Linking Vocational Schools to Industry: Effects on Teachers in Indonesia*

César Garro-Marín†
University of Edinburgh

Masyhur A. Hilmy‡
University of New South Wales

July 2024

Abstract

This paper evaluates a mass training program for in-service vocational school teachers in Indonesia. The government rolled out an intensive, field-specific professional development program to enhance teachers' vocational skills. Training was provided by experienced private sector firms. We assess its effects on teachers' knowledge, classroom practices, and expectations of students' outcomes using a randomized evaluation. The results show minimal impact on these areas. Survey data from training applicants suggest several reasons for the lack of effect: a mismatch between training offerings and existing skills gaps, insufficient post-training support, and untreated teachers accessing training from alternative providers.

Keywords: teacher training, vocational education

JEL Codes: I21, I25, I29

*We thank our partners at Mitrasdudi Kemendikbud for their support and assistance. This paper benefited from feedback and suggestions from Sam Bazzi, Joshua Goodman, Daniele Paserman, and seminar participants at the Boston University Development Reading Group and the 2024 Australasian Development Economic Workshop at the University of Melbourne. We thank Muhamad Haris Zamaludin Setiadiputra for providing outstanding research assistance. Alifia Firliani, Aulia Larasati, Ma'rifatul Amalia, and Terry Muthahhari provided invaluable support for this research project. This research is funded by the Australian Government through J-PAL Southeast Asia – IRF Fund and the Weiss Fund for Research in Development Economics. This project received approval from Boston University Charles River Campus IRB, protocol 5535X. This research was registered at the AEA RCT registry (AEARCTR-0014062). The views expressed here are those of the authors, and do not necessarily reflect those of the various individuals and organizations acknowledged above.

†School of Economics. Email: cgarrom@ed.ac.uk.

‡School of Economics. Email: m.hilmy@unsw.edu.au.

1 Introduction

Building a labor force with adequate technical skills is a pressing challenge for education systems in less developed countries (Atkin et al., 2019). Secondary education and, in particular, vocational high schools play a crucial role in addressing this issue. Globally, vocational high schools service more than 48 million students across low and middle-income countries (EdStats, 2022). Thus, policy interventions seeking to improve their effectiveness could be valuable tools for boosting the employment outcomes of young people in these countries.

In this paper, we evaluate a large policy intervention aimed at improving the effectiveness of vocational education in Indonesia. In 2020, the Indonesian Ministry of Education and Culture (MoEC) rolled out a professional development program for vocational high-school teachers called the Upkilling and Reskilling Training Program (UR). This program allocated thousands of teachers across dozens of trainings specific to their vocational field. UR introduced several innovations relative to training programs studied in the literature: (i) it was intensive, with the average course lasting between 6 and 8 weeks;[1] (ii) the trainings were designed and supplied by firms demanding vocational graduates in the labor market, and (iii) it was done at scale, with 2,700 teachers from more than 1,600 different high schools participating in the program. Through these features, MoEC hoped that teachers and schools would align their teaching with the needs of the private sector while also enhancing the links between vocational schools and potential employers of their graduates.

We worked closely with MoEC and took advantage of the substantial excess demand for the program by randomizing participant selection for six vocational subjects. We selected these subjects following two criteria: whether the training was oversubscribed, and whether MoEC had direct oversight of the training selection and implementation. Teachers assigned to treatment and control groups have balanced characteristics.

We focused our evaluation on understanding whether the training changed teachers' knowledge, their classroom practices, or their expectations of their students' success.[2] We collected data on these outcomes through an original survey developed in consultation with MoEC conducted approximately ten months after the end of the trainings.

We find little evidence that the UR had an economically relevant impact for most of the outcomes we considered. We find null effects on teachers' vocational knowledge, classroom practices, and expectations of their students. For the teachers' expectation of their students' employment rate three months after graduation, the 95% confidence intervals (CI) rule out positive effects larger than 0.12 standard deviations (SD; of the control group's outcome distribution). This bound is lower than the expected range of sizable 0.20-0.40 SD effects that can be attributed to intensive teacher training programs in Fryer (2017)'s review. The 95% CI of our estimates are also sufficiently narrow to rule out increases of US$2.87 per month or higher in the students' first salary. We also rule out

[1] This is substantially longer than the 2.5-week median length of the typical teacher professional development program worldwide (Popova et al., 2022).

[2] We focused our evaluation on teachers because they are the primary targets of the UR program are the teachers and the primary impacts of the program should arise on changes or teacher behavior first.

large changes in teachers' vocational knowledge, the distribution of classroom time, and the use of IT in the classroom. However, we find some evidence that treated teachers became more optimistic about how prepared their students are for the job market. We find similar results in alternative but larger sample where we matched treated teachers to suitable controls using propensity scores on a rich set of pre-training characteristics.

Analysis of our survey responses revealed three possible reasons for the lack of impact. First, the training program was not tailored to address teachers' skill gaps, leading to a mismatch between the training implementation and teachers' needs. MoEC's perceived skill gap among its teachers led to the creation of the upskilling program, but the training did not target teachers with identifiable weak vocational skills. This led to 80% of attendees reporting that they were already familiar with the materials covered during training. Second, after the conclusion of the program, the treated teachers lacked the support needed to make a sustained change. In particular, only 26% of teachers reported any follow-up sessions after the training and more than half mentioned the need for support advocacy to the school principals. Finally, we document that alternative professional development programs remain accessible to teachers in the comparison group. Teachers assigned to the control group reported participating in trainings other than UR, leading to no systematic difference in overall training hours between the two groups. This evaluation suggests that in an environment where training opportunities are not scarce, policymakers can instead provide clear guidelines for precisely targeted in-service vocational training that allows sustained collaboration with the private sector instead of launching an upskilling program from scratch.

Our study contributes to the literature of teacher professional development in developing countries. In-service professional development programs for teachers are widespread, but rigorous evaluations of such programs remain scarce. The evidence base on teacher training primarily comes from studies in high-income countries (Fryer, 2017; Yoon et al., 2007). Moreover, evaluations of teacher training programs implemented in low- and middle-income countries have shown mixed effectiveness (Popova et al., 2022). While programs in Argentina and South Africa showed positive effects (Albornoz et al., 2020; Cilliers et al., 2020), different programs in China, Nepal, and Rwanda did not show any effect (Blimpo and Pugatch, 2021; Loyalka et al., 2019; Schaffner et al., 2024), and a teacher training program in Costa Rica led to *worse* student outcomes (Berlinski and Busso, 2017).[3]

We add three distinct contributions to existing studies. First, we provide the first evidence of a program that wanted to improve teaching quality in vocational schools. Policymakers in developing countries place a high priority on vocational education, in marked contrast to international donors who recently redirected their focus toward improving a narrower measure of foundational learning outcomes (Crawfurd et al., 2021). This divergence suggests that efforts to bolster evidence-informed policymaking ought to take into account policymakers' priorities, which calls for more evidence on vocational education. Researchers have run RCTs of vocational programs, but most of these evaluated the effectiveness of short vocational training targeting low-skilled youth (Alfonsi et al.,

[3] Other studies bundled teacher training with inputs for students: tablets in Pakistan or textbooks in Papua (Beg et al., 2022; Zaw et al., 2021). A related literature has also investigated the impact of coaching as a form of teacher professional development. Studies include Cilliers et al. (2020) in South Africa, Majerowicz and Montero (2018) in Peru, Yoshikawa et al. (2015) in Chile, and Carneiro et al. (2022) in Ecuador. The first two studies show positive effect on student achievements, while the latter two do not.

2020; Attanasio et al., 2017, 2011) with a few RCTs examining secondary vocational education programs (Field et al., 2019; Hicks et al., 2011). These studies take the quality of vocational education as given, making our study as –to the best of our knowledge– among the first to evaluate interventions aiming at improving the effectiveness of vocational education itself.

Second, we provide evidence from a teacher professional development program that targets the upper secondary level. Until recently, the evidence on teacher training programs in developing countries consisted of interventions targeting primary school teachers (Null et al., 2017).[4] Nevertheless, recent papers have started to add evidence on training for lower secondary teachers: Berlinski and Busso (2017), Loyalka et al. (2019), and Schaffner et al. (2024) evaluated the effect of training programs that target junior secondary math teachers. Blimpo and Pugatch (2021) is a notable exception as they report the results of a comprehensive training program for upper secondary teachers in Rwanda. As developing countries continue to see the increase in (upper) secondary enrollment stemming from the near-universal access to primary schools, building the evidence base on post-primary education becomes a vital priority (Banerjee et al., 2013).

Third, our analysis is based on an evaluation of an intensive (260 hours), subject-specific, and at-scale teacher professional development program. Teacher training associated with specific methods has been highlighted in multiple systematic reviews to improve student learning in developing countries (Evans and Popova, 2016). Subject specificity and multiple-day training are both features of PD programs that are deemed promising to boost student learning outcomes (Popova et al., 2022). At the same time, implementations of promising interventions tested in smaller trials often pose challenges when they are being scaled up or delivered more cheaply. Ganimian (2020) finds null effects for an at-scale intervention on growth mindset, and Kerwin and Thornton (2021) find a weaker effect when a mother-tongue literacy program is delivered at a lower cost. Angrist and Meager (2023) find that variations in the literature of targeted instruction can be attributed to the degree of implementation and program delivery model. Al-Ubaydli et al. (2019) provide a framework to understand the threats to scaling experiments. With the UR role as an umbrella program to train Indonesian vocational school teachers, our data provides a unique window to look into how teacher training is implemented in diverse vocational streams.

2 Context: Vocational Secondary Schools and the Upskilling-Reskilling training

Indonesian vocational high schools (SMK by its acronym in Indonesian) prepare students for entry into the labor market upon graduation (Pritadrajati, 2018). They service approximately five million students every year and account for about half of the total upper-secondary enrolment in the country. Vocational high schools compete with General (SMA) and Islamic high schools (MA)

[4]Ganimian and Murnane (2016)'s review paper identified only three papers on increasing teacher's skills. Two of them are Abeberese et al. (2014)'s training for fourth-grade teachers in the Philippines and Yoshikawa et al. (2015)'s training for pre-K and kindergarten teachers in Chile.

to provide students with upper-secondary level education (grade 10-12).[5] The vocational school curriculum places significant emphasis on vocational training, progressively allocating more scheduled time to vocational subjects and field internships from grade 10 (26 percent) to grade 12 (72 percent).

Vocational schools offer programs in fields as diverse as performing arts, business, IT, energy, and engineering technology. The five most popular programs are computer and network technicians, accounting, light vehicle technicians, office administration, and motorcycle technicians. Three-quarters of all SMK in the country offer at least one of these five programs. While some vocational programs are widely available, others are more niche. Programs such as airplane frame constructions, Javanese shadow puppetry (*wayang*), crustacean aquaculture, thread manufacturing, and fiberglass boat constructions are only offered at a dozen or fewer schools across the country. Overall, MoEC records 256 unique vocational streams in 14 different fields offered across 14,178 vocational schools (Ditjen Vokasi, 2021).

The Indonesian Government's education policy places a lot of emphasis on vocational education. The number of vocational schools has more than doubled since 2005, and the Government intends to continue with this expansion in the near future (Pritadrajati, 2018). However, despite these investments, recent graduates still face high rates of unemployment. Graduates from vocational school graduates aged 30 or less face an unemployment rate that is 45% higher than the average for Indonesians under 30 –19% versus 13%– (Central Bureau of Statistics, Indonesia, 2016). A survey in 2021 by MoEC showed that 48 percent of graduates earn salary below the province minimum wage (Setditjen Vokasi, 2020).

To enhance the competitiveness of its graduates, the Indonesian government initiated a reform of vocational schools aimed at aligning them more closely with industry requirements. In a 2016 presidential decree, the Minister for Education was mandated to "link and match" the vocational curriculum with industry needs through partnerships with the private sector (Government of Indonesia, 2016). In 2020, MoEC introduced the Upskilling and Reskilling training program to fulfill this objective.

2.1 The Upskilling and Reskilling Training Program

In 2020 the Indonesian Ministry of Education rolled out the Upskilling and Reskilling Training Program (UR). This was a professional development program for vocational school teachers, who taught at schools specialized in five broad industries: hospitality, construction, repair of vehicles and machinery, healthcare services, and software and design services. UR had two primary objectives: (i) improving teachers' vocational knowledge, and (ii) creating links between schools and private sector firms.

UR was comprised of a collection of courses that offered content relevant to the program's target

[5] General schools provide a secular education. Islamic high schools use methods similar to secular schools but teach more religious content (Bazzi et al., 2020).

industries. Appendix Table A.1 provides illustrative examples of UR course offerings. These courses cover a range of specialized subjects, including photography, 3D animation, and programming and operation of computer numerical control machines (CNC), among others. While each teacher had the opportunity to apply for multiple trainings, they were only able to attend one if selected. The application process was free and conducted online, resulting in minimal application costs.

UR introduced several innovations relative to professional development programs evaluated elsewhere in the literature. First, the program had substantial private-sector involvement because MoEC partnered with private-sector firms operating in the target sectors to develop and deliver these courses.[6] This heavy cooperation with the private sector had two primary objectives: creating links between vocational schools and potential employers for their students and familiarizing teachers with the skills required by private-sector firms. With this in mind, MoEC only established general guidelines on the format of the courses, but the industry partner had ample autonomy on the course contents. According to the ministry guidelines, potential training providers only needed to meet three broad criteria: (i) being able to organize the training and provide instructors, (ii) having a curriculum and training materials available, and (iii) being able to provide certifications to the training participants (BBPPMPV BOE, 2020).

Moreover, UR provided intensive subject-specific training at scale in more than 50 different vocational subjects. The average training lasted for just over 6 weeks, and, overall, the program trained 2,701 teachers from more than 1,600 vocational high schools spread over 78% (403) of the Indonesian regencies.[7] This contrasts with the typically small professional development program studied in the literature, which usually lasts no more than a couple of weeks (Popova et al., 2022).

UR courses combined online and in-person sessions,[8] with a duration of between six and eight weeks for a total of approximately 260 contact hours. Each course followed a three-phase structure. Firstly, an online phase introduced trainees to the training materials through Zoom and pre-recorded videos. This phase lasted for approximately 30% of the course. Next, the trainees transitioned to in-person class sessions that delivered the contents using the traditional teaching approach. This second phase encompassed approximately 45% of the course. Lastly, teachers proceeded to on-the-job training sessions, where they interned at the facilities of the industry partner. These internships provided teachers with hands-on experience in the day-to-day operations of the company. Following the completion of the training, industry partners awarded certificates to the trainees upon successfully passing an examination. As reference, Appendix Table A.2 shows two illustrative course schedules.

To illustrate the type of training offered by UR, let us consider the course "*Network and Communications*" where teachers learned basic skills for setting up and maintaining internet networks, e.g., configuring and troubleshooting routers, and configuring data traffic priority for network users (see

[6] In a review of existing high-quality evaluations of teacher training intervention in developing countries, none of the trainings for post-primary education were implemented by private sector firms (Schaffner et al., 2024). Popova et al. (2022) surveyed 33 teacher professional development programs but did not report that any of the programs are designed or implemented by private firms.

[7] Regencies are Indonesian administrative units similar to municipalities in other countries.

[8] This was a necessary adaptation to the Covid-19 mobility restrictions in Indonesia.

Table A.3 in the appendix for a breakdown of the training curriculum). This course was provided by a certified training partner of MikroTik.[9] Upon completion, MikroTik issued certifications to teachers who scored higher than a passing grade (60/100) on the final exam.

Teacher participation in UR was voluntary, but it was greatly encouraged by the government. Besides advertising it on social media, MoEC also sent official invitations to apply to vocational school principals. Financial costs for the teachers were small: they could apply at no cost. If selected, the entire cost of the training program is borne by MoEC, including transportation, room and board, stipend, and Covid-19 swab tests during the in-person training. Nevertheless, teachers had to obtain an internet plan –if they lacked one– to ensure internet connectivity during the online portion of the training. On average, MoEC incurred on average cost of $2,907 per participant for a total cost of US$7 million in 2020 (Ditjen Vokasi, 2021).[10]

Although participation was voluntary, teachers also had multiple incentives to apply. First, teachers obtained a certification upon successfully completing the training. Second, schools as a whole also had an incentive to be in UR to unlock additional funding from MoEC. The ministry widely publicized that the number of certified teachers in a school would determine the school's eligibility to receive facility upgrading grants.

The selection of the UR participants was largely decentralized to the training providers. In most cases, MoEC gave ample discretion with regard to participant selection and the training curriculum. However, MoEC has direct control over participant selection for a subset of subjects which we were able to use for the research design.

Figure 1 illustrates UR's timeline. The teachers' application portal was released in June of 2020, after a delay of several months due to the effects of the Covid-19 emergency in Indonesia.[11] Starting in July, teachers had a two-month window to submit their applications. Then, the trainings took place between October and December of the same year. We collected data through a phone survey at the end of 2021, one year after the trainings concluded.

3 Research Design

Randomized Evaluation. Our main results come from a randomized control trial (RCT) we designed in collaboration with MoEC. This randomized evaluation takes advantage of the substantial over-subscription of the UR training program, with more than 24 thousand applications for only 2,468 available slots. Although the excess demand varied by vocational subject, for most courses

[9] MikroTik manufactures a popular brand of network routers and other network equipment. The training for UR participants was provided by PT AsiaVer, its partner based in East Java. Figures B.1-B.2 provide snapshots of the training content and the training activities. MikroTik is a competitor to Cisco, whom MoEC also partnered with to implement other UR training courses.

[10] The budgeted cost was 50% higher than the actual cost. The ministry budgeted IDR137.5 billion (USD 8.9 million) to train 2,160 teachers. Their end-of-year report stated that the actual cost of the program was IDR102.3 billion (USD 6.6 million) to train 2,426 teachers (Ditjen Vokasi, 2021), page 56. Currency conversion uses an exchange rate of IDR 14,500/USD.

[11] Originally, the application portal was intended for release in April of 2020.

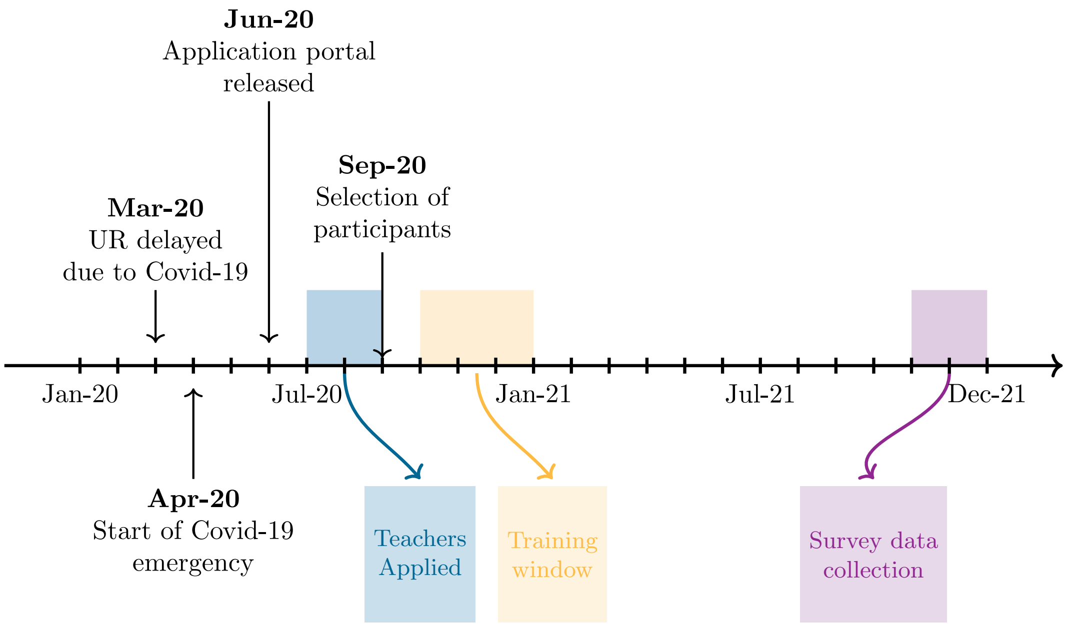


Figure 1: Timeline of the upskilling and reskilling program

there were at least twice as many applicants as slots available (see Table A.4 in the appendix). We were able to select randomly the successful applicants for a subset of over-subscribed vocational subjects.

We selected applicants randomly for several vocational subjects where MoEC had substantial direct control over applicant selection. As we discussed in Section 2.1, the degree of control of MoEC versus the training providers varied by subject. For many of them, MoEC delegated most of the course organization, including the selection of the attendees, to the training provider. This means that the provider had full discretion over the course syllabus and the selection criteria of the training participants for many, but not all, of the courses. Therefore, we randomly selected the trainees in six subjects where MoEC had substantial discretion on trainee selection: topography mapping, network and telecommunications, internet of things, 2D animation, Java programming, and database management. Panel A of Table A.4 details the number of applicants, budgeted slots, and actual attendees for each of the courses.

Our RCT sample is comprised of 400 teachers who applied to the UR training and were assigned either to the treatment or control group. We collected data from these teachers using a phone survey, which we describe in Section 4.1.1.

Matching. We supplemented the RCT design with a sample of nearly 1,100 teachers using a propensity score matching (PSM) design. We adopted this alternative strategy for two reasons: first, to achieve greater precision with a larger sample size, and second, to compare its results with those of the RCT. Similar results under both strategies would provide additional support to the results from the RCT. Appendix Table A.4 (Panel B) shows a list of the 15 courses in the PSM design with the most applicants. We selected the PSM sample by matching treated teachers to suitable control applicants using pre-treatment individual-level data in MoEC administrative databases. We used information about applicants' gender, education level, residence, teaching experience, and competency test score, among others. Further details on the matching approach are given in Appendix C.

4 Data

4.1 Data sources

We use two main data sources: an original phone survey that we use for our main results, and pre-training administrative data from MoEC.

4.1.1 Original survey data

We collected data on post-treatment outcomes through a survey deployed during November and December of 2021, approximately one year after the end of the UR trainings. See Figure 1 for more details on the timeline. We conducted the survey via phone to minimize the challenges posed by the spread of Covid-19 and the logistical costs of reaching teachers spread across the whole Indonesian archipelago.[12] We collected data on three main outcomes: the teacher's knowledge of their vocational field, their classroom practices, and their expectations about their student's success in the labor market. Besides these outcomes, we also collected detailed data on teachers' professional trainings, as well as basic demographic information.

Teachers' vocational knowledge and classroom outcomes are our primary outcomes. Any other impacts on, for instance, students' labor market outcomes would arise only as a result of improvements in teachers' quality or their teaching practices. Alternatively, they could also arise from the second-order effects of strengthened connections between schools and employers.

We measure improvements in teachers' vocational knowledge with a battery of training-specific true-false questions that test the contents taught in the UR trainings. We developed a set of ten questions for each of the six courses in the RCT design based on the training materials and the post-course test that the training partners administered to the participants.[13] We validated these questions with MoEC staff who were in charge of the training implementation. The question validation was done in several stages, and we tested the survey module extensively prior to the launch of the survey to ensure that respondents could answer the questions over the phone.[14]

[12]Overall, we collected information from respondents in 384 districts spread out over 36 provinces (i.e., 75% of all districts and 95% of all provinces). An established company with experience in phone surveys contacted the teachers on behalf of the research team using the teachers' phone number in the MoEC application database. The implementation of the survey by an independent survey company separate from MoEC helped minimize experimenter demand bias among respondents. Research staff at J-PAL SEA ran monitoring and quality assurance steps during the data collection period to ensure the fidelity of the recorded answers (backchecks, high-frequency checks, and spotchecks).

[13]See panel A of Table A.4. Due to the highly decentralized nature of UR, developing a set of ten questions for each of the 56 courses we surveyed was infeasible.

[14]Table A.7 shows the share of respondents who answered each item in the question module correctly, by vocational

Our measures of teaching practices come from a survey module capturing teachers' time use in the classrooms, the equipments they used to teach, as well as their teaching load. Our instrument is informed by the Stallings classroom observation instrument, a tool widely used in developed nations to assess teachers' effectiveness in managing their classrooms (Bruns et al., 2016). The original instrument was developed by Stallings to generate quantitative measures of teachers' use of instructional time and use of materials, including information and communications technology (World Bank, 2015). Stallings observed that teachers' distribution of time across activities can have an impact on student achievement (Jukes et al., 2013; Stallings, 1980). We also included questions designed to capture teachers' adaptation to the effects of the Covid-19 disruptions, which includes information about the share of students with internet access.

We proxy students' downstream outcomes with information about teachers' expectations of their students' labor market outcomes. Specifically, we ask them to estimate the share of their students who are employed three months post-graduation, their average salary, and their university enrollment rates. We ask them to estimate these outcomes for students graduating in May 2021 (after the treated teachers participated in the UR training) and May 2019 (prior to the program and also prior to the pandemic, which caused severe labor market disruption). While these proxies are somewhat limited, in Table A.8 we show that they are strongly correlated with actual student outcomes. The table regresses multiple measures of actual students' outcomes on teachers' expectations. Teachers' answers for the 2019 graduates are correlated with that year's students' national exam scores in math, Indonesian, English, their vocational subjects, and the overall average score. Given that the national exam was cancelled during the Covid-19 pandemic, teachers' expectations for the 2021 graduates provided us a practical source of information of the students' outcomes.

We also collected detailed data on professional trainings modeled on the ITTSI questionnaire in Popova et al. (2022) to systematically capture implementation characteristics. These characteristics allow us to investigate how training organization, content, and delivery may influence the targeted outcomes.

4.1.2 MoEC administrative data

Our pre-treatment administrative data comes from the 2015 Teacher Competency Exam database (UKG by its acronym in Indonesian). This is an administrative dataset containing individual-level information covering the universe of active school teachers in 2015. It contains data on demographic characteristics (age, education, gender, place of residence, etc), employment characteristics (school of employment, vocational specialty), and the teacher's test scores in the National Competency Exam.

The National Competency Exam was initially administered as a part of the teacher certification process that allows teachers to unlock a salary supplement. However, in 2015, the MoEC ran a

sector. Despite its simple true-false format, respondents still need to consider the content of the statement. None of our questions were too obvious for our respondents, with on average respondents able to score 72% on this simple test. Only a minority of respondents (7%) were inattentive, giving the same answer to all true-false statements.

nationwide exam to measure and identify gaps in teachers' quality across the country (Menteri Pendidikan dan Kebudayaan, 2015). Teachers were tested with a two-hour exam on two areas: proficiency in their teaching subjects (70%) and their pedagogy skills (30%). Teachers scoring below the passing threshold (55%) were referred to for a remedial professional development program.

We successfully matched 67% of applicants (66% of applications) to their test score data. Table A.6 describes the UR applicants according to their match status with the test score data. Matched applicants are older, and they are slightly more likely to live in Java. Because the UKG data only contains teachers that were active in 2015, it makes sense that we lack information from younger teachers.

4.2 Summary statistics

Table 1 summarizes the characteristics of the sample we use in our RCT design. Column (1) shows statistics for the whole sample, while Columns (2) and (3) present means by treatment status assignment. In addition, Column (4) shows the difference in means between the treatment and the control groups after netting out vocational sector (strata) fixed effects.

The typical teacher in our sample has a bachelor's degree, is employed full-time, has taught for more than 9 years and earns US$242 per month. Our sample is spread across 32 provinces and has good geographical representation: 51% and 24% are located in Java and Sumatra, which roughly corresponds to the shares of the Indonesian population in these islands. These teachers teach classes in multiple grades. Unsurprisingly, given the trainings we used in the randomization, they heavily teach Programming and ICT subjects (81%) and entrepreneurship (13%).

Columns (2) and (3) split the sample by treatment status. Overall, the treatment and control groups are well-balanced across most characteristics. Once we account for randomization strata, all but three differences are statistically significant at the 10% level. Treated teachers are 7 p.p. less likely to teach grade 12. Although the differences in the number of classes taught and the average number of students per class are also significant, they are fairly small: 0.2 classes and 0.3 students, respectively.

In addition, Table A.9 summarizes the characteristics of the UR trainings, as reported by the all survey respondents.[15] The average UR training lasted for 6.5 weeks and was provided by trainers from the private sector (90%). Other than activities such as lectures and discussions, participants also reported direct skill-building activities such as practice sessions with computers (60%), and internship at the industry partner (18%), as well as a pedagogical component in the form of teaching practices (28%). Teachers also reported receiving materials from the training (82%) and nearly half received lesson plans/videos. The most cited benefits from the trainings were certification from the industry (64%) and increases in knowledge and skills (59%). The top two training components rated as most helpful were on-the-job training (27%) and the training material (25%). After the conclusion of the training, 86% of the teachers reported they incorporated training content in their

[15]This table includes characteristics of all trainings, not just those in the RCT design.

day-to-day teaching and 77% reported sharing training materials with other teachers in their school. Three in five participants reported having knowledge of some of the training materials prior to the training, and one in five participants knowing all materials prior to the training, suggesting room for improvements in both the selection process and the training syllabus. Overall, teachers report high rate of satisfaction with an 8.8 average score on a scale of 10.

4.2.1 Who applied to UR?

Upon the launch of the UR program, the Ministry invited schools and teachers to apply for the training. The eligibility criteria advertised in the MoEC-issued guidelines, official letters, and YouTube live-stream launch were by no means restrictive. Schools needed to have at least two people teaching in the vocational sector, and they needed to be able to guarantee that students' learning will continue while the selected teacher participates in the UR training. The requirements for teachers were similarly broad. Participating teachers needed to be registered in the MoEC database, have at least a college degree, be no older than 50, be teaching a vocational subject in their schools, and be willing to apply the training materials in their original schools upon completion. The first two of these are basic requirements that any teacher must meet to be allowed to teach at a school.

Teachers and schools responded enthusiastically. In total, nearly 40% of all vocational schools in the country had at least one teacher submitting applications. Teachers could apply to several training subjects, and the average participating school submitted 3.8 applications from 3.3 teachers. These teachers represent 17% of all the teachers teaching vocational subjects nationwide.

Table 2 describes further the teachers who applied to UR. There we use applicant data along with the data from the universe of active teachers in 2015 to regress a dummy equal to one if the teacher is a UR applicant on a series of individual and school characteristics. All the regressions include province and vocational field fixed effects. In addition, column (3) controls for the teachers' alma mater.

Teachers in public schools and teachers who are permanent staff are more likely to apply to UR training. Public vocational schools, which are typically larger, are more likely to have a sufficiently large teaching force to minimize interruptions to regular teaching while a colleague participates in a several-weeks-long training program.

Applicants are also positively selected based on ability. They scored higher in their vocational subject tests in the 2015 competency examination and in the pedagogy test as well. UR training applicants also tend to be younger (with less teaching experience). Female teachers are less likely to apply.

5 Upskilling and Reskilling evaluation

We evaluate the effects of the UR on teachers' outcomes and expectations by comparing the results of individuals assigned to treatment and control as follows:

$$Y_i = \beta Treated_i + X_i\gamma + \delta_f + \varepsilon_i \tag{1}$$

where Y_i denotes the outcome of interest, $Treated_i$ is an assigned-to-treatment dummy, and X_i denotes additional controls that might be included. Because we stratified the randomization, we include field fixed effects δ_f in the specification.

In anticipation of the analysis, we registered a pre-analysis plan at the launch of the phone survey, prior to the completion of the survey data. Our pre-analysis plan is registered at the 3ie's Registry for International Development Impact Evaluations (RIDIE) platform, which allows the registration of studies using randomized evaluation and quasi-experimental designs.[16] The primary outcomes were teachers' vocational knowledge, classroom outcomes, and teachers' expectations of their students' labor market outcomes, as described in Section 4.1.1.

5.1 UR's impact

In Tables 3-5, we present our main estimations of the effects of UR on vocational knowledge, classroom practices, and teachers' expectations of their students' success. In Table 3, Columns (1)-(4) show the ITT estimates, and columns (5)-(8) show the 2SLS estimates. Columns (1) and (5) of Table 3 present the effects on vocational field knowledge. These estimates come from regressions where the dependent variable is the share of correct answers on a series of true and false statements about specific knowledge in six vocational sectors. Column (1) shows an ITT estimate that is very close to and statistically indistinguishable from zero. This estimate is precise enough to rule out positive effects larger than 3 percentage points (effectively zero questions) at the upperbound of the 95% confidence interval. With regard to the control group's standard deviation (0.14), this confidence interval means we can rule out any positive effect of larger than 0.18 standard deviations (SD). Column (5) shows 2SLS estimates of the local average treatment effect, also a small effect that is indistinguishable from zero. Overall, there is little evidence that training improved teachers' vocational knowledge.

We also study the effects of UR training on teachers' teaching practices. UR placed significant emphasis on day-to-day exposure to employers and hands-on learning. The Ministry thought that this could influence teachers to focus on hands-on learning in their classroom teaching (Ditjen Vokasi, 2021). However, Columns (2)-(4) and (6)-(8) in Table 3 indicate that the training has no discernible effect on teachers' distribution of time in the classroom. They spent as much time on lectures, independent student work, and discussion in class as the control group. The estimated

[16] Study ID: RIDIE-STUDY-ID-619b30d3ad31d, accessible at https://ridie.3ieimpact.org/

ITT coefficients are small and close to zero.[17] The 2SLS estimates suggest a potential substitution of lecture time for independent pupil work, but the estimates are imprecise and statistically insignificant.

Table 4 shows a similar picture for the use of information technology (IT) in the classroom. There are no significant effects for use of IT in lectures to cover material, in discussion, and for pupil to do independent work. The estimates in the table suggest that treated teachers could be more likely to use IT for discussion but, overall, our findings reflect that there is little systematic change in the way teachers taught after participating in UR. In line with the preceding findings, the 2SLS estimates lack precision, making it challenging to draw meaningful inferences.

For our final set of outcomes of interest, Table 5 shows results on teachers' expectations of their students' success. This table shows a picture that aligns well with the lack of treatment effects discussed earlier. We find null ITT estimates on the expected share of students employed after graduation (Column (1)) and the expected share of students continuing to university (Column (3)). The estimate for teachers' expectations of their graduates' employment allows us to rule out effects larger than 0.12 standard deviations. For expected students' average wages, Column (2) shows insignificant negative effects which also allows us to rule out wage increases of $2.87 per month or higher. Estimates from 2SLS in Columns (5)-(7) suggest a *negative* effect on expected wages and increases in teachers' optimism about their students.

In contrast to outcomes discussed above, teachers appear to shift their subjective expectations of student readiness for the labor market. Teachers are asked to indicate their agreement with the statement that their graduates are industry-ready. In both ITT (Column (4)) and 2SLS estimates (Column (8)), treated teachers are more likely to express strong agreement with the statement. It is thus possible that while the training did not move teachers' skills substantively, the training did update their beliefs on whether their students have the skills demanded by the private sector.

We also examine whether the program had differential effects depending on the teachers' characteristics. In Tables 6, we consider two main outcomes: teacher knowledge and graduates' employment expectations. We consider the following teacher characteristics: having a master's degree or higher, having teaching experience above the median (12 years), being a permanent employee, having MoEC certification, being a male teacher, and being from Java/Bali. We find there is no consistent pattern suggesting that the training is particularly effective for a subgroup of teachers. For the estimates in vocational knowledge in Panel A, the coefficients for treatment indicators and their interaction with the group indicator are fairly small, suggesting that any possible effects will be smaller than 1 question. Estimates for heterogeneity analysis for graduates' employment expectations along the above characteristics also show inconsistent patterns as shown in Panels B and C.

We contrast the estimates from our randomized evaluation against an alternative estimation strategy using propensity scores with a larger sample. Overall, in Appendix Table A.13-A.14, we find very similar patterns the PSM sample. The larger sample allows for a more precise estimation of

[17]Tables A.10 to A.12 show results from regressions analogous to those in Tables 3 to 5, but without covariates.

the null effects. In comparison with the control group's outcome distribution, the upper bounds of the 95% CI around the estimate indicate we can rule out effects on the share of classroom time that are larger than 0.03 SD for lectures, 0.15 SD for discussion, and 0.20 SD for independent pupil work. The main difference with the RCT results, is that Table A.13 suggests small and significant increases in the use of ICT to cover material and discussion, and a decline in the expected share going to university.

6 Why was UR ineffective?

We have documented that the program did not lead to a meaningful improvements for the treated teachers vis-a-vis the comparison group. In this section we explore three possible reasons for this finding.

6.1 Mismatch between teachers' needs and training offerings

For a teacher training program to be effective in improving student outcomes, it needs to improve the teachers' knowledge and classrooms practices. If teachers are not at all familiar with the curriculum, then training programs with more basic contents or refresher courses targeted at specific knowledge and skill gaps could be a first step to improve teachers' effectiveness. In other words, the theory behind effective Teaching at the Right Level interventions could be extrapolated to the teaching force to address teachers' skill gap.

MoEC is aware of a general skill gap among vocational teachers and sees the high unemployment rate among graduates as a symptom of quality issues in vocational high schools. However, addressing the skill gap with an at-scale program for such a diverse educational system is challenging. Teachers' responses to our survey, along with reports from the training, give clues as to why training mismatch may have contributed to the lack of impact of the program.

MoEC selection guidelines were broad and not targeted. Moreover, none of the UR attendees we surveyed perceived that the selection processes were targeted to address possible skill gaps. 47% reported that their school was selected because of the vocational sector they offered, 25% reported that there were no particular selection criteria, and 20% reported that selection was based on who submitted an application to MoEC's portal. Any targeting within school, if any, was coarse and did not go beyond their teaching duties: 72% teachers reported being selected for training (or being selected by the school to submit an online application) based on what subjects they taught at school. 28% of teachers reported that there is no within-school selection.

Moreover, teachers were also likely to report that they were already familiar with the materials delivered during the training. Eighty percent of the attendees reported that they had taught the materials to the students prior to the training. They also reported some degree of proficiency prior to the training: more than three-fifths of attendees reported that they were already proficient in some of the contents, and a full one-fifth reported proficiency in all of the contents. Respondents in our sample were also able to identify topics from the training that they had been teaching to their students. For example, multiple attendees in the Java programming training mentioned 'foundational Java', and attendees in the fiber optics training mentioned 'DHCP server' as materials they regularly had been teaching at school. These materials may be the same materials that UR training had covered. These findings could also help explain why we see an increase in teachers' optimism about their students without any meaningful change in their subject knowledge or teaching practices. If teachers received content they were already familiar with from the training providers, they could have interpreted this as evidence that they were already teaching skills demanded in the private sector.

6.2 Lack of post-training support

Centrally organized training brings teachers away from their daily environments where they deliver the curriculum to the students. For teachers to be able to apply the training that they received, they may require further support after the training was concluded. This may include following up with teachers or obtaining authorization from the school principal to incorporate the materials that they received from the training into their day-to-day teaching.

However, providing training follow-up remains a best practice that is rarely implemented. More than half (55%) of the teachers in our sample who attended UR reported that they needed further support to be able to incorporate training into classroom practices. At the same time, only a minority of teachers recalled any follow-up sessions from the training. The overwhelming majority (74%) did not receive any follow-up. The lack of post-training support has been argued as one of the explanations for the lack of impact for an at-scale training program for middle-grade teachers in Nepal (Schaffner et al., 2024). In comparison, Popova et al. (2022) noted that 85% of the top performing programs in their data include follow-up visits, while only 49% of the at-scale programs they analyzed include a follow-up visit.

Furthermore, teachers who found the training useful may also have to navigate negotiations with school principals. Among teachers attending the training, 53% reported that they would have needed management support from their principals to incorporate the materials from the training into their classroom practices. Slightly less than half of the teachers (47%) reported that they took steps to coordinate with their school principals. Without buy-in from the principals, this may have led to the lack of meaningful changes in teachers' practices in the classroom.

Hands-on training with industry may also reveal the infrastructure gap between industry and the teachers' vocational schools. Teachers may gain access to specialized equipments and softwares during the training, but the same facilities may not be available to the teachers to use in the classroom. Accordingly, 60% of teachers said that they needed specialized equipments to incorporate materials from the training, while also highlighting students' need to access computers (55%), specialized software (41%), and internet access (48%). In a setting where only 70% of students have sufficient internet access, adaptation becomes challenging.

6.3 Outside training provided alternative to untreated teachers

Measuring the impact of the training required constructing a counterfactual group to obtain an unbiased estimate. The counterfactual allows us to learn what would have happened to the trained individuals in absence of the training from the program. A training program could be successful in increasing the participants' skill if they do not have access to a comparable training in the counterfactual.

In our sample, we see that while the UR training was successful in increasing teachers' exposure to the industry, there is no significant difference in access to any training between treated and control teachers. In Table 7, we report coefficients from regressions of various training attendance indicators for both the ITT and 2SLS specifications. In Columns (1)-(2), our data shows that random assignment to the treatment group increased the likelihood that the treated group attended the training, and in Column (3) this training increased the exposure of teachers to the industry. However, there is no significant difference between the treated and control group with regard to whether they participated in any training in the past years, the number of trainings, and the hours that they were engaged in training from (Columns (4)-(6)). Column (7) also shows that there is no difference in the number of training follow-ups they received as well. Descriptive lists of trainings that non-UR participating teachers provide during our phone surveys reveal that various institutions offered trainings to teachers beyond the UR scheme. Among our respondents, 63 respondents report that they received training on utilization of e-learning platforms/Covid distance learning adjustments and 26 respondents listed trainings that are specific to a vocational sector as well. Examples from the latter group include training in Python programming language, IP address rooting, CSS and Javascript for web programming, CAD, welding, and machinery techniques. In this light, it may not be required for MoEC to implement the training on their own, so long as they provide clear guidelines for the private sector to collaborate with vocational high schools in a bid to improve the quality of the vocational education system in the country.

7 Conclusion

As education policymakers in developing countries prioritize vocational education, improving its effectiveness holds great potential to create meaningful impact for their students (Crawfurd et al., 2021). Teacher professional development programs that bring vocational teachers closer to the industry has a strong theoretical appeal, but challenges remain to implement an effective PD program at scale. This study finds that an at-scale intensive teacher PD program for vocational teachers in Indonesia did not have any significant impacts on teacher knowledge, teaching practices, and expectations of their graduates. Our evaluation adds to the PD literature that finds little impact of at-scale programs when they are being rigorously evaluated (Loyalka et al., 2019; Popova et al., 2022; Schaffner et al., 2024).

Our study makes three major contributions to the PD literature. First, we provide the first rigorous evidence of a program to improve teaching quality in upper secondary vocational schools. Second, we evaluated an at-scale PD program which serves as the umbrella program to train teachers in dozens of diverse vocational streams. Third, our analysis is based on an evaluation of an intensive program (260 hours) and subject-specific training, both of which are features of PD programs that are deemed promising to boost student learning outcomes.

Our evaluation offers valuable lessons from Indonesia to other policymakers interested in designing their teachers' professional development programs. Participating teachers' survey responses highlights the importance of needs assessment, which may help align interventions to target existing skill gaps better. While our findings are rooted in the specific context in which this program was implemented, our evaluation offers a rare case through which other policymakers wanting to improve their vocational education systems can build upon.

References

Abeberese, A. B., Kumler, T. J., and Linden, L. L. (2014). Improving reading skills by encouraging children to read in school:: A randomized evaluation of the sa aklat sisikat reading program in the philippines. *Journal of Human Resources*, 49(3):611–633.

Al-Ubaydli, O., List, J. A., and Suskind, D. (2019). The science of using science: Towards an understanding of the threats to scaling experiments. Technical report, National Bureau of Economic Research.

Albornoz, F., Anauati, M. V., Furman, M., Luzuriaga, M., Podesta, M. E., and Taylor, I. (2020). Training to teach science: experimental evidence from argentina. *The World Bank Economic Review*, 34(2):393–417.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling youth unemployment: Evidence from a labor market experiment in uganda. *Econometrica*, 88(6):2369–2414.

Angrist, N. and Meager, R. (2023). Implementation matters: Generalizing treatment effects in education. *SSRN 4487496*.

Atkin, D., Donaldson, D., Rasul, I., Teachout, M., Verhoogen, E., and Woodruff, C. (2019). Firms, trade, and productivity. Technical report, International Growth Centre, London.

Attanasio, O., GuarÍn, A., Medina, C., and Meghir, C. (2017). Vocational training for disadvantaged youth in colombia: A long-term follow-up. *American Economic Journal: Applied Economics*, 9(2):131–143.

Attanasio, O., Kugler, A., and Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 3(3):188–220.

Banerjee, A., Glewwe, P., Powers, S., and Wasserman, M. (2013). *Expanding Access and Increasing Student Learning in Post-Primary Education in Developing Countries: A Review of the Evidence*. Abdul Latif Jameel Poverty Action Lab, Cambridge, MA.

Bazzi, S., Hilmy, M., and Marx, B. (2020). Religion, education, and the state. Working Paper 27073, National Bureau of Economic Research.

BBPPMPV BOE (2020). Panduan pelaksanaan diklat program upskilling dan reskilling guru kejuruan pada smk berstandar industri tahun 2020. Technical report, Balai Besar Pengembangan Penjaminan Mutu Pendidikan Vokasi Bidang Otomotif dan Elektronika Direktorat Jenderal Pendidikan Vokasi Kementerian Pendidikan dan Kebudayaan.

Beg, S., Halim, W., Lucas, A. M., and Saif, U. (2022). Engaging teachers with technology increased achievement, bypassing teachers did not. *American Economic Journal: Economic Policy*, 14(2):61–90.

Berlinski, S. and Busso, M. (2017). Challenges in educational reform: An experiment on active learning in mathematics. *Economics Letters*, 156:172–175.

Blimpo, M. P. and Pugatch, T. (2021). Entrepreneurship education and teacher training in rwanda. *Journal of Development Economics*, 149:102583.

Bruns, B., De Gregorio, S., and Taut, S. (2016). Measures of effective teaching in developing countries. Technical report, RISE Working Paper 16/009.

Carneiro, P., Cruz-Aguayo, Y., Intriago, R., Ponce, J., Schady, N., and Schodt, S. (2022). When promising interventions fail: Personalized coaching for teachers in a middle-income country. *Journal of Public Economics Plus*, 3:100012.

Central Bureau of Statistics, Indonesia (2016). National labor force survey (sakernas).

Cilliers, J., Fleisch, B., Prinsloo, C., and Taylor, S. (2020). How to improve teaching practice?: an experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*, 55(3):926–962.

Crawfurd, L., Hares, S., Minardi, A., and Sandefur, J. (2021). *Understanding Education Policy Preferences: Survey Experiments with Policymakers in 35 Developing Countries*. Center for Global Development.

Ditjen Vokasi (2021). Laporan kinerja direktorat jenderal vokasi tahun 2020. Technical report, Direktorat Jenderal Pendidikan Vokasi Kementerian Pendidikan dan Kebudayaan.

EdStats (2022). Education statistics (https://datatopics.worldbank.org/education).

Evans, D. K. and Popova, A. (2016). What really works to improve learning in developing countries? an analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, 31(2):242–270.

Field, E. M., Linden, L. L., Malamud, O., Rubenson, D., and Wang, S.-Y. (2019). Does vocational education work? evidence from a randomized experiment in mongolia. Technical Report w26092, National Bureau of Economic Research.

Fryer, R. G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of economic field experiments*, volume 2, pages 95–322. Elsevier.

Ganimian, A. J. (2020). Growth-mindset interventions at scale: Experimental evidence from argentina. *Educational Evaluation and Policy Analysis*, 42(3):417–438.

Ganimian, A. J. and Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, 86(3):719–755.

Government of Indonesia (2016). Instruksi presiden (inpres) nomor 9 tahun 2016 tentang revitalisasi sekolah menengah kejuruan dalam rangka peningkatan kualitas dan daya saing sumber daya manusia indonesia. Legislation.

Hicks, J. H., Kremer, M., Mbiti, I., and Miguel, E. (2011). Vocational education voucher delivery and labor market returns: A randomized evaluation among kenyan youth. *Report for Spanish Impact Evaluation Fund (SIEF) Phase II.*

Jukes, M., Vagh, S. B., and Kim, Y.-S. (2013). Development of assessments of reading ability and classroom behavior - a report prepared for the world bank. Technical report, World Bank.

Kerwin, J. T. and Thornton, R. L. (2021). Making the grade: The sensitivity of education program effectiveness to input choices and outcome measures. *Review of Economics and Statistics*, 103(2):251–264.

Loyalka, P., Popova, A., Li, G., and Shi, Z. (2019). Does teacher training actually work? evidence from a large-scale randomized evaluation of a national teacher training program. *American Economic Journal: Applied Economics*, 11(3):128–154.

Majerowicz, S. and Montero, R. (2018). Can teaching be taught? experimental evidence from a teacher coaching program in peru. Job Market Paper.

Menteri Pendidikan dan Kebudayaan (2015). Peraturan menteri pendidikan dan kebudayaan republik indonesia nomor 43 tahun 2015 tentang uji kompetensi guru atau pendidik lainnya dan tenaga kependidikan. 43.

Null, C., Cosentino, C., Sridharan, S., and Meyer, L. (2017). Policies and programs to improve secondary education in developing countries. *Mathematica Policy Research.*

Popova, A., Evans, D. K., Breeding, M. E., and Arancibia, V. (2022). Teacher professional development around the world: The gap between evidence and practice. *The World Bank Research Observer*, 37(1):107–136.

Pritadrajati, D. (2018). From school to work: Does vocational education improve labour market outcomes? an empirical analysis of indonesia. Technical report, UNESCAP.

Schaffner, J., Glewwe, P., and Sharma, U. (2024). Why programs fail: Lessons for improving public service quality from a mixed-methods evaluation of an unsuccessful teacher training program in nepal. *The World Bank Economic Review*, 0(0):1–24.

Setditjen Vokasi (2020). Laporan akhir riset evaluasi pelaksanaan program pendidikan vokasi tahun 2020. Technical report, Sekretariat Direktorat Jenderal Pendidikan Vokasi Kementerian Pendidikan dan Kebudayaan.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher*, 9(11):11–16.

World Bank (2015). Conducting classroom observations: Analyzing classroom dynamics and instructional time using the stallings "classroom snapshot" observation system. Technical report.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., and Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement. issues & answers. rel 2007-no. 033. *Regional Educational Laboratory Southwest (NJ1).*

Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M., Weiland, C., Gomez, C. J., Moreno, L., Rolla, A., D'Sa, N., et al. (2015). Experimental impacts of a teacher professional development program in chile on preschool classroom quality and child outcomes. *Developmental psychology*, 51(3):309.

Zaw, H. T., Mizunoya, S., Richardson, D., Karamperidou, D., Hattori, H., and Oledzka-Nielsen, M. (2021). Teacher training and textbook distribution improve early grade reading: evidence from papua and west papua. *Comparative Education Review*, 65(4):691–722.

Table 1: Summary statistics by RCT treatment status

	All (1)	**Control** (2)	**Treated** (3)	**Difference** (4)
Male	0.674	0.714	0.630	-0.040
				(0.031)
Age (years)	36.263	36.153	36.382	-0.364
				(1.039)
Has a bachelor's degree or higher	0.995	1.000	0.990	-0.014
				(0.008)
Has a master's degree or higher	0.147	0.138	0.156	0.022
				(0.021)
Civil servant	0.381	0.395	0.365	-0.071
				(0.043)
				(0.072)
Teaching experience (years)	9.793	9.700	9.895	-0.265
				(0.642)
Salary (USD)	242.786	241.040	244.752	-32.916
				(20.151)
Java-Bali	0.517	0.495	0.542	-0.008
				(0.031)
Sumatra	0.244	0.233	0.255	-0.003
				(0.035)
Kalimantan and other eastern islands	0.239	0.271	0.203	0.011
				(0.032)
Teaches grade 10 at school	0.656	0.633	0.681	0.079
				(0.066)
Teaches grade 11 at school	0.768	0.743	0.796	0.067
				(0.048)
Teaches grade 12 at school	0.753	0.786	0.717	-0.085*
				(0.038)
Total classes taught at school (grade 10-12)	4.903	4.852	4.958	-0.198*
				(0.545)
Average students per class	30.287	29.881	30.733	-0.301*
				(0.881)
Vocational subject[a]				
Programming/ICT/digital subjects	0.811	0.829	0.792	
Machinery/automotive subjects	0.007	0.010	0.005	
Accounting/business subjects	0.030	0.029	0.031	
Entrepreneurship subjects	0.139	0.138	0.141	
Hospitality subjects	0.000	0.000	0.000	
Fashion subjects	0.002	0.000	0.005	
Other technical subjects	0.072	0.057	0.089	

Notes: Columns (1), (2), and (3) show means for the whole sample, for teachers assigned to the control, and for teachers assigned to treatment, respectively. Column (4) shows the difference in means after netting out vocational sector (strata) fixed effects. [a]Difference excluded as there is little variation in vocational subject taught within vocational sector. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Selection into training

	Applied (1)	**Applied** (2)	**Applied** (3)
Public school	0.032***	0.034***	0.032***
	(0.005)	(0.005)	(0.005)
Permanent staff	0.028***	0.033***	0.032***
	(0.005)	(0.005)	(0.005)
2015 Vocational subject test score	0.010***	0.012***	0.010***
	(0.002)	(0.003)	(0.003)
2015 Pedagogy test score	0.008***	0.006***	0.005***
	(0.002)	(0.002)	(0.002)
Age	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.000)
Teaching certification	-0.009***	-0.006*	-0.005*
	(0.003)	(0.003)	(0.003)
Female	-0.013***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)
Tenure length		-0.002***	-0.002***
		(0.000)	(0.000)
Constant	0.182***	0.174***	0.158***
	(0.022)	(0.020)	(0.016)
Province FE	✓	✓	✓
Vocational sector FE	✓	✓	✓
University FE			✓
Observations	213,884	154,890	153,766

Notes: Table presents coefficients from multivariate OLS of SMK application dummy on pre-treatment characteristics. The sample is restricted to vocational high school teachers who took the 2015 competency test. All regressions include province and vocational sector fixed effects. Standard errors clustered at the vocational subject group in parentheses. * p$\leq$ 0.10, ** p$\leq$ 0.05, *** p$\leq$ 0.01.

Table 3: Effect of training on teachers' vocational knowledge and classroom practices

	ITT				2SLS			
	Vocational test	**Lectures**	**Independent work**	**Discussion**	**Vocational test**	**Lectures**	**Independent work**	**Discussion**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-0.004	-0.027	0.026	0.011	0.017	-0.185	0.177	0.076
	(0.015)	(0.040)	(0.039)	(0.015)	(0.069)	(0.284)	(0.277)	(0.067)
Dep. Var. Mean	0.719	0.465	0.195	0.169	0.719	0.465	0.195	0.169
Dep. Var. SD	0.144	0.236	0.239	0.168	0.144	0.236	0.239	0.168
Observations	395	318	318	318	395	318	318	318

Notes: The table shows coefficients from regressions of the outcome variables in column titles. Columns (1) to (4) show ITT (intention to treat) estimates from OLS regression on an indicator of being assigned to the treatment group. Columns (5) to (8) show 2SLS estimates regression on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. Columns (2) to (4) and (6) to (8) use as dependent variable the share of classroom time dedicated to the activity in column title. All regressions include the following covariates: gender, years of education, years of teaching experience, indicator of being civil servant or a full time staff, indicator of certification status, an array of province dummies, and an array of vocational sector dummies. Clustered standard errors at the vocation sector level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect of training on teachers' use of ICT

	ITT			**2SLS**		
	Cover material	**Discussion**	**Pupil work**	**Cover material**	**Discussion**	**Pupil work**
	(1)	(2)	(3)	(4)	(5)	(6)
Training	-0.032	0.063	0.013	-0.161	0.311	0.063
	(0.034)	(0.049)	(0.046)	(0.199)	(0.275)	(0.178)
Dep. Var. Mean	0.739	0.359	0.423	0.739	0.359	0.423
Dep. Var. SD	0.440	0.480	0.495	0.440	0.480	0.495
Observations	395	395	395	395	395	395

Notes: The table shows coefficients of regressions of share of ICT used for the classroom activities indicated in the column titles. Columns (1) to (3) show ITT (intention to treat) estimates from OLS regressions on an indicator of being assigned to the treatment group. Columns (4) to (6) show 2SLS estimates from regressions on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. All regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and an array of vocational sector dummies. Clustered standard errors at the vocation sector level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effect of training on teachers' expectations of their students' outcomes

	ITT				2SLS			
	Employed	**Wage (USD)**	**Going to university**	**Industry-ready**	**Employed**	**Wage (USD)**	**Going to university**	**Industry-ready**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	0.007	-6.903	0.010	0.055**	0.055	-47.589**	0.071	0.313***
	(0.009)	(4.987)	(0.012)	(0.017)	(0.059)	(23.804)	(0.093)	(0.121)
Dep. Var. Mean	0.351	137.161	0.298	0.107	0.351	137.161	0.298	0.107
Dep. Var. SD	0.208	71.895	0.206	0.309	0.208	71.895	0.206	0.309
Observations	326	314	350	394	326	314	350	394

Notes: Coefficients from regressions of outcome variables in various estimation samples. Column 1 shows ITT (intention to treat) estimates from OLS regression of outcome on an indicator of being assigned to the treatment group in the RCT sample. Column 2 shows 2SLS estimates regression of outcome on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. Column 3 shows estimates from regression of outcome on an indicator of being assigned to the treatment group in the PSM sample. All regressions include the following covariates: gender, years of education, years of teaching experience, indicator of being civil servant or a full time staff, indicator of certification status, an array of province dummies, and an array of vocational sector dummies. The PSM regression includes matching group dummies. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Heterogeneity in treatment effects by teacher's characteristics

	Master's degree	**Above-median tenure**	**Permanent employee**	**Certified**	**Male**	**In Java/Bali**
	(1)	(2)	(3)	(4)	(5)	(6)
A. Outcome: Vocational knowledge test						
Training	-0.003	-0.013	-0.015	-0.014	0.003	-0.039
	(0.013)	(0.014)	(0.024)	(0.015)	(0.023)	(0.027)
Training × group	-0.001	0.049	0.017	0.035*	-0.011	0.062*
	(0.056)	(0.031)	(0.039)	(0.015)	(0.042)	(0.027)
Dep. Var. Mean	0.719	0.719	0.719	0.719	0.719	0.719
Dep. Var. SD	0.144	0.144	0.144	0.144	0.144	0.144
Observations	395	395	395	395	395	395
B. Outcome: Share of graduates employed 3 months after graduation						
Training	0.008	0.008	0.033**	0.010	0.039	-0.003
	(0.011)	(0.010)	(0.011)	(0.008)	(0.022)	(0.018)
Training × group	-0.011	-0.004	-0.041	-0.012	-0.051	0.016
	(0.052)	(0.027)	(0.023)	(0.018)	(0.031)	(0.018)
Dep. Var. Mean	0.351	0.351	0.351	0.351	0.351	0.351
Dep. Var. SD	0.208	0.208	0.208	0.208	0.208	0.208
Observations	326	326	326	326	326	326
C. Outcome: Average monthly salary in first job (USD)						
Training	-7.929	-11.261**	-8.325	-7.431	-1.745	-7.681
	(5.185)	(4.211)	(9.703)	(5.459)	(4.231)	(5.538)
Training × group	7.312	19.225***	2.302	2.005	-7.714**	1.322
	(8.085)	(3.319)	(8.070)	(4.317)	(2.503)	(6.840)
Dep. Var. Mean	137.161	137.161	137.161	137.161	137.161	137.161
Dep. Var. SD	71.895	71.895	71.895	71.895	71.895	71.895
Observations	314	314	314	314	314	314

Notes: Each column shows estimates for the coefficients of an attended-training indicator, and the interaction between the training indicator and a group indicator. The group indicator is equal to one for the group indicated the column title. Besides the shown variables, all regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and an array of vocational sector dummies. Clustered standard errors at the vocational sector level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effect on training participation

	(1) **Attended training MOEC record**	(2) **Attended training self-reported**	(3) **Trained by industry**	(4) **Had any training past year**	(5) **Number of trainings past year**	(6) **Hours in training**	(7) **Training follow-ups**
A. RCT ITT							
Treatment	0.185***	0.108*	0.107*	0.004	0.306	-5.109	-0.430
	(0.053)	(0.056)	(0.057)	(0.048)	(0.411)	(11.747)	(0.627)
Constant	-0.309	-0.250	-0.325	1.266**	1.479	80.201	-8.313
	(0.569)	(0.599)	(0.609)	(0.516)	(4.383)	(125.298)	(6.683)
B. RCT - 2SLS							
Attended		0.586***	0.575**	0.019	1.656	-27.684	-2.330
		(0.204)	(0.230)	(0.246)	(2.128)	(62.263)	(3.315)
Dep. Var. Mean	0.40	0.41	0.45	0.78	2.89	69.66	0.60
Dep. Var. SD	0.49	0.49	0.50	0.41	3.48	100.05	5.10
Observations	401	401	397	401	401	401	401

Notes: Coefficients from regressions of outcome variables in various estimation samples. All regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and an array of vocational sector dummies. Clustered standard errors at the vocational level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendices

Appendix A Tables

Table A.1: Examples of Upskilling and Reskilling courses

Target industry	Course name	Provider
Construction	Furniture and building wood finishing	Bojong Westplas LLC
	Gypsum board finishing	Petrojaya Boral Plasterboard LLC
Healthcare	Toddler care industry process	Koba Mirai
	Industrial processes to support the elderly	Koba Mirai
Hospitality	Meat processing training	Badranaya, Bandung
	Organic vegetable processing training	Prosperous Agro Mandiri, Banjarnegara
	Preparation of guest rooms / porter services	Royal Hotel Padjadjaran Bogor, The Mirah Hotel Bogor
	Manufacturing of Indonesian food	IPA, MGM Horizon
Repair of vehicles and machinery	Operation and maintenance of pneumatic equipment and systems	Festus
	Programming and operation of computer numerical control machines (CNC)	Siemens LLC
	Light vehicle chassis maintenance	Toyota Delta More Medan
	Engine management system	Toyota Delta More Medan
Software and design services	3D animation	Kayon Tunggal Ika LLC
	Computer graphics	Publisher Diamond Pariwara
	Publication design	Creative Media E&E Studio
	Photography	CPC Photo Design
	Creative digital marketing with adobe creative cloud	Adobe
	Blouse manufacturing process	Creative Creative Ditali LLC
	Network and communications	MicroTik

Notes: course names translated by the authors.

Table A.2: Schedule and training partners for selected UR courses

UR 2020 Courses	OJT training partners	Schedule and time allocation		
A. BBPPMPV BOE Batch 1				
Programming and operation of CNC machines	PT. INKA, PT.PAL		Hours/Days	Dates
Maintenance of injection system light vehicles	Auto 2000 Malang and Surabaya; Nissan Surabaya; Nasmoco Solo; Borobudur Otomobil Yogyakarta	Online training	65/12	21 Sep-3 Okt
Topographic mapping	PT. Rasicipta Consultama, PT. Onward Bladgoud	In-person training	102/12	19-31 Okt
Gypsum board finishing	PT. Petrojaya Boral Plasterboard	On-the-job training	60/6	02-07 Nov
Smartphone optimization	Service Center Samsung Electronics Indonesia	Certification	40/4	09-12 Nov
		total	267/34	
B. BBPPMPV BOE Batch 3				
Steel plates welding techniques	PT. INKA, PT. PAL		Hours/Days	Dates
Light vehicle spooring and suspension system	PT Malintra	Online training	65/12	19-31 Okt
Using 2D AutoCAD	PT. Tiga Dinamika Solusi Indonesia	In-person training	102/12	16-28 Nov
Furniture and building wood finishing	PT. Natania Furniture Singosari, PT Propan Raya Industrial Coating Chemical	On-the-job training	60/6	30 Nov-05 Dec
Building plumbing installations	PT. Bojong Wesplas	Certification	40/4	07-10 Dec
		total	267/34	

Note: This table combines and adapts information from BBPPMPV BOE (2020), page 5 (OJT partner), page 11 (time allocation), and page 26 (schedule).

Table A.3: Curriculum for Network and Communication UR training

Content	Training Modules (Contact Hours)
Policies	MoEC policy (2), Upskilling and Reskilling policy (2)
MTCNA	Introduction (2), DHCP Server and Client, ARP (2), Bridging, Wireless Bridging (4), Foundations of Routing (4), Wireless (4), Firewall (4), Quality of Service (4)
	Mikrotik Certified Network Associate (MTCNA) Certification test (3)
MTCRE	Static Routing (5), Point to Point Address (2), VPN (2), Open Shortest Path First (9)
	MikroTik Certified Routing Engineer (MTCRE) certification test (3)
Fiber optics	Intro to Fiber Optic (3), Fiber Optic Network Design FTTX FFTH (5), Fiber Optics Cable Installations, Optical Distribution Panel Adapter, Optical Terminal Box (4), Fusion Splicer and Mechanical Fiber Optics Termination (8), Damping Measurement (2), Fiber Optic Cables Implementation for Internet access using Mikrotik and SFP Module (2), Troubleshooting (3)
	Fiber Optics test (5)

Notes: This table presents the curriculum delivered during the offline training organized by BBPPMPV KPTK in two batches. Batch 1 took place 26 Oct-06 Nov 2020 at BBPPMPV KPTK building, and Batch 2 took place between 9-20 November in Hotel Gammara Makassar. The curriculum table is taken from *'Laporan Singkat Program Upskilling dan Reskilling Guru Kejuruan SMK Kompetensi Keahlian Teknik Komputer Jaringan "Mikrotik dan Fiber Optik"'*, a document issued by BPPMPV KPTK Direktorat MitrasDUDI Dirjen Pendidikan Vokasi Kemdikbud. The above table is adapted from the table on page 2 of the report. Nineteen out of twenty participants Batch 1 scored 89.9 in the final exam, while one scored 83.95. Nineteen out of 20 participants in Batch 2 scored 94.1, while one scored 90.53. Participants' post-training scores are summarized in tables on pages 7 and 11 of the report.

Table A.4: Distribution of applicants, budgeted slots available, and number of actual attendees by course

Sector code	Training name	(1) Total applicants	(2) Budgeted slots	(3) Attendees
A. Trainings in RCT design				
TKJ 1	Mikrotik and fiber optics	1355	20	60
RPL 2	Java programming	357	150	206
RPL 4	Database management	280	150	146
GEO 1	Topographic mapping	87	40	33
SIJ 2	Internet of things	79	20	15
ANI 3	2D animation	77	25	16
B. Trainings in PSM design				
AK 1	Accounting processing training	1575	80	75
OTO 2	Maintenance of injection system light vehicles	1383	40	96
OTKP 1	Training for administrative staff	1348	80	76
ANI 1	Creative digital marketing with adobe creative cloud	1341	100	39
OTO 3	Automotive mechanic junior, light vehicle chassis maintenance	595	60	57
HOT 1	Industrial process preparation of guest rooms	586	80	66
SAM 1	Smartphone optimization and smartphone troubleshooting	579	40	47
RPL 6	Android programming	571	28	67
OTO 1	Light vehicle suspension and spooring balancing system	532	40	53
CNC 1	CNC machine programming and operation	528	92	99
LIS 1	Center of excellence training for vocational school teachers	525	108	9
BUS 1	Blouse manufacturing process	507	80	79
TEI 1	Operation and maintenance of pneumatic equipment and systems	451	40	33
BOG 1	Continental and oriental food manufacturing industry processes	445	80	70
AGRI 3	Fruit and vegetable processing training	422	14	18

Note: Panel A shows a list of all the sectors slotted for the RCT design. Panel B shows the list of the 15 largest classes in the PSM design by the number of applicants. Column (2) shows the number of student slots *initially* budgeted by MoEC. In some cases, these slots were expanded.

Table A.5: Summary statistics by treatment status, PSM design

	All (1)	**Control** (2)	**Treated** (3)	**Difference** (4)
Male	0.527	0.501	0.557	0.001
				(0.021)
Age (years)	39.906	40.066	39.631	-0.555*
				(0.285)
Has a bachelor's degree or higher	0.998	0.999	0.995	-0.003*
				(0.003)
Has a master's degree or higher	0.192	0.178	0.215	0.045**
				(0.021)
Civil servant	0.610	0.615	0.601	0.003**
				(0.021)
Civil servant or full time employee	0.772	0.779	0.760	-0.004**
				(0.021)
Teaching experience (years)	12.852	12.917	12.739	-0.307**
				(0.222)
Salary (USD)	318.199	318.402	317.843	13.905**
				(9.024)
Java-Bali	0.568	0.561	0.581	0.022**
				(0.023)
Sumatra	0.247	0.244	0.253	0.021**
				(0.023)
Kalimantan and other eastern islands	0.185	0.195	0.167	-0.043**
				(0.021)
Teaches grade 10 at school	0.575	0.593	0.543	-0.023**
				(0.029)
Teaches grade 11 at school	0.801	0.796	0.808	-0.003**
				(0.023)
Teaches grade 12 at school	0.837	0.846	0.821	-0.042*
				(0.021)
Total classes taught at school (grade 10-12)	4.500	4.666	4.210	-0.381***
				(0.126)
Average students per class	30.971	30.793	31.283	0.600*
				(0.330)
Programming/ICT/digital subjects	0.333	0.340	0.321	
Machinery/automotive subjects	0.071	0.069	0.076	
Accounting/business subjects	0.103	0.103	0.104	
Entrepreneurship subjects	0.135	0.145	0.119	
Hospitality subjects	0.124	0.128	0.116	
Fashion subjects	0.087	0.085	0.091	
Other technical subjects	0.210	0.211	0.210	
Observations[a]	1,397	760	637	

Notes: Column (1) shows means for the whole sample. Columns (2) and (3) show means for teachers assigned to the control and treatment groups, respectively. Column (4) shows the difference in means after netting out matched-group fixed effects. Control observations are weighted by the inverse of the number of observations in the control group. [a]Total number of observations in the PSM design. The number of observations varies from row to row depending on the number of valid responses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Summary statistics for UR applicants by match status with test score data

	Age in 2020	**In Java island**	**No. applications**	**Share of applicants**
	(1)	(2)	(3)	(4)
Unmatched	31.20	0.52	1.62	32.55
	(0.08)	(0.01)	(0.02)	
Matched	38.93	0.56	1.54	67.45
	(0.05)	(0.00)	(0.01)	
Number of applicants	18,448			

Notes: The table shows the characteristics of UR applicants my match status with the 2015 UKG data.

Table A.7: Item analysis of questions in the vocational knowledge module

	All	**GEO1**	**ANI3**	**TKJ1**	**SIJ2**	**RPL2**	**RPL4**
Q1	0.85	0.94	0.71	0.89	0.79	0.78	0.92
Q2	0.57	0.58	0.20	0.62	0.67	0.75	0.39
Q3	0.88	0.52	0.97	0.98	0.88	0.90	0.81
Q4	0.83	0.87	0.89	0.91	0.79	0.81	0.76
Q5	0.86	0.87	0.80	0.98	0.67	0.80	0.89
Q6	0.69	0.32	0.17	0.74	0.95	0.78	0.67
Q7	0.75	0.81	0.83	0.84	0.67	0.73	0.67
Q8	0.81	0.74	0.77	0.83	0.69	0.87	0.77
Q9	0.86	0.77	0.80	0.78	0.88	0.86	0.95
Q10	0.28	0.19	0.11	0.075	0.95	0.31	0.32
Answers all 'True' (N)	37	5	7	4	0	7	14
Answers all 'True' (share)	0.07	0.16	0.20	0.03	0	0.05	0.11
Mean score	0.72	0.66	0.63	0.77	0.79	0.76	0.72
Observations	506	31	35	133	42	134	131

Note: This table tabulates the share of respondents answering with the correct answer for each question in the knowledge module, disaggregated by vocational sectors.

Table A.8: Correlations between national exam scores and teachers' reported expectations of their schools' graduates

	Dep. Var.: 2019 Grade 12 National Exam				
	Average	**Math**	**Indonesian**	**English**	**Vocational**
	(1)	(2)	(3)	(4)	(5)
A. Share of 2019 graduates in employment three months after graduation					
Share employed	1.86***	1.00*	1.62**	1.68***	3.15***
	(0.45)	(0.52)	(0.61)	(0.47)	(0.65)
Dep. Var. Mean	47.70	35.99	67.06	42.47	45.27
Dep. Var. SD	7.47	7.72	8.77	8.09	7.50
R2	0.39	0.33	0.52	0.32	0.27
Observations	1442	1442	1442	1442	1442
B. Average monthly salary of 2019 graduates in their first job (USD)					
Salary (USD)	0.0047*	0.0068**	0.0058*	0.0094**	-0.0033
	(0.0026)	(0.0026)	(0.0030)	(0.0035)	(0.0024)
Dep. Var. Mean	47.91	36.12	67.34	42.70	45.48
Dep. Var. SD	7.44	7.67	8.69	8.06	7.54
R2	0.38	0.33	0.52	0.31	0.26
Observations	1385	1385	1385	1385	1385
C. Share of 2019 graduates continuing to university after graduation					
Share in university	7.31***	6.36***	7.69***	9.98***	5.22***
	(1.53)	(1.66)	(1.48)	(1.89)	(1.55)
Dep. Var. Mean	47.71	35.98	67.11	42.52	45.25
Dep. Var. SD	7.45	7.76	8.75	8.05	7.47
R2	0.42	0.35	0.55	0.36	0.28
Observations	1518	1518	1518	1518	1518
D. Share of 2019 graduates working in vocational sector after graduation					
Share in vocational work	2.39***	1.12*	2.00**	2.77***	3.66***
	(0.77)	(0.61)	(0.93)	(0.85)	(0.98)
Dep. Var. Mean	47.67	35.91	67.09	42.44	45.21
Dep. Var. SD	7.36	7.62	8.59	7.96	7.45
R2	0.39	0.33	0.52	0.32	0.27
Observations	1450	1450	1450	1450	1450
E. Subjective assessment whether 2019 graduates are industry-ready					
Industry ready	0.89***	0.71*	0.90***	1.04***	0.91**
	(0.30)	(0.36)	(0.31)	(0.31)	(0.36)
Dep. Var. Mean	47.80	36.05	67.17	42.65	45.31
Dep. Var. SD	7.51	7.82	8.77	8.16	7.51
R2	0.40	0.33	0.53	0.32	0.27
Observations	1625	1625	1625	1625	1625

Note: Coefficients from regressions of 2019 national examination scores on teachers' reported expectations of their graduates' outcomes. All regressions include an array of province dummies. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: UR training characteristics reported by participants

	N	Mean
Selection criteria w/in school: Subject taught by the teacher	786	0.71
Training duration (in hours)	855	262.57
Trainers from the industry/private	858	0.90
Facilitator trainers from Balai Besar	858	0.88
Activity		
Discussion	861	0.84
Practice with computers	861	0.60
Teaching practice	861	0.28
Internship at the industry	861	0.19
Facilities received		
Craft material	859	0.82
Lesson plan/video	859	0.47
Benefit		
Certification from the industry	861	0.64
Increases in teachers knowledge and skills	861	0.59
After training		
Incorporate in day-to-day teaching	862	0.86
Share material with other teachers	862	0.77
Most helpful training component:		
On the job training	860	0.27
Training material	860	0.25
Recommending graduates from school to industry	800	0.76
Overall score given for the training	861	8.81
Materials mastered before the training: All	861	0.21
Materials mastered before the training: Some	860	0.63
Has taught materials before training	840	0.81

Source: original survey data.

Table A.10: Effect of training on teachers' vocational knowledge and classroom practices, no covariates

	ITT				2SLS			
	Vocational test	**Lectures**	**Independent work**	**Discussion**	**Vocational test**	**Lectures**	**Independent work**	**Discussion**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	0.004	-0.015	0.028	-0.002	0.017	-0.088	0.165	-0.013
	(0.017)	(0.034)	(0.027)	(0.011)	(0.069)	(0.199)	(0.189)	(0.060)
Dep. Var. Mean	0.719	0.465	0.195	0.169	0.719	0.465	0.195	0.169
Dep. Var. SD	0.144	0.236	0.239	0.168	0.144	0.236	0.239	0.168
Observations	395	318	318	318	395	318	318	318

Notes: The table shows coefficients from regressions of the outcome variables in column titles. Columns (1) to (4) show ITT (intention to treat) estimates from OLS regression on an indicator of being assigned to the treatment group. Columns (5) to (8) show 2SLS estimates regression on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. Columns (2) to (4) and (6) to (8) use as dependent variable the share of classroom time dedicated to the activity in column title. All regressions include vocational sector dummies. Clustered standard errors at the vocation sector level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Effect of training on teachers' use of ICT, no covariates

	ITT			2SLS		
	Cover material	**Discussion**	**Pupil work**	**Cover material**	**Discussion**	**Pupil work**
	(1)	(2)	(3)	(4)	(5)	(6)
Training	-0.014	0.066	-0.008	-0.061	0.293	-0.036
	(0.034)	(0.053)	(0.048)	(0.157)	(0.226)	(0.208)
Dep. Var. Mean	0.739	0.359	0.423	0.739	0.359	0.423
Dep. Var. SD	0.440	0.480	0.495	0.440	0.480	0.495
Observations	395	395	395	395	395	395

Notes: The table shows coefficients of regressions of share of ICT used for the classroom activities indicated in the column titles. Columns (1) to (3) show ITT (intention to treat) estimates from OLS regressions on an indicator of being assigned to the treatment group. Columns (4) to (6) show 2SLS estimates from regressions on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. All regressions include vocational sector dummies. Clustered standard errors at the vocation sector level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Effect of training on teachers' expectations of their students' outcomes

	ITT				2SLS			
	Employed	**Wage (USD)**	**Going to university**	**Industry-ready**	**Employed**	**Wage (USD)**	**Going to university**	**Industry-ready**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	0.001	-5.920*	0.003	0.037**	0.008	-30.441	0.019	0.182*
	(0.014)	(2.923)	(0.008)	(0.012)	(0.076)	(23.057)	(0.047)	(0.093)
Dep. Var. Mean	0.351	137.161	0.298	0.107	0.351	137.161	0.298	0.107
Dep. Var. SD	0.208	71.895	0.206	0.309	0.208	71.895	0.206	0.309
Observations	326	314	350	394	326	314	350	394

Notes: Coefficients from regressions of outcome variables in various estimation samples. Column 1 shows ITT (intention to treat) estimates from OLS regression of outcome on an indicator of being assigned to the treatment group in the RCT sample. Column 2 shows 2SLS estimates regression of outcome on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. Column 3 shows estimates from a regression of outcome on an indicator of being assigned to the treatment group in the PSM sample. All regressions include vocational sector dummies. The PSM regression includes matching group dummies. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Effect of training on teacher outcomes in PSM design

		Share of classroom time used in			Share of ICT use to		
	Vocational test	**Lectures**	**Independent work**	**Discussion**	**Cover material**	**Discussion**	**Pupil work**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.048	-0.023	0.020	0.005	0.038*	0.064**	0.028
	(0.040)	(0.015)	(0.016)	(0.010)	(0.022)	(0.026)	(0.028)
Dep. Var. Mean	0.679	0.435	0.236	0.164	0.795	0.398	0.455
Dep. Var. SD	0.128	0.245	0.260	0.168	0.404	0.490	0.498
Observations	116	1001	1001	1001	1200	1200	1200

Notes: The table shows coefficients of the outcomes in the column titles on a dummy of attendance to UR training. All regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and matching group dummies. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Effect of training on teachers' expectations of students' outcomes in PSM design

	Employed	**Wage (USD)**	**Going to university**	**Industry-ready**
	(1)	(2)	(3)	(4)
Training	0.016	-0.250	-0.015**	0.040***
	(0.011)	(2.496)	(0.008)	(0.015)
Dep. Var. Mean	0.376	139.718	0.241	0.082
Dep. Var. SD	0.218	65.656	0.187	0.274
Observations	903	878	999	1181

Notes: The table shows coefficients of the outcomes in the column titles on a dummy of attendance to UR training. All regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and matching group dummies. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Heteregeneity in treatment effects by teacher's characteristics in PSM design

	Master's degree	**Above-median tenure**	**Permanent employee**	**Certified**	**Male**	**In Java/Bali**
	(1)	(2)	(3)	(4)	(5)	(6)
A. Outcome: Share of graduates employed 3 months after graduation						
Training	0.018	0.025	0.037	0.008	0.012	0.055***
	(0.013)	(0.018)	(0.026)	(0.018)	(0.018)	(0.018)
Training × group	-0.006	-0.015	-0.024	0.018	0.011	-0.060**
	(0.036)	(0.028)	(0.031)	(0.025)	(0.025)	(0.026)
Dep. Var. Mean	0.375	0.375	0.375	0.375	0.375	0.375
Dep. Var. SD	0.217	0.217	0.217	0.217	0.217	0.217
Observations	903	903	903	903	903	903
B. Outcome: Average monthly salary in first job (USD)						
Training	-1.208	0.131	-5.837	-8.581**	2.519	1.826
	(2.886)	(3.562)	(5.832)	(4.181)	(3.279)	(3.380)
Training × group	4.533	-0.266	7.122	15.020***	-5.553	-3.559
	(8.206)	(6.440)	(6.604)	(5.195)	(5.457)	(5.482)
Dep. Var. Mean	139.834	139.834	139.834	139.834	139.834	139.834
Dep. Var. SD	65.943	65.943	65.943	65.943	65.943	65.943
Observations	878	878	878	878	878	878

Notes: Each column shows estimates for the coefficients of an attended-training indicator and the interaction between the training indicator and a group indicator. The group indicator is equal to one for the group indicated by the column header. Besides the shown variables, all regressions include the following covariates: gender, years of education, years of teaching experience, an indicator of being a civil servant or a full-time staff, an indicator of certification status, an array of province dummies, and an array of vocational sector dummies. Observations are weighted by the inverse of the number of unites in the control group. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

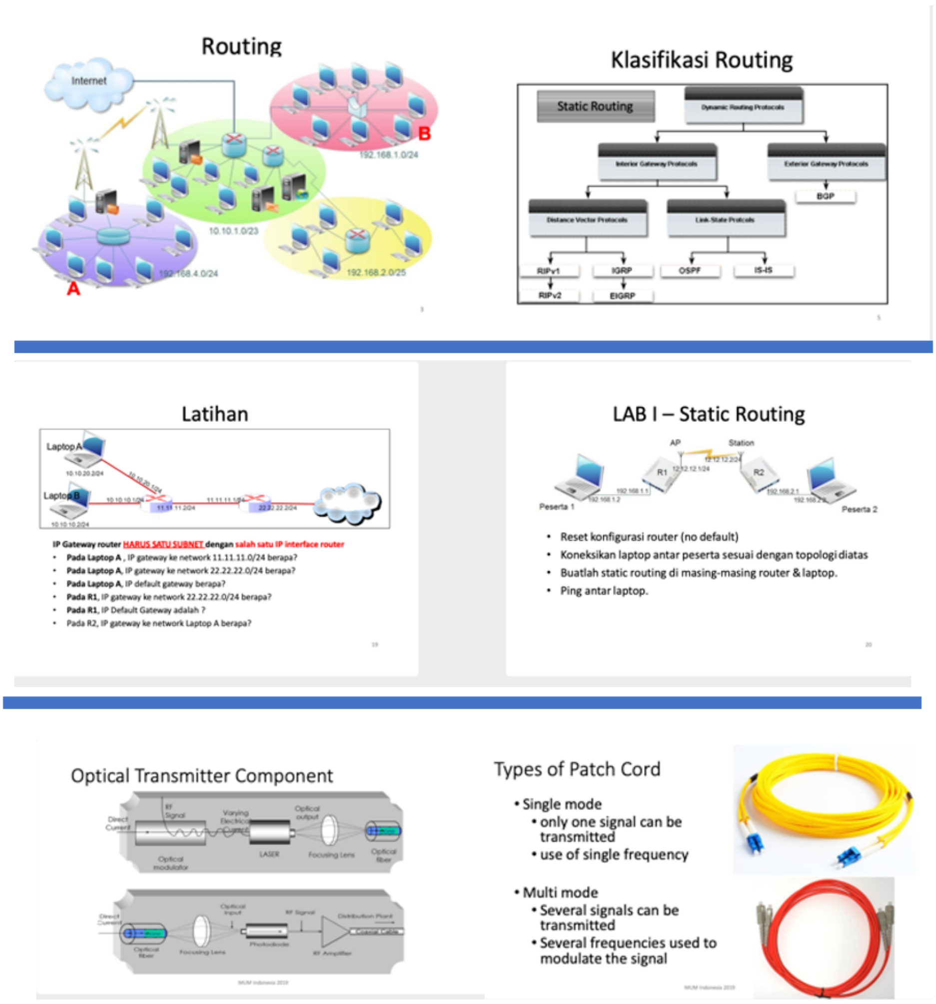
Table A.16: Effect of training on teachers' knowledge and classroom practices, by vocational sectors

	(1) Vocational Knowledge	(2) ICT cover material	(3) ICT independent work	(4) ICT class discussion	(5) % class time lectures	(6) % class time independent work	(7) % class time discussion
A. RPL2 – Java Programming UR course							
Treatment	0.009	-0.040	0.133*	-0.013	-0.057	0.056	-0.002
	(0.023)	(0.068)	(0.080)	(0.082)	(0.042)	(0.043)	(0.035)
Control Group Mean	0.71	0.80	0.33	0.48	0.49	0.17	0.18
Control Group SD	0.13	0.40	0.47	0.50	0.25	0.23	0.20
Observations	147	150	150	150	124	124	124
B. RPL4 – Database Programming UR course							
Treatment	-0.053	-0.011	-0.020	-0.052	0.074	-0.029	-0.040
	(0.038)	(0.112)	(0.114)	(0.122)	(0.068)	(0.071)	(0.039)
Control Group Mean	0.75	0.73	0.32	0.45	0.42	0.21	0.20
Control Group SD	0.14	0.46	0.48	0.51	0.27	0.26	0.15
Observations	89	89	89	89	63	63	63
C. TKJ1 – Mikrotik and fiber optics UR course							
Treatment	0.034	0.026	0.227	0.139	-0.066	0.070	-0.008
	(0.037)	(0.141)	(0.140)	(0.141)	(0.077)	(0.077)	(0.051)
Control Group Mean	0.76	0.67	0.31	0.32	0.48	0.18	0.17
Control Group SD	0.13	0.47	0.47	0.47	0.25	0.25	0.17
Observations	107	106	106	106	80	80	80

Note: Coefficients from regressions of outcome variables in various estimation samples. Column 1 shows ITT (intention to treat) estimates from OLS regression of outcome on an indicator of being assigned to the treatment group in the RCT sample. Column 2 shows 2SLS (treatment on the treated) estimates from 2SLS regression of outcome on an indicator of UR training attendance, instrumented with assignment to the treatment group in the RCT sample. Column 3 shows estimates from regression of outcome on an indicator of being assigned to the treatment group in the PSM sample. All regressions include the following covariates: outcome variable in the year 2019 and an array of vocational sector dummies. The PSM regression includes matching group dummies. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix B Figures

Figure B.1: Snippets of UR 2020 Training Materials

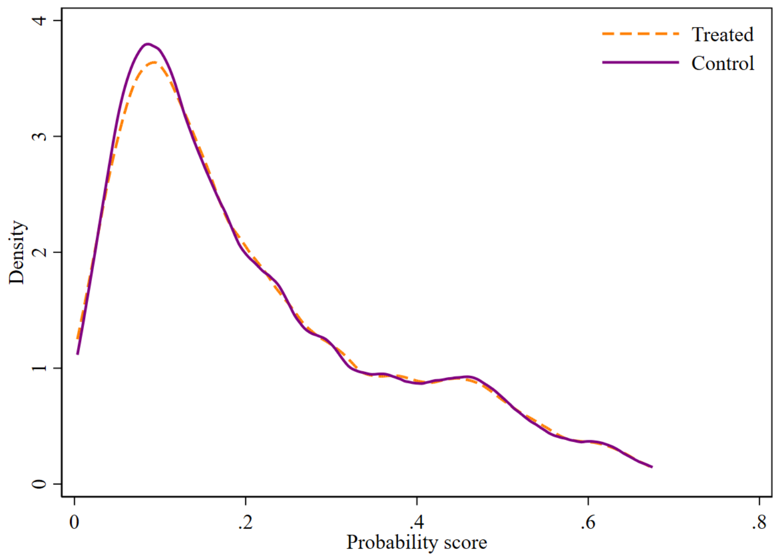


Note: Snippets of training materials from the Network and Communication training, covering modules on routing and fiber optics. Snippets are reproduced from the BPPMPV KPTK training report.

Figure B.2: Snapshots of UR 2020 Training Activities

Note: Photos taken during offline training of one of the Upskilling Reskilling 2020 training. Top left: trainers and MoEC officials on a panel in front of a backdrop with the UR course title name, dates, and location. Top right and bottom left: teachers participated in hands-on activities with network equipments during the training. Bottom left: participating teachers working on individual laptops as part of the training session. Photographs are reproduced from BPPMPV KPTK training report.

Figure B.3: UR: propensity score by treatment status



Note: The figure shows estimates of the propensity score for UR attendees (treated) and people in the control group. The figure combines the scores for all trainings. The distribution of the control group is weighted by the inverse of the number of units in the control group.

Appendix C Propensity Score Matching and Sample Selection

We supplemented the RCT design with a sample of nearly 1,200 teachers who were selected using Propensity Score Matching. We used this second strategy to improve precision and to compare its results with the RCT. Our rationale was that if we obtained similar results under both strategies, this would bolster the –potentially imprecise– results of the RCT.

We included a total of 48 courses in this supplementary design. We matched each of the attendees to these 48 courses to control applicants using a Propensity Score:

$$P(T_i = 1 | \text{ subject} = j) = X_i \beta_j \tag{C.1}$$

we calculate the propensity score by OLS using a series of pre-treatment characteristics available in the 2015 Teacher Competency Exam database. These are: years of education, years of teaching experience, gender, whether the teacher resides in Java island, whether the teacher was certified,[18] school type (public/private), a state-employee dummy, type of contract (permanent/temporary), field of specialization (care services, construction, creative economy, hospitality, machinery, other), and the standardized test score in the 2015 competency exam.

We estimated (C.1) on the set of attendees and applicants with 2015 test score data. We ran a separate regression for each of the 48 training courses. We computed the propensity scores separately by subject because the selection of applicants was fairly decentralized, and the selection procedures could vary by course. For each subject, we included all attendees and, as potential controls, we used all people who applied for admission to that subject. Because people often applied to several trainings, this means that the same individual can appear in the control pool for several subjects. Figure B.3 shows estimates of the propensity score by treatment group for all the trainings in the sample. Note that the estimated scores for the control group match quite closely the UR attendee's scores (treated group).

We matched attendees to controls using the four nearest neighbors with replacement with a caliper of 0.05; that is, for each treated individual we matched up to four controls as long as the difference between the treated and control propensity scores was within 5 percentage points. We slated all the attendees and their four matched controls for the survey. To be included in the results, we had to successfully survey the UR attendee and at least one of their matched controls.

[18]Teachers can get certified on their teaching fields, which unlocks salary supplements.