by ability. A high-performing college student with an education major may choose to exit the field for a better paying job than a low-paying entry-level teaching job (Chang et al. 2014).
3. *Pengerahan Tenaga Mahasiswa*/College Student Send-down.
4. Agreement by the district depended on the cooperation of the district's education office but in early cohorts the district head (Bupati) and the head of district education office would be honoured with a reception at the Vice President's office before the deployment of the Indonesia Mengajar teachers to the destination districts. Then Vice President Boediono was a personal supporter of the programme.
5. Initially, there were fourteen districts in the first year (2010–11). However, conflict between the state military and the Aceh separatist movements forced the programme's relocation from Aceh Utara to Musi Banyuasin and Muara Enim in South Sumatra. In November 2012, Indonesia Mengajar re-added Aceh Utara and added Banggai to its programme districts, bringing the total to seventeen districts.
6. While a longer-term evaluation with a panel data that extends beyond 2011 would also be of interest, the author does not have access to this dataset.
7. The comparability is harder to establish for examinations in other subjects such as the Indonesian language and science. The 2011 Ministerial Decree listed 34 to 43 per cent more competencies to cover in the examination for the subjects of Indonesian language and science (thirty and twenty-six, respectively, compared to seventeen for mathematics). These stemmed from heavier loads in its grade 4 to 6 curriculum with twenty-four and twenty-seven competencies to cover for Indonesian and science without overlaps across grades, whereas mathematics only has twenty-one competencies with significant overlaps.
8. An ideal evaluation using the same difference-in-differences approach for this programme would prospectively collect grade-level measures of academic ability using the same test for students in both treated and comparison schools. The econometrician could then estimate the programme effect while taking into accounts the difference in teacher assignments across treated schools (cf. Banerjee et al. 2007). Unfortunately, the Indonesia Mengajar programme did not embed such an evaluation plan in their roll-out and the ministry only collected school-level statistics for the exit examination for grade six.
9. Education Minister Decree 59/2011 stipulated that secondary school students can graduate if they score at least 4.0 in their final score in all of their examination subjects. The final score is a weighted average of the examination score (60 per cent) and semester report cards (40 per cent).

REFERENCES

Allen, R., and J. Allnutt. 2017. "The Impact of Teach First on Pupil Attainment at Age 16". *British Educational Research Journal* 43, no. 4: 627–46. https://doi.org/10.1002/berj.3288

Ashraf, N., et al. 2020. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services". *American Economic Review*. https://doi.org/10.1257/aer.20180326

Banerjee, A.V., et al. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India". *Quarterly Journal of Economics* 122, no. 3: 1235–64. https://doi.org/10.1162/qjec.122.3.1235

Bold, T., et al. 2019. *The Lost Human Capital: Teacher Knowledge and Student Achievement in Africa*. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-8849

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering". *Journal of Business and Economic Statistics* 29, no. 2: 238–49. https://doi.org/10.1198/jbes.2010.07136

Chang, M.C., et al. 2014. *Teacher Reform in Indonesia: The Role of Politics and Evidence in Policy Making Human Development*. https://doi.org/10.1596/978-0-8213-9829-6

Chaudhury, Nazmul, et al. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries". *Journal of Economic Perspectives*. https://doi.org/10.1257/089533006776526058

Chen, Y., et al. 2020. "Arrival of Young Talent: The Send-down Movement and Rural Education in China". *American Economic Review* 110, no. 11: 3393–430. https://doi.org/10.1257/AER.20191414

Chetty, R., J.N. Friedman, and J.E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-added and Student Outcomes in Adulthood". *American Economic Review* 104, no. 9: 2633–79. https://doi.org/10.1257/aer.104.9.2633

Clark, M.A., and E. Isenberg. 2020. "Do Teach for America Corps Members Still Improve Student Achievement? Evidence from a Randomized Controlled Trial of Teach for America's Scale-up Effort". *Education Finance and Policy* 15, no. 4: 736–60.

Clark, M.A., et al. 2013. *The Effectiveness of Secondary Math Teachers from Teach for America and the Teaching Fellows Programs*. Tech. rep.

Decker, P.T., D.P. Mayer, and S. Glazerman. 2004. "The Effects of Teach for America on Students: Findings from a National Evaluation". University of Wisconsin—Madison, Institute for Research on Poverty Discussion Paper.

Eble, A., et al. 2021. "How Much Can We Remedy Very Low Learning Levels in Rural Parts of Low-income Countries? Impact and Generalizability of a Multi-pronged Para-teacher Intervention from a Cluster-randomized Trial in the Gambia". *Journal of Development Economics* 148, September 2020, p. 102539. https://doi.org/10.1016/j.jdeveco.2020.102539

Evans, D.K. and A. Popova 2016. "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews". *World Bank Research Observer* 31, no. 2: 242–70. https://doi.org/10.1093/wbro/lkw004

Finan, F., B.A. Olken, and R. Pande 2017. "The Personnel Economics of the Developing State". In *Handbook of Economic Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, pp. 467–514. North-Holland.

Glewwe, Paul, et al. 2013. "School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010". In *Education Policy in Developing Countries*, pp. 13–64. Chicago: Chicago University Press.

Gozali, C. 2020. "From Early Childhood to Adulthood: Leader Development in Indonesia". PhD thesis, Claremont Graduate University.

Lakshminarayana, R., et al. 2013. "The Support to Rural India's Public Education System (STRIPES) Trial: A Cluster Randomised Controlled Trial of Supplementary Teaching, Learning Material and Material Support". *PLoS ONE* 8, no. 7. https://doi.org/10.1371/journal.pone.0065775

McEwan, P.J. 2015. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments". *Review of Educational Research* 85, no. 3: 353–94. https://doi.org/10.3102/0034654314553127

Teach For All. 2021. "Teach For All Network Partners". https://teachforall.org/network-partners

Thomas, M.A.M., K. Crawford-Garrett, and E. Rauschenberger. 2021. "A Growing Global Network". In: *Examining Teach For All*: *International Perspectives on a Growing Global Network.* Routledge.

Turner, H., et al. 2018. "What Are the Effects of Teach for America on Math, English Language Arts, and Science Outcomes of K–12 Students in the USA?". *Campbell Systematic Reviews* 14, no. 1: 1–60. https://doi.org/10.4073/csr.2018.7

Wicaksono, T.Y. and D. Friawan. 2011. "Recent Developments in Higher Education in Indonesia: Issues and Challenges". In *Financing Higher Education and Economic Development in East Asia*, edited by Shiro Armstrong and Bruce Chapman. Canberra: ANU E Press.