AUTHOR ACCEPTED MANUSCRIPT

The role of preschool quality in promoting child development: Evidence from rural Indonesia[*]

Sally Anne Brinkman[1], Amer Hasan[2], Haeil Jung[3], Angela Kinnell[4], Nozomi Nakajima[2], Menno Pradhan[5]

Corresponding author email: hijk@korea.ac.kr

Abstract: This article examines the relationship between preschool quality and children's early development in a sample of over 7,900 children enrolled in 578 preschools in rural Indonesia. Quality was measured by: (1) classroom observations using the Early Childhood Environment Rating Scale-Revised (ECERS-R), (2) teacher characteristics, and (3) structural characteristics of preschools. Children's development was measured using the Early Development Instrument (EDI). The article proposes two methodological improvements to preschool quality studies. First, an instrumental variable approach is used to correct for measurement error. Second, ECERS-R is adjusted to the local context by contrasting items with Indonesia's national preschool standards. Results show that observed classroom quality is a significant and meaningful positive predictor of children's development once models correct for measurement error and apply a locally-adapted measure of classroom quality. In contrast, teacher characteristics and structural characteristics are not significant predictors of child development, while holding observed classroom quality constant.

Keywords: Early Childhood Education and Development (ECED), quality, teachers, playgroups, kindergartens, classrooms, Indonesia

---

[*] The authors are listed alphabetically to indicate that this paper is a team product in which all the authors contributed in different ways.

[1] Fraser Mustard Centre, Telethon Kids Institute, University of Western Australia, Perth, Australia.

[2] Education Global Practice, World Bank, Washington, DC, USA.

[3] Corresponding Author: Department of Public Administration, Korea University, Seoul, Republic of Korea. Email: hijk@korea.ac.kr Phone: 82-2-3290-2275

[4] School of Psychology, The University of Adelaide, Adelaide, Australia

[5] University of Amsterdam and VU University, Amsterdam, the Netherlands.

characteristics and structural characteristics are not significant predictors of child development, while holding observed classroom quality constant.

Keywords: Early Childhood Education and Development (ECED), quality, teachers, playgroups, kindergartens, classrooms, Indonesia

Introduction
The quality of preschool education[6] plays an important role in early development. Current research has shown that children who have high-quality early childhood education experiences have better cognitive and socio-emotional outcomes (Engle et al. 2011). Providing quality early childhood education services can be particularly challenging in poor, rural contexts – especially so in developing countries. As a result, few studies have been able to examine the link between quality and child development outcomes in remote, resource-constrained environments using an internationally comparable measure of quality.

This study investigates the quality of early childhood education programs in 303 villages in Indonesia. We use data collected in 2013 as part of an impact evaluation of the Indonesia Early Childhood Education and Development (ECED) Project, which provides rich information on early childhood classrooms, teacher characteristics, and child development outcomes. Indonesia offers a useful setting to explore early childhood education as it is a middle income country where the government has been highly committed to improving both access to and quality of preschools in the country (Hasan, Hyson and Chang 2013).

This paper is organised as follows. First, we review the literature on early childhood education quality by discussing three components of quality: observed classroom quality (focusing on the ECERS-R), teacher characteristics (measured using their education and experience, and structural characteristics. Next, we provide an overview of the early childhood education policy landscape in Indonesia and the research questions in this study. We then introduce the data and methods of the study. This is followed by an outline of the empirical strategy. Results are presented, followed by a discussion of findings and conclusion.

Literature review
While many studies have looked at the impact of preschool or early education quality on children's later development outcomes, there has been no consensus in the literature about how quality should be defined and/or operationalised. This paper will therefore look at three measures of preschool quality which have appeared regularly in the literature and investigate the relationship of each measure with the child development outcomes of those in attendance. These three measures are observed preschool quality, teacher characteristics, and structural characteristics and will be discussed in turn.

Observed preschool quality
Decades of research on child development has established the benefits of providing high-quality early childhood programs to young children (Campbell and Ramey, 1994; NICHD, 2005). One of the most widely used observational measures of the quality of early childhood classroom environments is the revised version of the Early Childhood Environment Rating Scale (ECERS-R) (Harms, Clifford and Cryer, 2005).

Empirical studies have generally found positive associations between quality (measured using the using ECERS-R) and various measures of child development. These range from language and cognitive skills (Burchinal et al., 2008; Peisner-Feinberg and Burchinal, 1997; Peisner-Feinberg et al., 2001) to

---

[6] In this paper, the terms "preschool" and "early childhood education" are used interchangeably, and refer broadly to organized, center-based early learning environments for children from the ages of 3 until they enter primary school – usually playgroups and kindergartens.

social-behavioral development (Burchinal et al., 2008; Sylva et al., 2006). However, the magnitude of these associations has been small—as low as 0.04 when models include extensive control variables (Duncan, 2003). Such findings, however, are not specific to the ECERS-R alone (Gordon et al., 2013; Vandell and Wolfe, 2000). Other widely-used measures of observational classroom quality – such as the Classroom Assessment Scoring System (CLASS) – have also shown fairly small associations with child development outcomes (Howes et al. 2008, Vandell and Wolfe 2000). In studies that have produced particularly small effect sizes, authors suggest that effect sizes of observed classroom quality are likely to be underestimated due to measurement error (Duncan, 2003). Recent studies recommend researchers and practitioners alike pay more attention to how measures of quality like ECERS-R align with local regulations and accreditation standards (Gordon et al., 2013). Taken together, further work in this field must (1) address the measurement error present in observational measures of classroom quality and (2) use internationally comparable measures of quality in locally appropriate ways.

Teacher characteristics
In addition to directly observing the quality of the learning environment, teacher characteristics are another common way of measuring the quality of early childhood programs. There is considerable policy interest in understanding the link between teacher qualifications and children's developmental outcomes since governments can more readily regulate a teacher's level of education, experience, and training than observed classroom quality. However, prior research on the link between teacher characteristics in early childhood programs and children's developmental outcomes is not uniformly positive. Some studies show that classrooms with more highly educated and trained teachers are associated with higher quality care (Burchinal et al., 2002; Phillips et al., 2001). In contrast, other studies have demonstrated that once unobserved differences across centers are controlled for using center fixed effects, the effect of teacher's education and training on child development outcomes disappear (Blau, 2000; Early et al., 2006, Early et al., 2007).

Structural characteristics
There are other characteristics of preschools that can perhaps be regulated even more easily than teacher characteristics. These so-called structural characteristics are usually easier to mandate and less costly to assess than observing the quality of individual classrooms.

Two structural characteristics that are often investigated in studies of preschool quality are teacher-child ratio and hours of operation. For teacher-child ratio, studies have generally shown that in lower-ratio settings, teachers spend less time managing children in the classroom and are able to provide more stimulating, responsive, and warm care (Phillips et al. 2001, Burchinal et al., 2000). Meanwhile, research on the relationship between hours spent in early childhood education and children's developmental outcomes are mixed. For cognitive and language development, studies find positive correlation between time spent in care and children's development but the association disappears once children's family background and other preschool aspects are controlled (Duncan, 2003; NICHD 2000; Vandell, 2004). Similarly, studies that have looked at children's social competence and behavioral problems have also found mixed results. In the US, researchers found that even when holding family factors constant, children who spend more hours in preschool have more behavioral problems than their peers who spend fewer hours in preschool settings (NICHD, 2003; Vandell, 2004). In contrast, more recent research from Norway finds little evidence of behavioral issues in children from spending more hours in child care (Zachrisson et al, 2013). To our knowledge, early childhood education studies from developing settings have yet to examine whether structural characteristics matter to children's developmental outcomes even after controlling for key confounding factors such as observed classroom quality (which captures teacher and student interactions within classrooms) and teacher characteristics.

The basic conceptual framework for this paper is depicted in Figure 1. Observed classroom quality, teacher characteristics, and structural characteristics are each thought to be associated with child

development when family and child characteristics are controlled for. Based on this conceptual framework, this study examines the differential associations between the three aspects of preschool quality and children's developmental outcomes, and their relative importance.

<<Insert Figure 1 here>>

This study addresses several limitations in the prior literature on the relationship between the quality of early childhood programs and children's outcomes. First, we correct for possible measurement error in observational measures of classroom quality using an instrumental variable approach. Second, we contrast items in the ECERS-R with Indonesia's national quality standards for early childhood education to propose an alternative measure of classroom quality that is more locally-relevant. Third, we contribute to the literature on the link between teacher and structural aspects of preschool quality and early developmental outcomes, which remains understudied in resource-constrained contexts.

Indonesian context
Several types of early childhood services exist in Indonesia; they serve different purposes and are administered by different ministries. Table 1 summarises the key features of some of these programs:
<<Insert Table 1 here>>
Among the many existing types of early childhood programs in Indonesia, kindergartens and playgroups focus on providing education directly to children. Generally, playgroups emphasise learning through play while kindergartens help prepare students for primary school (Hasan, Hyson and Chang, 2013). Given the focus on early childhood education in this study, we analyse quality and child outcomes in kindergartens (both *Kindergarten* and *Islamic kindergarten*) and playgroups.

In recent years, the government of Indonesia has prioritised early childhood education and development. The national standard for ECED was developed by the Ministry of Education and Culture in 2009 to ensure overall quality in early childhood programs across the country.[7] Key areas of the national standard included teacher qualifications such as a higher education degree for teachers and a demonstration of a wide range of professional competencies in teaching young children, which would require significant experience in the classroom. The Indonesia ECED standard also defines policies on structural aspects of early childhood centers, such as class size and duration. Although these teacher and structural policies have been in place since 2009, little is known about whether they are adhered to and whether they actually help promote child development outcomes.

As the government of Indonesia moves to expand and improve early childhood education programs across the country, there is immense interest among policymakers for evidence on whether classroom quality, teacher qualifications, and other structural characteristics do in fact support positive developmental outcomes for children. Thus, this study focuses on the following three research questions in the context of 303 poor villages in Indonesia:

(1)   Does observed classroom quality of early childhood education programs predict children's developmental outcomes?
(2)   Do teacher education, experience, and training predict children's developmental outcomes in early childhood programs?
(3)   Do structural characteristics—student-to-staff ratio and hours of operation—predict children's developmental outcomes?

---

[7] As of late 2014 these standards have been updated. However, at the time data for this study was being collected, the 2009 version of the Indonesia standards were in effect. These are the standards described in this section.

Data and Measures
Data
Data for this analysis were collected in 2013 as part of an impact evaluation of the Indonesia ECED Project (see Pradhan et al. 2013 for detailed study protocol). The Indonesia ECED Project aimed to improve poor children's access to ECED services and enhance children's school readiness. Through the project, participating villages created playgroups (hereafter referred to as *project playgroups*) and received training for teachers in these playgroups. As part of the project evaluation, data on children enrolled in project playgroups as well as their peers living in the same village attending other types of early childhood programs (i.e., kindergartens, Islamic kindergartens, and non-project playgroups) were collected.

The quality of 578 early childhood programs—located in 303 poor villages across 9 districts—was observed. In total the development outcomes of over 7,900 children attending these centers were measured.

Measures
Early Development Instrument (EDI)
The key dependent variables of interest are children's developmental outcomes measured using the Early Development Instrument (EDI), which has been shown to be a valid and reliable measure of child development internationally (Ip et al., 2013; Janus et al., 2011; Janus et al., 2007; Janus and Offord, 2007; Brinkman et al., 2013). The EDI is comprised of five domains: physical health and well-being, social competence, emotional maturity, language and cognitive skills, and communication skills and general knowledge. Each domain is scored from 1 (low) to 10 (high). A teacher in the early childhood education center in which the child was enrolled completed the child's EDI. The EDI was adapted and translated for use in the Indonesia ECED Project by the authors and members of the research team. Descriptive statistics of the EDI, along with child covariates, are shown in Table 2.

<<Insert Table 2 here>>

Both girls and boys are equally represented across kindergartens and playgroups. Children attending kindergarten are slightly older (mean age of 5 years) than those enrolled in playgroups (mean age of 4 years). The highest level of education attained by mothers of children in the sample are similar across all types of preschools with about 40% having completed elementary school or less, followed by junior and senior high (about 25% each), and higher education (less than 10%).
Observed preschool quality
One of the main predictor variables is the quality of early childhood programs as measured by ECERS-R. The instrument has good test-retest reliability, high inter-rater reliability (Clifford et al., 2010), and many studies have demonstrated its predictive validity (Burchinal et al., 2008; Montes et al., 2005; Peisner-Feinberg et al., 2001). Each center was assessed by two raters on a seven-point Likert scale, which ranges from inadequate (1), minimal (3), good (5) to excellent (7). Seven subscales make up the ECERS-R and the total ECERS-R is the average score of the subscales. Descriptive statistics of the ECERS-R are shown in Table 3.[8]

<<Insert Table 3 here>>

On average, preschools in Indonesia score slightly under 3 (minimal) on the ECERS-R. However, when we look across a number of other studies that have carried out an assessment of pre-school

---

[8] It is important to note that while the range of the ECERS-R only goes from 1 to 7 – the differences in quality between a 3 and a 5 are substantial.

quality using ECERS-R, we find that services in rural Indonesia are not unique in their inability to score well (see Figure 2). Even services in Sweden or parts of Canada do not score above a 5 (good) on this scale on average. This suggests that the ECERS-R sets a high bar for excellence.

<<Insert Figure 2 here>>

As noted earlier, recent work in this field highlights the need for researchers and practitioners alike to pay more attention to how measures of quality like ECERS-R align with local regulations and accreditation standards (Gordon et al., 2013). In an effort to align the ECERS-R data with the reality of the Indonesian context, we turn to the Indonesian national standard for ECED as an alternative way to look at quality. Using textual analysis, we compared ECERS-R to Indonesia's national standard and find 28 out of 43 ECERS-R items discussed in the national standard (see Table 4). Based on this crosswalk of ECERS-R items to Indonesia's standards, we calculate the mean ECERS-R score using only the 28 items that were found to be common between ECERS-R and the national standard; we call this alternative measure of quality the "Indonesia standard." The mean and standard deviation of this alternative measure is presented in Table 5. Thus, our study examines the relationship between observed classroom quality and child development outcomes using two measures: the total ECERS-R score and the Indonesia standard.

<<Insert Table 4 & Table 5 here>>

Teacher and structural characteristics
Teacher characteristics are averaged at the center level. They are constructed using the information of teachers who were present on the day of the ECERS-R assessment. These teachers completed the student's EDI but may not necessarily be the child's classroom teacher. Each teacher variable is created by dividing the number of teachers in a center with a particular characteristic by the total number of teachers observed in that center on that day. In our sample, between 1 and 7 teachers were observed in each center. Thus, teacher characteristics in our analysis are continuous variables that measure the average teacher characteristics at the center-level. In addition, student-to-staff ratio and hours of operation per week are included to account for structural characteristics in our models.

Summary statistics of teacher and structural characteristics are described in Table 3. A larger proportion of teachers in kindergartens (64.8% and 48.1%) had post-secondary education than those in playgroups (27.3% and 23.3%). Similarly, teachers in kindergartens had more years of teaching than teachers in playgroups. Teachers were also asked to report on their past experience. Their responses fell in the following categories; ECED teacher experience (they had been teachers in kindergartens or playgroups before); ECED non-teacher experience (they had worked in other early childhood programs like village health services (*Posyandu*), family planning, or a woman's group); or had no prior ECED experience. About a third of teachers in the sample did not have any prior ECED experience. Finally, teachers were asked about their training since one of the key features of the Indonesia ECED Project was provision of training to teachers. The training under the project emphasised interactive learning, demonstrations, and role-play to become effective instructors for young children. The project training was 200 hours, which was considerably longer than training typically offered by accredited teacher training programs in Indonesia at the time.

The total sample consists of 578 early childhood education centers with observations on child development outcomes for over 8,300 children.[9] We have missing data on teacher and structural characteristics of the centers, which reduces our analytic sample size to 566 centers. This results in usable

---

[9] The sample size for each EDI domain is slightly different. N=8,348 for physical health, 8,317 for social competence, 8,306 for emotional maturity, 8,319 for language and cognitive development, and 8,345 for communication and general knowledge.

data on child development outcomes for 7,946 children.[10]

Empirical strategy

We examine the differential associations between the three aspects of preschool quality and children's developmental outcomes while controlling for child, center, and district characteristics using Model 1 as follows:

$$Y_{ijk} = \beta_0 + \beta_1 Q_{jk} + \alpha_2 T_{jk} + \alpha_3 C_{jk} + \alpha_4 X_{ijk} + \alpha_4 S_{jk} + \nu_k + u_{ijk} \qquad (1)$$

where Y_{ijk} is the developmental outcome (one of the EDI domains) for child i enrolled in center j located in district k. First, by adding Q_{jk}, the mean ECERS-R score of center j in district k, we study whether ECERS-R predicts children's developmental outcomes. Second, we add average teacher characteristics (T_{jk}) at the center level in district k to equation 1 in order to examine whether teacher characteristics predict child development outcomes, over and above observed classroom quality. Third, we include center-level variables in district k (C_{jk}) to examine whether other structural characteristics (i.e., student-to-staff ratio and hours of operation per week) predict child development outcomes. As control variables of child and center characteristics, X_{ijk} represents a vector of observable child characteristics for child i, in center j located in district k and S_{jk} is a dummy variable indicating the type of early childhood service (1=kindergarten/Islamic kindergarten, 0=project/non-project playgroup) provided in center j in district k. District characteristics are controlled for using district-specific unobserved variables (using district fixed effects) embedded in ν_k and the error term is u_{ijk}.

In Model 2, we use an instrumental variable approach to correct for possible measurement errors in our score of observational classroom quality. Such quality ratings are subject to measurement error as they involve judgment on the part of each observer. Measurement error increases the noise in the quality variable leading to a downwards bias in the estimated correlations between the quality and child development outcomes.[11] To correct for this potential bias, we exploit the fact that two observers rated each center independently. Under this assumption, we can apply an instrumental variable approach where one observer's rating serves as an instrument for the rating of the other observer for the same center. This corrects for the downward bias due to measurement error.[12] Thus instead of using this average ECERS-R score as we have done in equations 1 to 3, we use the first rater's score Q_j^1, as an instrument for the second rater's score Q_j^2, in a 2 Stage Least Squares (2SLS) model as follows:

First stage regression:

$$Q_{jk}^2 = \alpha_0 + \alpha_1 Q_{jk}^1 + \alpha_2 T_{jk} + \alpha_3 C_{jk} + \alpha_4 X_{ijk} + \alpha_4 S_{jk} + \nu_k + u_{ijk} \qquad (2.1)$$

Second stage regression:

$$Y_{ijk} = \gamma_0 + \gamma_1 \widehat{Q_{jk}^2} + \gamma_2 T_{jk} + \gamma_3 C_{jk} + \gamma_4 X_{ijk} + \gamma_4 S_{jk} + \nu_k + \varepsilon_{ijk} \qquad (2.2)$$

Finally, in Model 3, we re-estimate the instrumental variable approach from Model 2 using the Indonesia standard instead of the total ECERS-R score to examine whether a more locally-relevant measure of observed classroom quality improves estimates of the association between quality and child outcomes.

Ethical considerations

This study complied with all local ethical and permission requirements prevailing at the time it was conducted. Survey studies conducted in Indonesia do not require ethics approval. As such, neither the Ministry of National Education (MoNE) nor the World Bank required formal ethics approval for this study. Despite this the research team in collaboration with MoNE and the survey firm undertook several measures to ensure that participants in the survey were aware that their participation in the study was

---

[10] Sample size for each EDI domain also varies for these 566 centers, going up to N=7,984 for emotional maturity.

[11] Please see the discussion on classical errors in independent variables on pages 295-296 in Wooldridge (2002).

[12] Note that this strategy corrects for bias due to measurement error, not due to endogeneity of the quality ratings. The parameter estimates should still be interpreted as correlations, not as causal effects of quality.

voluntary and that data collected would be analysed anonymously. In addition, the survey firm operated with a letter of approval from the MoNE (Directorate General of Non-Formal and Informal Education). The MoNE then issued a letter to each survey district education office requesting the survey firm's permission to conduct the survey over a specified time period. The survey firm showed this letter to the district education office, which in turn issued a permit letter to the survey team, which could be presented to the village and dusun offices to gain permission to survey in these areas.

Results

Tables 6 through 10 present the unstandardised (raw EDI scores) regression results of the three models. Each table shows the outputs for a different EDI domain and each column presents the results from a separate model. Below, we organise our results around observed classroom quality (ECERS-R and Indonesia standard), aggregate teacher characteristics, and structural characteristics.

Observed classroom quality

The OLS results in Model 1 indicate that there are few considerable associations between observed classroom quality (ECERS-R) and child development outcomes, which is similar to the findings in Gordon et al. (2013) and Sylva et al. (2006).[13] It seems that one of the limitations of observational measures of classroom quality is that its effect size is likely to be underestimated due to measurement error. As a result, we attempt to correct for measurement error to yield more precise estimates of observed classroom quality on children's developmental outcomes in the following models.

The results in Model 2 suggest that using an instrumental variables approach produces stronger associations between ECERS-R and EDI. Compared to Model 1, Model 2 shows that a one-unit increase in ECERS-R is associated with a 0.107 unit increase in the physical health and well-being domain, holding all else constant. Similarly, the coefficient of ECERS-R is 0.111 for social competence, while controlling for all other variables. For other domains the same pattern is observed, although the correlations are not statistically significant at the conventional levels.

Finally, results in Model 3 show that using a measure of observed classroom quality that aligns with the local context (in this case, Indonesia's national standards) yields significant, positive relationships between observed quality and EDI. A one-unit increase in classroom quality is associated with a 0.167 unit increase in children's language and cognitive skills, holding all else constant. Similarly, the coefficient of quality is significant for physical health (0.107 units), social competence (0.121 units), and communication and general knowledge (0.133 units).

Teacher characteristics

For teacher characteristics aggregated at the center level, we generally find mixed results across the various models. In terms of teacher's education level, results from Models 1-3 show that increasing the percent of teachers with a secondary or post-secondary degree predicts better EDI scores in social competence and communication and general knowledge, while controlling for all other variables. For example, a 100% increase in teachers with a post-secondary education degree is associated with a 1.073-1.086 unit increase in social competence, holding all else constant. For both social competence and communication and general knowledge, the coefficient on the percent of teachers with post-secondary education is larger than that of senior secondary education. For the other three EDI domains, we find no significant relationship between increasing the share of teachers with higher levels of education and EDI.

Across the models, we find that increasing the mean years of teaching experience at the center level does not predict higher EDI scores. In fact, for the physical health and well-being domain and the communication and general knowledge domain, an additional year of average teaching experience in a center is associated with a 0.014-0.015 unit decrease (physical health and well-being) or a 0.043-0.045 unit decrease (communication and general knowledge), holding all else constant. In interpreting these results, it is important to keep in mind that these data are aggregated at the center level and do not link specific children to specific teachers.

---

[13] Regression results for Models 1 to 3 with standardized beta coefficients are available upon request.

For previous experience in ECED, we find null results on children's development. Similarly, we find no significant associations between teachers who have received teacher training and EDI.

Structural characteristics

Over and above aspects of the quality of the learning environment, we find that the quantity of exposure (or dose) of early childhood education is important to children's developmental outcomes in two out of five EDI domains. An additional hour per week of early childhood education is associated with a 0.036-0.039 unit increase in language and cognitive skills and a 0.027-0.028 unit increase in social competence, controlling for all other variables. In contrast, we find null results for student-to-staff ratio on EDI, except in physical health and well-being.

Discussion

Our study addresses two limitations in the current literature on preschool quality. First, classroom observations of ECED quality are subject to considerable measurement error. We correct for measurement error using an instrumental variable approach (with the first rater's score as an instrument for the second rater's score), which allows us to improve the use of ECERS-R as a predictor of child development outcomes.

Second, studies have not sufficiently addressed how measures of quality like ECERS-R align with local regulations and accreditation standards, particularly in resource-constrained environments. We show that using a subset of the ECERS-R items that correspond with local preschool standards can provide researchers with an alternative measure of classroom quality that aligns closely to the particular context of study. The estimated effect sizes of observational classroom quality for four out of five EDI domains (physical health and well-being, social competence, language and cognitive skills, and communication skills and general knowledge) suggests that observed classroom quality is a modest, reliable predictor of children's developmental outcomes during early childhood in rural Indonesia.

For the most complex model in this study (Model 3), a one s.d. increase in classroom quality related to a 0.071 to 0.082 s.d. increase in children's developmental outcomes. This range of effect sizes appears to be within the range of other empirically rigorous studies that include extensive control variables or use fixed effects, both in the United States and in developing countries. For example, large-scale studies from the United States have effect sizes ranging from 0.04 (Duncan, 2003) to 0.18 (Peisner-Feinberg et al., 2001) on children's language and cognitive skills. For socio-emotional skills, prior studies report effect sizes as small as 0.02 in the U.S. (Peisner-Feinberg et al., 2001) to as large as 0.13 in the U.K. (Sylva et al., 2006). Meanwhile, an East African study reports an effect size as large as 0.15 on children's cognitive skills but the positive relationship is reported only for high-quality Madrasa centers and not for other community preschools (Malmberg et al., 2011). A study from Ecuador, which uses CLASS instead of ECERS-R, shows similar relationships between classroom quality and children's learning outcomes with effect sizes ranging from 0.06 to 0.13 s.d.

Compared to previous studies on preschool quality from developing contexts, our regression analyses introduced an extensive set of controls to adjust for potential biases that might lead to inconsistent estimates of observational classroom quality. Thus, our study confirms that even with these extensive controls (i.e., child, teacher, structural characteristics with district fixed effects), observed classroom quality is a significant and meaningful positive predictor of child development outcomes.

In addition to observed classroom quality, teacher's education is a significant predictor of children's social competence as well as communication and general knowledge. The effect size of teacher's education, particularly post-secondary education, is relatively large and ranges from 0.184-0.264 s.d. in these two EDI domains.

In contrast, mean years of teaching is negatively correlated with physical health and well-being and communication and general knowledge—although the effect sizes are relatively small, ranging from -0.056 to -0.117 s.d. In addition, experience in preschool and receiving teacher training yielded null results. It is important to note, however, that teacher characteristics used in this study are averaged at the center-level, which means we cannot attribute student outcomes directly to their teachers. As a result, we interpret

the small but negative coefficient on years of teaching and the null results of ECED experience and teacher training as follows: policies focused solely on hiring teachers with more experience and training will be insufficient to improve children's development in rural Indonesia. Rather, policies must address the quality of professional development activities for teachers in order to ensure the effectiveness of early childhood education programs. Our mixed findings on teacher characteristics are not surprising given that prior studies on teacher qualification have also shown mixed evidence in relation to children's developmental outcomes. The few studies that have successfully produced statistically significant results often employ a value-added model (Hanushek and Rivkin, 2006).

Finally, our analysis suggests that the quantity of exposure to early childhood education also matters. On average, children enrolled in more hours of early childhood programs in rural Indonesia scored higher on the EDI, when controlling for various child, teacher, and structural characteristics. Although previous studies examining the relationship between duration of preschool and child development outcomes have been mixed (Vandell, 2004), we find a positive association despite the comparatively low dose of early childhood education in rural Indonesia. On average, children in our sample were attending kindergartens or playgroups for 2-3 hours a day for 4-6 days per week. The maximum number of hours in a week was 24 hours, which was only found in 3 kindergartens in the sample. In contrast, studies that have found negative effects of quantity of early care often focus on children who spend a substantial amount of time in care settings (i.e., 45 hours a week) over an extended period (Vandell, 2004). As such, in rural Indonesia—where early childhood programs are relatively low dose—children are likely to benefit from attending somewhat longer hours of playgroups and kindergartens.

Limitations and areas for future work

There are a few limitations to our study. First, our data on student EDI is likely to have measurement error. The study protocol was to collect the EDI of approximately 15 children in each early childhood education center and that the child's main teacher would fill out the EDI. In practice, however, children in these centers were taken care of by multiple teachers and it was not feasible for the main teacher to fill out the EDI for all sampled children in the center. As a result, a teacher who was familiar with the child's development filled out the EDI questionnaire, but this teacher may or may not have been the child's main teacher.

Second, our teacher level data does not allow us to match teacher observations with student EDI. As a result, the measure of teacher characteristics used in this study is averaged at the center-level, which is a less precise measure of teacher quality than could be derived from matched teacher-student datasets. Future studies of early childhood quality in developing countries could benefit from matched teacher-student data to improve the estimates of teacher quality and better understand what kinds of qualification of ECED teachers are particularly successful in producing high-level student outcomes.

Another explanation for the mixed results on teacher characteristics may be due to collinearity between measures of teacher qualifications and items on the ECERS-R that focus on teacher behavior in the classroom. It is likely that those with higher levels of education, more experience, and better training tend to provide higher quality care as measured by the items on the ECERS-R, making it difficult to tease apart classroom quality and teacher qualifications.

Conclusion

Our analysis of early childhood education quality in Indonesia makes two important contributions to the literature. First, classroom observations of ECED quality are subject to considerable measurement error. Not correcting for this yields an underestimation of the association between quality and child development outcomes. Correcting for measurement error using an instrumental variable approach (with the first rater's score as an instrument for the second rater's score) allows us to improve the use of ECERS-R as a predictor of child development outcomes.

Second, in countries with a national early childhood education standard, using a subset of the ECERS-R items that correspond with the national standard can provide researchers with an alternative

measure of classroom quality that aligns closely to the particular context. The estimated effect sizes of observational classroom quality for four out of five EDI domains (physical health, social competence, language and cognitive development, and communication and general knowledge) suggests that observed classroom quality is a modest, reliable predictor of children's developmental outcomes during early childhood in rural Indonesia.

From a policy perspective, our study confirms the importance of investing in high-quality early childhood education programs. As Indonesia considers expanding early childhood education, a necessary first step is to ensure that programs meet existing standards. However, as the country concentrates on raising the quality of services, the standard itself will need to be revised to ensure if early childhood education programs continue to have meaningful impact on children's development.

References

Aboud, F.E. 2006. "Evaluation of an early childhood preschool program in rural Bangladesh." *Early Childhood Research Quarterly* 21(1): 46-60. doi:10.1016/j.ecresq.2006.01.008

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., and Schady, N. 2014. *A helping hand? Teacher quality and learning outcomes in kindergarten.* Washington, D.C.: Inter-American Development Bank.

Barnett, W. S. 1995. "Long-term effects of early childhood programs on cognitive and school outcomes." *The Future of Children* 5(3): 25-50.

Blau, D. M. 2000. "The production of quality in child-care centers: Another look." *Applied Developmental Science* 4(3): 136-148. doi:10.1207/S1532480XADS0403_3

Brinkman, S., Gregory, T., Harris, J., Hart, B., Blackmore, S., and Janus, M. 2013. "Associations between the Early Development Instrument at Age 5 and Reading and Numeracy Skills at Ages 8, 10 and 12: A Prospective Linked Data Study." *Child Indicators Research* 6(4): 695-708. doi: 10.1007/s12187-013-9189-3

Burchinal, M., Cryer, D., Clifford, R., and Howes, C. 2002. "Caregiver training and classroom quality in child care centers." *Applied Developmental Science* 6(1): 2-11. doi:10.1207/S1532480XADS0601_01

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., and Barbarin, O. 2008. "Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction." *Applied Developmental Science* 12(3): 140-153. doi:10.1080/10888690802199418

Burchinal, M. R., Roberts, J. E., Riggins Jr, R., Zeisel, S. A., Neebe, E., and Bryant, D. 2000. "Relating quality of center-based child care to early cognitive and language development longitudinally." *Child Development* 71(2): 339-357.

Campbell, F. A. and Ramey, C. T. 1994. "Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families." *Child Development* 65(2): 684-698. doi: 10.2307/1131410

Clifford, R., Reszka, S., and Rossbach, H. 2010. "Reliability and validity of the early childhood environment rating scale." *Working Paper.* Chapel Hill: FPG Child Development Institute, University of North Carolina.

Duncan, G. J. 2003. "Modeling the impacts of child care quality on children's preschool cognitive development." *Child Development* 74(5): 1454-1475. doi:10.1111/1467-8624.00617

Duncan, G. J. and Magnuson, K. 2013. "Investing in preschool programs." *The Journal of Economic Perspectives: A Journal of the American Economic Association* 27(2): 109.

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., ... and Barbarin, O. 2006. "Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten?" *Early Childhood Research Quarterly* 21(2): 174-195.

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., ... and Zill, N. 2007. "Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs." *Child development* 78(2): 558-580.

Engle, P. L., Fernald, L. C., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., ... and Global Child Development Steering Group. 2011. "Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries." *The Lancet* 378(9799): 1339-1353.

Esposito, Y. L., Rosemberg, F., de Andrade, D.F., Gimenes, N., and Unbehaum, S. 2010. *An overview of the study: Early Childhood Education in Brazil. A qualitative and quantitative assessment.* Fundação Carlos Chagas, Brazilian Ministry of Education, and the Inter-American Development Bank.

Glewwe, P. and Kremer, M. 2006. "Schools, teachers, and education outcomes in developing countries." *Handbook of the Economics of Education* 2: 945-1017.

Goelman, H., Forer, B., Kershaw, P., Doherty, G., Lero, D., and LaGrange, A. 2006. "Towards a predictive model of quality in Canadian child care centers." *Early Childhood Research Quarterly* 21(3): 280-295.

Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., and Abner, K. 2013. "An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development." *Developmental psychology* 49(1): 146. doi: 10.1037/a0027899

Hanushek, E. A. 2003. "The Failure of Input-based Schooling Policies." *The economic journal* 113(485): F64-F98. doi: 10.1111/1468-0297.00099

Hanushek, E. A. and Rivkin, S. G. 2006. "Teacher quality." *Handbook of the Economics of Education* 2: 1051-1078.

Harms, T., Clifford, R. M., and Cryer, D. 2005. *Early Childhood Environment Rating Scale - Revised Edition*. New York: Teachers College Press.

Hasan, A., Hyson, M., and Chang, M. (Eds.). 2013. *Early Childhood Education and Development in Poor Villages of Indoensia: Strong Foundations, Later Success.* Washington DC: World Bank.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., and Barbarin, O. 2008. "Ready to learn? Children's pre-academic achievement in pre-kindergarten programs." *Early Childhood Research Quarterly* 23(1): 27-50.

Ip, P., Li, S. L., Rao, N., Ng, S. S. N., Lau, W. W. S., and Chow, C. B. 2013. "Validation study of the Chinese Early Development Instrument (CEDI)." *BMC Pediatrics* 13(1): 146. doi: 10.1186/1471-2431-13-146

Janus, M., Brinkman, S., and Duku, E. 2011. "Validity and psychometric properties of the Early Development Instrument in Canada, Australia, United States and Jamaica." *Social Indicators Research* 103(2): 283-297. doi: 10.1007/s11205-011-9846-1

Janus, M., Brinkman, S., and Duku, E. 2007. *The EDI Handbook*. Ontario Canada: McMaster University Press.

Janus, M. and Offord, D. 2007. "Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness." *Canadian Journal of Behavioural Science* 39: 1-22.

Liang, X., Zhang, Y., and Fu, Y. 2013. *Early childhood education in Yunnan: Challenge and opportunities.* Washington, D.C.: World Bank Group.

Magnuson, K. A., Ruhm, C., and Waldfogel, J. 2007. "Does prekindergarten improve school preparation and performance?" *Economics of Education Review* 26(1): 33-51. doi: 10.1016/j.econedurev.2005.09.008

Malmberg, K. A., Mwaura, P., and Sylva, K. 2011. "Effects of a preschool intervention on cognitive development among East-African preschool children: A flexibly time-coded growth model." *Early Childhood Research Quarterly* 26(1): 124-133. doi: 10.1016/j.ecresq.2010.04.003

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., ... and Howes, C. 2008. "Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills." *Child Development* 79(3): 732-749.

Montes, G., Hightower, A. D., Brugger, L., and Moustafa, E. 2005. "Quality child care and socio-emotional risk factors: No evidence of diminishing returns for urban children." *Early Childhood Research Quarterly* 20(3): 361-372. doi: 10.1016/j.ecresq.2005.07.006

National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network. 2000. The relation of child care to cognitive and language development. *Child Development* 71(4): 960-980.

National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network. 2003. "Does amount of time spent in child care predict socioemotional adjustment during the transition to kindergarten?" *Child Development* 74(4): 976-1005.

National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network. 2005. "Early Child Care and Children's Development in the Primary Grades: Follow-up Results from the NICHD Study of Early Child Care." *American Educational Research Journal 42*(3): 537-570. doi: 10.2307/3700462

Peisner-Feinberg, E. S. and Burchinal, M. R. 1997. "Relations Between Preschool Children's Child-Care Experiences and Concurrent Development: The Cost, Quality, and Outcomes Study." *Merrill-Palmer Quarterly* 43(3): 451-477. doi: 10.2307/23093333

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., and Yazejian, N. 2001. "The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade." *Child Development* 72(5): 1534-1553. doi: 10.1111/1467-8624.00364

Phillips, D. A., Mekos, D., Scarr, S., McCartney, K., and Abbott–Shim, M. 2001. "Within and beyond the classroom door: assessing quality in child care centers." *Early Childhood Research Quarterly* 15(4): 475-496. doi: 10.1016/S0885-2006(01)00077-1

Pradhan, M., Brinkman, S. A., Beatty, A., Maika, A., Satriawan, E., de Ree, J., and Hasan, A. 2013. "Evaluating a community-based early childhood education and development program in Indonesia: study protocol for a pragmatic cluster randomized controlled trial with supplementary matched control group." *Trials*,14 (1), 259. doi:10.1186/1745-6215-14-259

Rao, N. 2010. "Preschool quality and the development of children from economically disadvantaged families in India." *Early Education and Development* 21(2): 167-185. doi:10.1080/10409281003635770

Sheridan, S., Giota, J., Han, Y. M., and Kwon, J. Y. 2009. "A cross-cultural study of preschool quality in South Korea and Sweden: ECERS evaluations." *Early Childhood Research Quarterly*, 24(2), 142-156. doi:10.1016/j.ecresq.2009.03.004

Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., and Totsika, V. 2006. "Capturing quality in early childhood through environmental rating scales." *Early Childhood Research Quarterly* 21(1): 76-92. doi: 10.1016/j.ecresq.2006.01.003

Vandell, D. L., and Wolfe, B. 2000. *Child Care Quality: Does It Matter and Does It Need to Be Improved?* Washington, D.C.: Department of Health and Human Services.

Vandell, D. 2004. "Early child care: The known and the unknown." *Merrill-Palmer Quarterly* 50(3): 387-414.

Wooldridge, J. 2002. *Introductory Econometrics: Modern Approach, 2ⁿᵈ edition*. Mason, OH: South-Western College Pub.

Zachrisson, H., Dearing, E., Lekhai, R., and Toppelberg, C. 2013. "Little Evidence That Time in Child Care Causes Externalizing Problems During Early Childhood in Norway." *Child Development* 84(4): 1152-1170.

Zaslow, M., Halle, T., Martin, L., Cabrera, N., Calkins, J., Pitzer, L., and Margie, N. G. 2006. "Child Outcome Measures in the Study of Child Care Quality." *Evaluation Review* 30(5): 577-610. doi: 10.1177/0193841x06291529

Figure 1: Conceptual Framework of Preschool Quality on Child Development
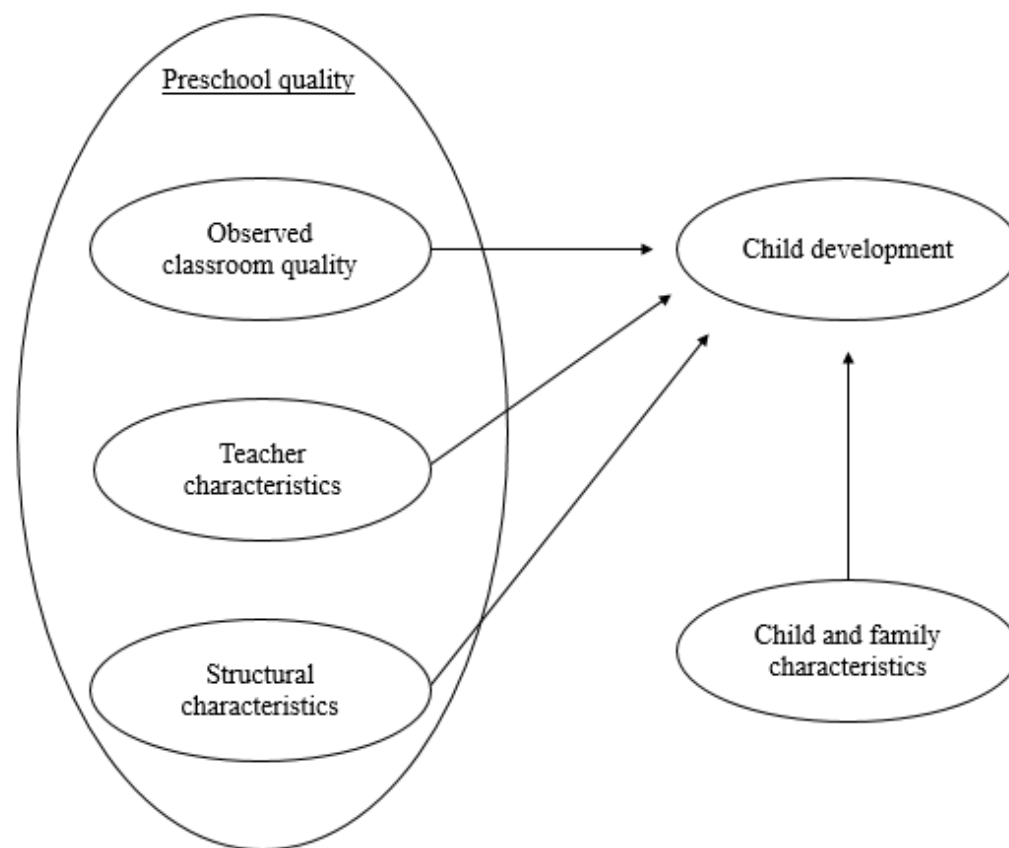
Table 1. Types of early childhood programs in Indonesia

Type	Name	Responsible Ministry
Pre-primary/Kindergarten (typically ages 5-6)	Kindergarten (*Taman Kanak-kanak, TK*) & Islamic kindergarten (*Radhatul Athal, RA*)	Ministry of Education and Culture Ministry of Religious Affairs
Pre-primary/Playgroup (typically ages 3-5)	Playgroup (*Kelompok Bermain, KB*)	Ministry of Education and Culture
Care service for children of working parents	Daycare (*Taman Penitipan Anak, TPA*)	Ministry of Social Welfare & Ministry of Education and Culture
Health care service for children and parenting information for mothers	Integrated health service unit (*Posyandu*)	Ministry of Health
Parenting classes for mothers	Toddler family groups (*Bina Keluarga Balita, BKB*)	National Family Planning Board

Note: This is a non-exhaustive list to show some of the most common types of early childhood programs in Indonesia.
Source: Hasan, Hyson and Chang, 2013

Table 2. Summary statistics of child-level variables by type of ECED

	Kindergarten		Islamic Kindergarten		Non-Project Playgroup		Project Playgroup		Total	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Child development outcomes										
EDI: Physical health & well-being	8.234	1.375	8.228	1.430	7.927	1.498	7.996	1.478	8.101	1.442
EDI: Social competence	6.927	1.724	6.812	1.627	6.755	1.683	6.634	1.679	6.778	1.698
EDI: Emotional maturity	6.532	1.433	6.274	1.245	6.099	1.316	6.320	1.336	6.373	1.373
EDI: Language & cognitive	7.229	2.456	7.240	2.534	6.285	2.565	6.189	2.653	6.696	2.607
EDI: Communication & general knowledge	6.188	2.114	6.193	1.995	6.033	2.030	5.773	2.020	6.002	2.065
Child characteristics										
Age (years)	5.124	0.723	5.010	0.948	4.407	1.014	4.331	1.155	4.707	1.043
Female (1=Yes)	0.504	0.500	0.520	0.500	0.513	0.500	0.500	0.500	0.505	0.500
Special needs (1=Yes)	0.038	0.192	0.056	0.229	0.023	0.151	0.037	0.190	0.038	0.190
Mother's highest education level:										
Primary or less	0.378	0.485	0.410	0.492	0.408	0.492	0.430	0.495	0.405	0.491
Junior secondary	0.277	0.448	0.267	0.443	0.271	0.444	0.263	0.441	0.270	0.444
Senior secondary	0.267	0.442	0.225	0.418	0.239	0.427	0.250	0.433	0.253	0.435
Post secondary	0.078	0.269	0.099	0.299	0.082	0.275	0.056	0.230	0.071	0.258

Note: EDI scores range from 1-10. All data are for children who are interviewed while enrolled in the type of center noted in the column heading.

Table 3. Summary statistics of center-level variables by type of ECED service

|	Kindergarten (N=221)		Islamic kindergarten (N=50)		Non-project playgroup (N=70)		Project playgroup (N=236)		Total (N=578)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Observed classroom quality										
ECERS-R: Space & furnishing	2.915	1.223	2.636	0.993	2.522	1.075	3.235	1.251	2.974	1.223
ECERS-R: Personal care & routine	2.534	1.113	2.195	0.811	2.435	0.990	2.612	1.091	2.524	1.070
ECERS-R: Language-reasoning	3.762	1.593	3.023	1.131	3.204	1.401	3.518	1.215	3.531	1.404
ECERS-R: Activities	2.429	0.974	1.920	0.689	2.122	0.924	2.752	0.893	2.480	0.951
ECERS-R: Interactions	4.076	1.581	3.926	1.305	3.849	1.559	4.158	1.436	4.069	1.498
ECERS-R: Program structure	2.743	1.470	2.178	1.138	2.510	1.422	2.883	1.314	2.723	1.387
ECERS-R: Parents & staff	2.784	0.961	2.387	0.821	2.473	1.065	2.465	0.867	2.582	0.938
ECERS-R: Total (mean) score	3.035	1.049	2.609	0.725	2.731	0.995	3.089	0.911	2.983	0.974
									(N=566 centers)	
Teacher characteristics										
Highest edu. level: Primary	0.000	0.000	0.000	0.000	0.014	0.084	0.006	0.051	0.004	0.044
Highest edu. level: Junior secondary	0.005	0.067	0.017	0.084	0.029	0.145	0.019	0.090	0.014	0.091
Highest edu. level: Senior secondary	0.347	0.402	0.502	0.393	0.684	0.392	0.742	0.342	0.563	0.417
Highest edu. level: Post secondary	0.648	0.403	0.481	0.402	0.273	0.388	0.233	0.328	0.419	0.418
Mean years of teaching	9.815	6.392	8.249	5.429	5.799	4.372	5.189	2.614	7.300	5.273
Teacher's prior experience is:										
ECED teacher experience	0.425	0.437	0.330	0.422	0.334	0.423	0.245	0.389	0.333	0.422
ECED teacher & non-teacher exp.	0.142	0.309	0.163	0.324	0.133	0.310	0.132	0.308	0.139	0.309
ECED non-teacher experience	0.171	0.333	0.152	0.269	0.173	0.316	0.318	0.403	0.229	0.363
No ECED experience	0.262	0.390	0.355	0.414	0.361	0.427	0.304	0.396	0.299	0.400
Teacher's training is:										
No training	0.164	0.317	0.323	0.398	0.144	0.287	0.117	0.255	0.157	0.303
Non-project training	0.834	0.317	0.677	0.398	0.856	0.287	0.259	0.347	0.592	0.433
100 hours of Project training	0.000	0.000	0.000	0.000	0.000	0.000	0.101	0.265	0.040	0.175
200 hours of Project training	0.002	0.022	0.000	0.000	0.000	0.000	0.522	0.412	0.210	0.366
Structural characteristics of center										
Student-to-staff ratio	12.162	5.788	9.419	5.152	9.182	4.821	10.214	4.383	10.769	5.199
Hours of operation per week	15.197	2.163	14.790	3.077	11.289	3.810	9.834	3.792	12.508	4.063

Note: N represents number of centers. ECERS-R scores are the averages of two observers who each rated the same centers at the same time using the stop-score rule. All teacher characteristics are averaged at the center-level and thus, continuous. Each variable is constructed by dividing the # of observed teachers in a center who have the particular characteristic by the total number of observed teachers in that center.

Figure 2: Average ECERS-R across various settings
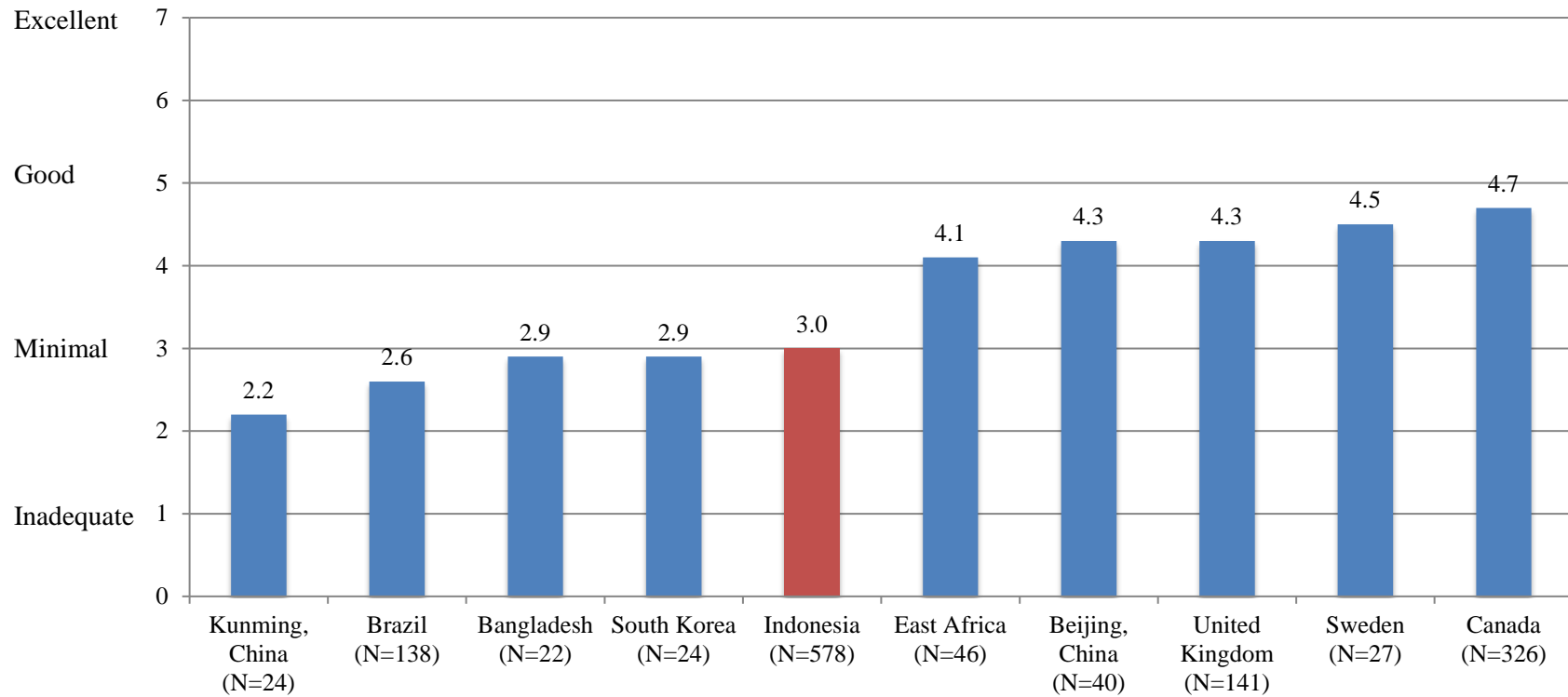
N is the number of classrooms or ECED centers observed

Sources: Authors' calculations for Indonesia, Aboud (2006) for Bangladesh, Esposito et al. (2010) for Brazil, Goelman et al. (2006) for Canada, Liang et al (2013) for Kunming and Beijing, Malmberg et al. (2009) for East Africa, Sheridan et al. (2009) for South Korea and Sweden, Sylva et al. (2009) for United Kingdom.

Table 4. Cross-walk of ECERS-R items and Indonesia ECED Standard

ECERS-R items	Is the ECERS-R item discussed in the Indonesia ECED Standard (2009)?	Citation from the Indonesia ECED Standard (2009)[a]
Space and Furnishing		
1. Indoor space	Yes	Section IV. A.2
2. Furniture for routine care, play and learning		
3. Furnishings for relaxation and comfort		
4. Room arrangement for play		
5. Space for privacy		
6. Child-related display		
7. Space for gross motor play	Yes	Section IV.A.2
8. Gross motor equipment	Yes	Section IV.A.2
Personal Care Routines		
9. Greeting/departing		
10. Meals/snacks	(No, only for daycare)	Section IV.A.2
11. Nap/rest	(No, only for daycare)	Section IV.A.2
12. Toileting/diapering	Yes	Section IV.A.2
13. Health practices	Yes	Section III.B.2
14. Safety practices	Yes	Section III.B.2
Language Reasoning		
15. Books and pictures	Yes	Section I.B.2,3
16. Encouraging children to communicate	Yes	Section II.A.2.b.3
17. Using language to develop reasoning skills	Yes	Section I.B.2-4
18. Informal use of language		
Activities		
19. Fine motor	Yes	Section I.B.1-4
20. Art	Yes	Section I.B.3
21. Music/movement	Yes	Section I.B.2-4
22. Blocks	Yes	Section I.B.2-4
23. Sand/water	Yes	Section II.B.3-4
24. Dramatic play		
25. Nature/science	Yes	Section I.B.4
26. Math/number	Yes	Section I.B.2-4
27. Use of TV, video and/or computers		
28. Promoting acceptance of diversity		
Interaction		
29. Supervision of gross motor activities	Yes	Section II.A.2.b.3
30. General supervision of children	Yes	Section II.A.2.b.3
31. Discipline	Yes	Section II.A.2.b.3
32. Staff-child interactions	Yes	Section II.A.2.b.3,4
33. Interactions among children		
Program Structure		
34. Schedule	Yes	Section III.A.3
35. Free play	Yes	Section III.A
36. Group time	Yes	Section III.A
37. Provisions for children with disabilities	Yes	Section III.C.5
Parents and Staff		
38. Provisions for parents	Yes	Section II.A.2.b.2
39. Provisions for personal needs of staff		
40. Provisions for professional needs of staff		
41. Staff interaction and cooperation	Yes	Section II.A.2.b.4
42. Supervision and evaluation of staff	Yes	Section III.C
43. Opportunities for professional growth	Yes	Section III.C.3

Note: [a] Citation of Indonesia's national standards are coded as follows:

I. Standards for Development Achievement   II. Standards for Educators and Education Personnel
III. Standards for Content, Process, and Assessment   IV. Standards for Facility and Infrastructure, Management, and Financing

Table 5. Mean and standard deviation of Indonesia standard

	Indonesia standard[a]	
	Mean	S.D.
Kindergarten	2.944	1.041
Islamic kindergarten	2.521	0.756
Non-project playgroup	2.659	1.007
Project playgroup	3.090	1.002
Overall	2.932	1.014

Note: [a] The Indonesia ECED standard is an alternative measure of observed classroom quality that only includes the 28 common items between the ECERS-R and the national standards for ECED.

Table 6. Analysis of classroom quality, teacher, and structural characteristics on children's physical health & well-being

	Physical health & well-being (Raw score)		
	OLS	IV	IV & Indonesia
	(1)	(2)	(3)
Observational classroom quality			
ECERS-R	0.083*	0.107*	
	(0.050)	(0.059)	
Indonesia standard			0.107*
			(0.058)
Teacher characteristics			
% of teachers w/ sr. secondary edu.	0.297	0.283	0.288
	(0.381)	(0.386)	(0.383)
% of teachers w/ post secondary edu.	0.285	0.278	0.285
	(0.390)	(0.395)	(0.392)
Mean years of teaching	-0.014*	-0.015**	-0.015**
	(0.008)	(0.008)	(0.008)
% of teachers w/ no ECED experience	-0.046	-0.057	-0.058
	(0.099)	(0.099)	(0.099)
% of teachers w/ no training	0.018	0.017	0.019
	(0.134)	(0.134)	(0.133)
Structural characteristics			
Student-to-staff ratio	0.013*	0.013*	0.013*
	(0.007)	(0.007)	(0.007)
Hours of operation per week	0.008	0.008	0.007
	(0.012)	(0.012)	(0.012)
Constant	3.906***	3.860***	3.876***
	(0.512)	(0.522)	(0.518)
Observations	7,984	7,984	7,984
R-squared	0.157	0.155	0.157
No. of clusters (centers)	565	565	565
First-stage F-statistic of excluded instrument		647.760	842.486
Control included:			
Child characteristics	Y	Y	Y
Kindergarten dummy	Y	Y	Y
District dummy	Y	Y	Y

Note: Robust standard errors clustered at the center level in parentheses. Child characteristics include age, age-squared, gender, whether child has special needs (reported by teacher), and mother's highest education level. Kindergarten dummy is 1 for kindergartens or Islamic kindergartens and 0 for non-project playgroups or project playgroups. District dummies are included for the 8 districts where centers and children are located. Model 1 is OLS regression model. Models 2 and 3 are two-stage least squares model to correct for measurement error in the measure of center quality. The F-statistic of the first stage is reported.

* $p<0.1$,   ** $p<0.05$, *** $p<0.01$

Table 7. Analysis of classroom quality, teacher, and structural characteristics on children's social competence

	Social competence (raw score)		
	OLS	IV	IV & Indonesia
	(1)	(2)	(3)
Observational classroom quality			
ECERS-R	0.070	0.111*	
	(0.052)	(0.063)	
Indonesia standard			0.121**
			(0.062)
Teacher characteristics			
% of teachers w/ sr. secondary edu.	0.826**	0.808**	0.813**
	(0.336)	(0.340)	(0.342)
% of teachers w/ post secondary edu.	1.086***	1.073***	1.078***
	(0.351)	(0.356)	(0.358)
Mean years of teaching	-0.006	-0.007	-0.008
	(0.009)	(0.009)	(0.009)
% of teachers w/ no ECED experience	0.067	0.055	0.053
	(0.126)	(0.126)	(0.125)
% of teachers w/ no training	-0.110	-0.111	-0.110
	(0.156)	(0.156)	(0.155)
Structural characteristics			
Student-to-staff ratio	0.001	0.002	0.002
	(0.009)	(0.009)	(0.009)
Hours of operation per week	0.028*	0.028*	0.027*
	(0.016)	(0.015)	(0.015)
Constant	2.257***	2.188***	2.192***
	(0.556)	(0.565)	(0.562)
Observations	7,957	7,957	7,957
R-squared	0.144	0.141	0.143
No. of clusters (centers)	565	565	565
First-stage F-statistic of excluded instrument		646.386	842.505
Control included:			
Child characteristics	Y	Y	Y
Kindergarten dummy	Y	Y	Y
District dummy	Y	Y	Y

Note: Robust standard errors clustered at the center level in parentheses. Child characteristics include age, age-squared, gender, whether child has special needs (reported by teacher), and mother's highest education level. Kindergarten dummy is 1 for kindergartens or Islamic kindergartens and 0 for non-project playgroups or project playgroups. District dummies are included for the 8 districts where centers and children are located. Model 1 is OLS regression model. Models 2 and 3 are two-stage least squares model to correct for measurement error in the measure of center quality. The F-statistic of the first stage is reported.
* $p<0.1$,   ** $p<0.05$, *** $p<0.01$

Table 8. Analysis of classroom quality, teacher, and structural characteristics on children's emotional maturity

	Emotional maturity (raw score)		
	OLS	IV	IV & Indonesia
	(1)	(2)	(3)
Observational classroom quality			
ECERS-R	0.063	0.072	
	(0.051)	(0.062)	
Indonesia standard			0.088
			(0.062)
Teacher characteristics			
% of teachers w/ sr. secondary edu.	0.191	0.183	0.184
	(0.411)	(0.415)	(0.413)
% of teachers w/ post secondary edu.	0.271	0.268	0.269
	(0.431)	(0.434)	(0.433)
Mean years of teaching	0.010	0.009	0.009
	(0.008)	(0.008)	(0.008)
% of teachers w/ no ECED experience	0.063	0.055	0.052
	(0.101)	(0.100)	(0.100)
% of teachers w/ no training	0.113	0.113	0.114
	(0.129)	(0.129)	(0.128)
Structural characteristics			
Student-to-staff ratio	0.000	0.000	0.001
	(0.007)	(0.007)	(0.007)
Hours of operation per week	0.001	0.002	0.000
	(0.013)	(0.013)	(0.012)
Constant	3.812***	3.792***	3.781***
	(0.502)	(0.504)	(0.502)
Observations	7,946	7,946	7,946
R-squared	0.118	0.117	0.119
No. of clusters (centers)	565	565	565
First-stage F-statistic of excluded instrument		647.316	842.323
Control included:			
Child characteristics	Y	Y	Y
Kindergarten dummy	Y	Y	Y
District dummy	Y	Y	Y

Note: Robust standard errors clustered at the center level in parentheses. Child characteristics include age, age-squared, gender, whether child has special needs (reported by teacher), and mother's highest education level. Kindergarten dummy is 1 for kindergartens or Islamic kindergartens and 0 for non-project playgroups or project playgroups. District dummies are included for the 8 districts where centers and children are located. Model 1 is OLS regression model. Models 2 and 3 are two-stage least squares model to correct for measurement error in the measure of center quality. The F-statistic of the first stage is reported.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 9. Analysis of classroom quality, teacher, and structural characteristics on children's language and cognitive development

	Language & cognitive development (raw score)		
	OLS	IV	IV & Indonesia
	(1)	(2)	(3)
Observational classroom quality			
ECERS-R	0.102	0.132	
	(0.081)	(0.096)	
Indonesia standard			0.167*
			(0.092)
Teacher characteristics			
% of teachers w/ sr. secondary edu.	-0.109	-0.127	-0.127
	(0.550)	(0.544)	(0.538)
% of teachers w/ post secondary edu.	-0.170	-0.180	-0.180
	(0.573)	(0.568)	(0.563)
Mean years of teaching	-0.003	-0.004	-0.006
	(0.012)	(0.012)	(0.012)
% of teachers w/ no ECED experience	0.255	0.241	0.236
	(0.170)	(0.169)	(0.170)
% of teachers w/ no training	0.042	0.040	0.041
	(0.217)	(0.217)	(0.215)
Structural characteristics			
Student-to-staff ratio	0.016	0.017	0.017
	(0.011)	(0.011)	(0.011)
Hours of operation per week	0.039*	0.039*	0.036*
	(0.020)	(0.020)	(0.020)
Constant	-1.437*	-1.496*	-1.524*
	(0.786)	(0.785)	(0.778)
Observations	7,958	7,958	7,958
R-squared	0.270	0.269	0.270
No. of clusters (centers)	565	565	565
First-stage F-statistic of excluded instrument		663.973	862.420
Control included:			
Child characteristics	Y	Y	Y
Kindergarten dummy	Y	Y	Y
District dummy	Y	Y	Y

Note: Robust standard errors clustered at the center level in parentheses. Child characteristics include age, age-squared, gender, whether child has special needs (reported by teacher), and mother's highest education level. Kindergarten dummy is 1 for kindergartens or Islamic kindergartens and 0 for non-project playgroups or project playgroups. District dummies are included for the 8 districts where centers and children are located. Model 1 is OLS regression model. Models 2 and 3 are two-stage least squares model to correct for measurement error in the measure of center quality. The F-statistic of the first stage is reported.

* $p<0.1$,   ** $p<0.05$, *** $p<0.01$

Table 10. Analysis of classroom quality, teacher, and structural characteristics on children's communication & general knowledge

	Communication & general knowledge (raw score)		
	OLS	IV	IV & Indonesia
	(1)	(2)	(3)
Observational classroom quality			
ECERS-R	0.080	0.120	
	(0.062)	(0.075)	
Indonesia standard			0.133*
			(0.072)
Teacher characteristics			
% of teachers w/ sr. secondary edu.	0.926**	0.908*	0.911*
	(0.470)	(0.476)	(0.481)
% of teachers w/ post secondary edu.	1.196**	1.183**	1.188**
	(0.477)	(0.484)	(0.489)
Mean years of teaching	-0.043***	-0.045***	-0.045***
	(0.012)	(0.012)	(0.012)
% of teachers w/ no ECED experience	-0.009	-0.023	-0.025
	(0.139)	(0.139)	(0.138)
% of teachers w/ no training	0.073	0.072	0.073
	(0.178)	(0.178)	(0.178)
Structural characteristics			
Student-to-staff ratio	-0.005	-0.004	-0.004
	(0.012)	(0.012)	(0.012)
Hours of operation per week	0.005	0.005	0.004
	(0.018)	(0.018)	(0.018)
Constant	0.084	0.015	0.014
	(0.675)	(0.684)	(0.682)
Observations	7,981	7,981	7,981
R-squared	0.151	0.149	0.150
No. of clusters (centers)	565	565	565
First-stage F-statistic of excluded instrument		647.668	842.419
Control included:			
Child characteristics	Y	Y	Y
Kindergarten dummy	Y	Y	Y
District dummy	Y	Y	Y

Note: Robust standard errors clustered at the center level in parentheses. Child characteristics include age, age-squared, gender, whether child has special needs (reported by teacher), and mother's highest education level. Kindergarten dummy is 1 for kindergartens or Islamic kindergartens and 0 for non-project playgroups or project playgroups. District dummies are included for the 8 districts where centers and children are located. Model 1 is OLS regression model. Models 2 and 3 are two-stage least squares model to correct for measurement error in the measure of center quality. The F-statistic of the first stage is reported.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$