

Effects of water and health on primary school enrolment and absenteeism in Indonesia

Ahmad Komarulzaman, Eelke de Jong and Jeroen Smits

ABSTRACT

Clean water provision remains a serious problem in low- and middle-income countries. In 2017, about 30% of the world population relied on unimproved water sources located outside of the dwellings. Often women and children are occupied in fetching water. This situation increases the prevalence of water-related diseases such as diarrhoea and reduces children's study time. School attendance may decrease due to the combined effects of diarrhoea and time spent on fetching water. We investigate the effects on school absenteeism and primary school enrolment in Indonesia, using a panel data set for 295 districts over the period 1994–2014. Districts with higher diarrhoea prevalence are found to have lower school enrolment ($B: -0.202$, sig $p < 0.01$) and higher school absenteeism ($B: 2.334$, sig $p < 0.001$). Districts where more households have access to private water facilities have higher school enrolment ($B: 0.025$, sig $p < 0.01$) and lower school absenteeism ($B: -0.027$, sig $p < 0.01$). More use of piped and bottled water in a district is associated with a lower diarrhoea prevalence ($B: -0.004$, sig $p < 0.05$). Policy-makers should take the benefits of improved water supply into account when making cost-benefit analyses regarding investments in water infrastructure.

Key words | diarrhoea, drinking water, Indonesia, panel, school absenteeism, school attendance

Ahmad Komarulzaman (corresponding author)
Department of Economics, Faculty of Economics
and Business,
Universitas Padjadjaran,
Bandung 40132,
Indonesia
E-mail: ahmad.komarulzaman@fe.unpad.ac.id

Ahmad Komarulzaman
Eelke de Jong
Jeroen Smits
Institute for Management Research,
Economics Department,
Radboud University,
Nijmegen,
The Netherlands

This article has been made Open Access thanks to the generous support of a global network of libraries as part of the Knowledge Unlatched Select initiative.

INTRODUCTION

The provision of clean water remains a serious problem, especially in developing countries. Globally, 2.1 billion people (11% of the world population) lack safe water at the premises (WHO & UNICEF 2017). The unsafe sources are often located far from the house so that some members of the household (mainly women and children) have to invest a lot of time walking and waiting. They may have to walk as far as 6 km to fetch water, which may take up to 60 min a day (Sorenson *et al.* 2011). This leads to two essential problems: water contamination and wasted time. The risk of contamination is high because the water source is often unprotected and the water is transported over a long distance in barrels or tanks that are not completely clean. Consequently, this

may increase the prevalence of enteric diseases causing diarrhoea, which often damages children's health (Wolf *et al.* 2014; Komarulzaman *et al.* 2017). Furthermore, as children are frequently involved in helping their family to fetch water, they might lose precious study time and attend school less often (Dreibelbis *et al.* 2013; Nauges & Strand 2013). Thus, school attendance may decrease due to the combined effects of diarrhoea disruption and fetching water.

A large and growing body of literature has reported on the association between clean water provision and diarrhoea morbidity and mortality (Wolf *et al.* 2014; Komarulzaman *et al.* 2017) as well as the impact of water and sanitation improvement on school attendance (Dreibelbis *et al.* 2013; Nauges & Strand 2013; Hunter *et al.* 2014). However, as far as we are aware, no empirical investigation has studied these three aspects simultaneously. Given the link between water provision and school attendance as

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

doi: 10.2166/wh.2019.044

well as the link between water quality and diarrhoea, it seems important to incorporate diarrhoea prevalence in the analysis when studying the effects of clean water access on school attendance.

The current study aims to investigate both the direct effect of clean water access on school attendance as well as its indirect effect that runs through diarrhoea prevalence. We consider the case of Indonesia, where water access and diarrhoea as well as school attendance are still major problems. The provision of piped water is limited (covering 20% of the population) and unevenly distributed over the country. In relation to this, Indonesia needs to cope with diarrhoea, which is the third major cause of child mortality (UNICEF 2012). Lastly, even though primary school enrolment in Indonesia reached 95% in 2014, this figure does not capture the true reality of school participation as about 12.5% of the students have been absent from school due to illnesses (Statistics Indonesia 2014). Providing a more comprehensive understanding of the relationship between access to clean water and school attendance might support policy-makers in finding more effective interventions.

A THEORETICAL FRAMEWORK OF DIRECT AND INDIRECT EFFECTS OF WATER ON SCHOOL ATTENDANCE

The demand for primary education can originate from policy measures, which oblige children to attend school in order to acquire the basic skills needed for economic participation (Checchi 2006). In Indonesia, the 1990 basic education law regulates 9 years of compulsory education for all children aged 7–15 years. However, this top-down approach is not sufficient to explain the variation in primary school attendance. Education can also be seen as an investment in human capital made by individuals to acquire future benefits (Eide & Showalter 2010; Todaro & Smith 2012). Based on this view, school attendance depends on both the benefits and costs of education accrued by the individual or the population. The costs of education include the direct costs of tuition fees, books, and transport costs, and the indirect or opportunity costs of forgone income (Eide & Showalter 2010; Todaro & Smith 2012). In low-income

countries, parents who decide on their children's education tend to be more myopic than the ones in high-income countries and base their choices more on current costs than on future gains, resulting in a lower demand for education (Checchi 2006). Other factors influencing the optimal demand for education are talent (i.e. personal intelligence or family background), current and expected future gains (such as employment conditions and expectation of higher returns in the future), the initial level of human capital, and the available resources for education in the region.

Based on these views, there are at least two channels through which water access influences school attendance. First, households with no access to water in the dwelling have to transport the water into the house from outside, a job usually carried out by women and children. This water-fetching role can be a burden on the children, who lose energy and have less time to study (Dreibelbis *et al.* 2013; Nauges & Strand 2013). In such a situation, the parents might consider the current costs of sending their children to school as being higher than expected future benefits. Second, the lack of access to clean water increases the probability of diarrhoea (Wolf *et al.* 2014; Komarulzaman *et al.* 2017) and other hygiene issues among children, which might further reduce school attendance (Neuzil *et al.* 2002; Azor-Martínez *et al.* 2014). Thus, sub-optimal water conditions may force children to be absent from school (Dreibelbis *et al.* 2013) or at worst drop out from school completely (Nauges & Strand 2013).

Figure 1 gives an overview of the relations between water access, diarrhoea and other factors that affect school attendance. The solid lines indicate the most important relations, whereas the dashed lines (the diarrhoea prevalence model) indicate relations that are of secondary importance, as they indirectly influence school attendance. This figure forms the theoretical framework of empirical analysis.

DATA

Susenas data set

This study uses a district panel data set constructed on the basis of the annual Indonesia National Socio-Economic Survey (Susenas) data sets from 1994 to 2014. Susenas

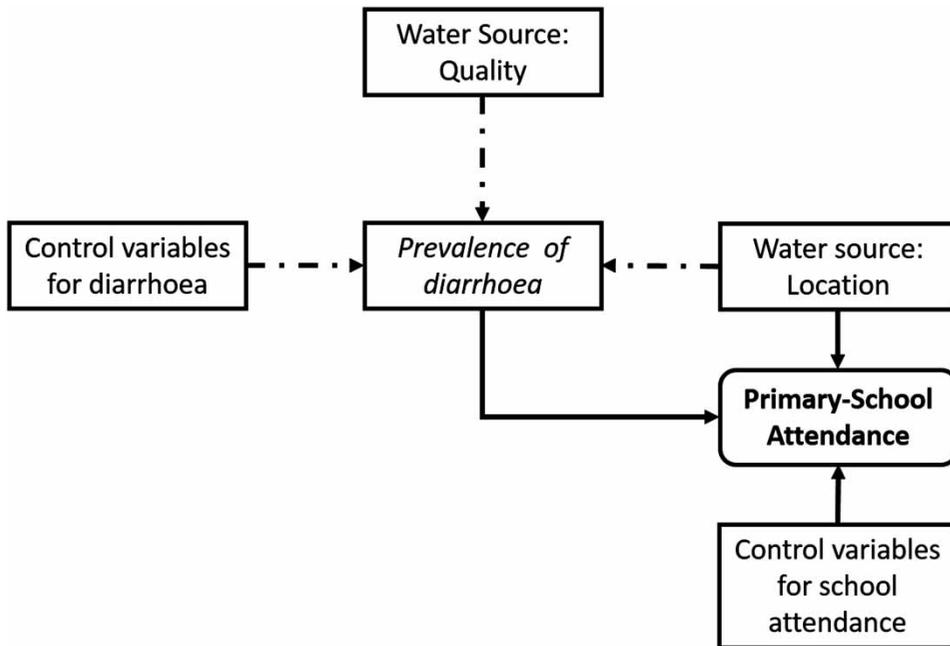


Figure 1 | Theoretical model of water, diarrhoea and school attendance.

was initiated in the period of 1963–1964 and is composed of two questionnaires. The first is a core questionnaire which contains household characteristics and household members' information on age, sex, education, health, and working activities. The core questionnaire is fielded to about 200,000–286,000 households and 0.7–1.1 million household members. With this sample size, the core data are representative at national, provincial, and district level. The survey is supplemented by a module questionnaire that collects additional information on consumption, expenditure, socio-cultural characteristics and education, as well as on health and housing. The module is fielded to about 65,000 households and representative at the national and provincial levels.

The Susenas sampling frame is prepared in accordance with the recent population census, e.g. the sampling frame for the 2000's Susenas originated from the 2000 population census. The Susenas sample is selected using a three-phase sampling design with two strata (urban and rural areas) in each district. In Phase 1, a number of census blocks are selected from the sample frame of census blocks using the probability proportional to the size (PPS) method. Size is the number of households in each census block in the

sampling frame. In the second phase, one segment group is selected from each census block with the PPS method. Households listing is then conducted for all selected segment groups. Lastly, in each enumeration area, 16 households are selected by systematic linear sampling. Given this sampling procedure, Susenas is the only annual socio-economic data source that is representative at the district level in Indonesia.

The annual cross-sectional data from the core questionnaire are used to construct a balanced district panel of 259 of the 291 Indonesian districts between 1994–2014. The 32 districts that were not included are located in the provinces of Aceh, Sulawesi Tenggara, Maluku, Papua and Papua Barat. These districts were not surveyed in one or more years due to conflicts in the regions. In the current study, the district code is based on the 1994 geographical definition of the districts that were taken as the basic definition to accommodate the splitting up of districts during the period. Thus, the split districts were merged into their original parent district. Details of the procedure for generating homogenous districts over time can be obtained from the authors. All district-level variables were aggregated from the individual and household data.

Long-term trend of piped water coverage, diarrhoea prevalence, and school attendance

This section presents a descriptive analysis of the long-term trend in the main variables of this analysis, namely primary school attendance, clean water access, and diarrhoea prevalence among school-aged children in 259 districts in Indonesia over the period 1994–2014. Primary school attendance is measured by two variables, namely (1) school enrolment, which is the percentages of children aged 7–15 enrolled in primary school and (2) school absenteeism, which is measured by the percentage of primary school students who, during the past month, were absent from school due to health complaints for at least one day. Good quality water is captured by the percentage of households with access to piped and bottled water. Nearby (location) water source is measured by the percentage of households having access to a private water facility as opposed to sources that are shared or public. This procedure is similar to the one used in [Mangyo \(2008\)](#), who uses the water access within the yard as the physical proximity of a drinking water source to the dwelling. Diarrhoea prevalence

is measured by the percentage of school-aged children who suffered from diarrhoea in the past month (preceding the Susenas survey).

Primary school enrolment shows a clear increasing trend for Indonesia as a whole as well as for the different regions ([Figure 2](#)). At the national level, primary school enrolment increased 8.5 percentage points from 86.5% in 1994 to 95.0% in 2014. [Figure 2](#) also shows disparities in school attendance across regions of Indonesia. The districts in Nusa Tenggara (NT), Kalimantan, and Sulawesi islands (which are all in eastern Indonesia) are lagging behind those in Sumatera, Java and Bali islands (which are situated in western Indonesia).

To get more insight into actual school attendance, [Figure 3](#) shows the absenteeism rate due to health complaints in the districts as well as its average by island and nationally. Nationally, on average, 12.9% of the students was absent (for at least one day) from school due to health complaints. In 1994, about 9.4% of students missed school days, and there is hardly any improvement over the following two decades, as in 2014, even 12.5% of the students were absent from school.

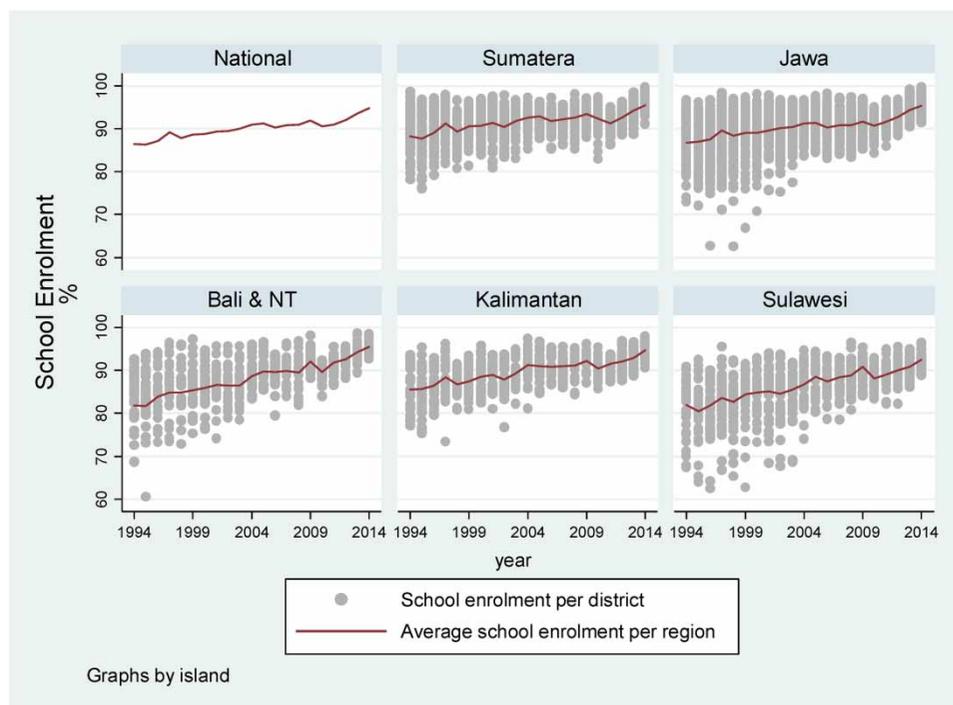


Figure 2 | Percentage of primary school enrolment in districts in Indonesia, by islands and total, 1994–2014.

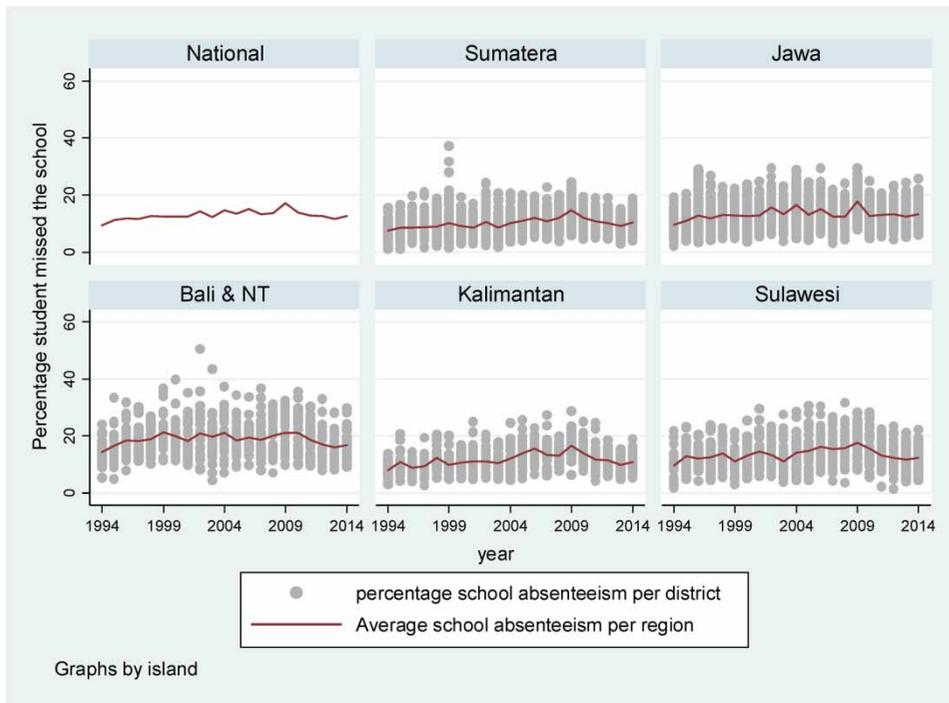


Figure 3 | Percentage of primary school students absent from school due to health complaints (from the enrolled student), by island and total, 1994–2014.

The most important explanatory variables of school attendance are the quality and location of the water source and the prevalence of diarrhoea. An indication of the development of the quality of water is provided in Figure 4, which shows the distribution of the main drinking water sources at the national as well as island level from 1994 to 2014.

In general, water provision in Indonesia improved significantly over the last two decades. Nationally, the percentage of households relying on safe water sources is increasing. Regarding the source of improved water, we see an increase in the use of bottled water and a decrease in the coverage of piped water. This increase in bottled water and a decrease in piped water mainly took place in the big cities, such as Kota Surabaya (East Java province), all districts in the Jakarta Province, Kota Banda Aceh (Aceh Province), and Kota Bukit Tinggi (West Sumatera province). Kota Surabaya experienced the biggest fall in the use of piped water as the main source of drinking water: in 2014, the percentage of households using piped water was 72.7 points lower than in 1994. It seems that the ability of the government to provide a reliable piped water service cannot keep up with the rapid increase in

population. As a consequence, the consumption of bottled water has increased substantially. Only 83 districts (out of 259 districts) show an increase in piped water coverage. Kota Tegal (Central Java province) experienced the largest increase (26.5 percentage points over 21 years).

In 1994, about 50% of the households had access to a private source of drinking water (Figure 5). This proportion is steadily increasing over time and reached 71% in 2014. Meanwhile, the rest of the population has to obtain drinking water from a shared or public source. This means that these households need to walk from home to the source and queue there to get the water. Having a private source of drinking water is more common in Sumatera and Jawa islands. This is related to the relatively high coverage of piped and bottled water and urbanization at these two islands. The lowest percentages of households with a private drinking water source are found at Bali and NT islands. Here only 28% of households had access to a private water source in 1994 and 45% in 2014.

Figure 6 shows that diarrhoea prevalence among school-aged children fluctuates at a quite low level (less than 2%), over the years and between regions. Diarrhoea is most

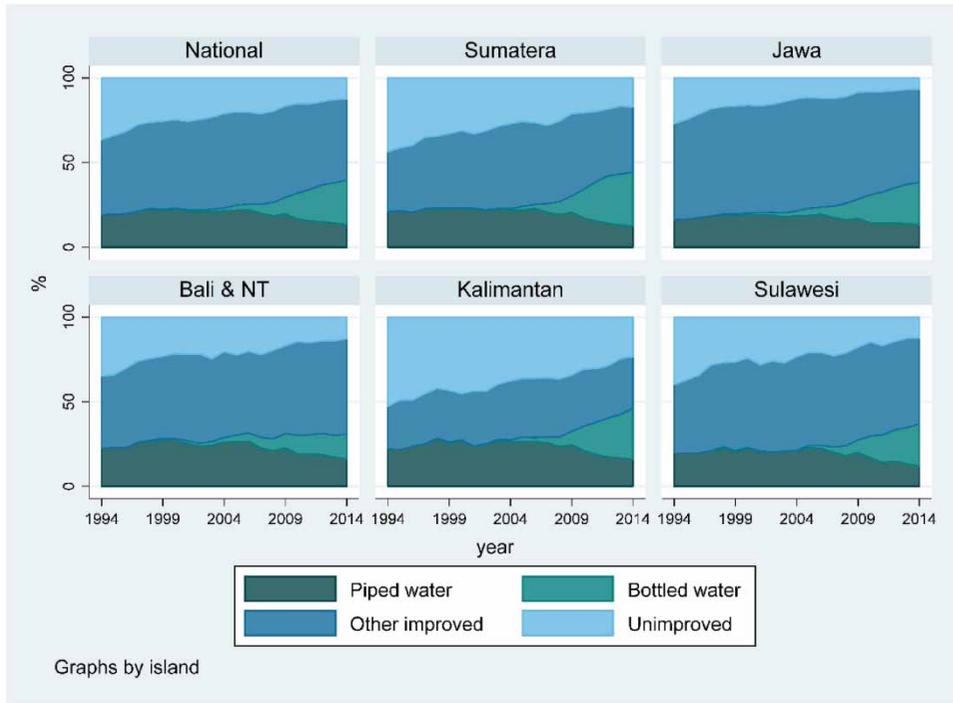


Figure 4 | Percentage of households by main drinking water sources in Indonesia 1994–2014.

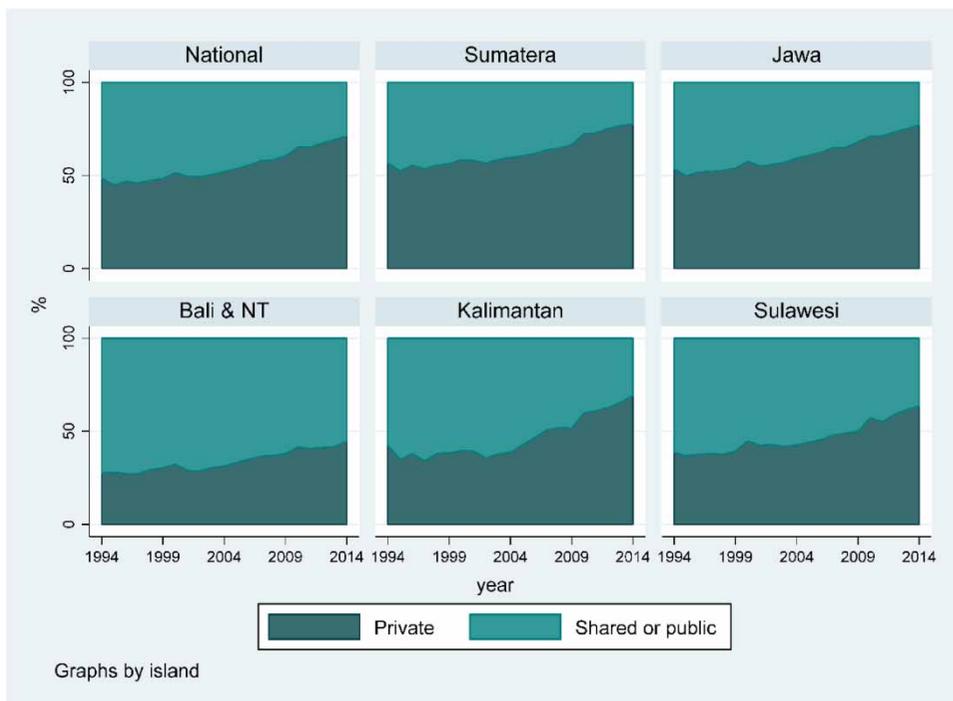


Figure 5 | Percentage of households by the type of main drinking water sources in Indonesia 1994–2014.

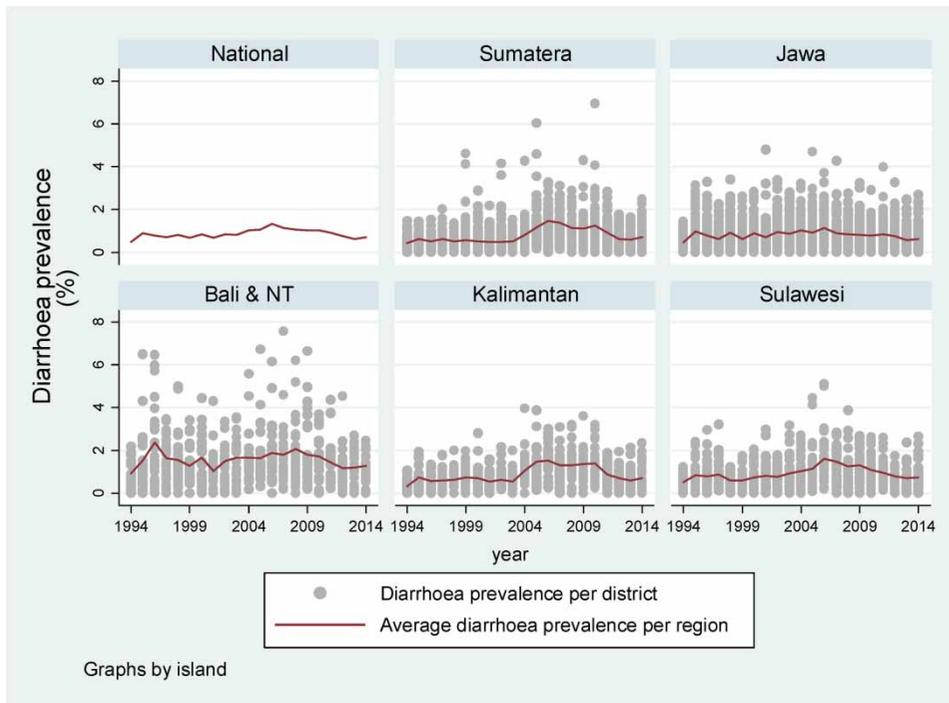


Figure 6 | Diarrhoea prevalence among school-aged children in Indonesia, by island and Nationally, 1994–2014.

prominent in Bali and NT islands, specifically in the Nusa Tenggara Timur province. The rest of the provinces in Indonesia experienced a considerably lower prevalence of diarrhoea. However, if we look at the district level, the picture is worse. The highest prevalence of diarrhoea, 7.6%, was experienced in the Sumba Timur District (Bali and NT islands) in 2007.

Independent variables

Following our conceptual framework that was explained previously, we include a set of control variables in our models. This section explains the theoretical foundation of each of these control variables in our model. Control variables are households' wealth, parents' education, urbanization, health facilities coverage, access to improved sanitation, as well as cultural factors. All control variables are at the district level.

Poverty is often a key reason for children not attending school (Checchi 2006; Huisman & Smits 2015). Wealthy households typically have more resources to support their children and to offer them good education. Poorer

households also have higher opportunity costs to keep their children at home as the latter can help the family or earn additional income. In addition, when financial markets are imperfect or absent, children from poor families are often constrained in getting credit and tend to acquire less education compared to children from richer families (Checchi 2006; Todaro & Smith 2012). These influences are captured by households' wealth, which is measured by the average monthly food expenditure in the district.

Parents often set their own education level as a benchmark for their children, so that children from a well-educated family have a higher probability of achieving a higher level of education (Glick & Sahn 2000; Huisman & Smits 2015). For each district, parents' education is calculated as the average years of education of both father and mother.

Moreover, gender differences of parents' education level reflect the effect of cultural conditions in the region and the degree of the mother's position in decision-making processes. A more educated mother might have more power within the household, and this increases the chances of the children, especially girls, attending school (Glick & Sahn 2000). She

might also be better able to use the facilities for the benefit of her children. [Jalan & Ravallion \(2003\)](#), for example, found a larger health gain from piped water for children with better-educated mothers. This effect is even significant among poor households, probably because an educated mother may have a better understanding of how to get and treat clean water. She will also be more likely to practice better hygiene in the family, specifically for her children. Consequently, less prevalence of diarrhoea is found ([Mangyo 2008](#)). The position of the women (cultural factor) is measured by the ratio of the mother's number of years of education divided by the father's number of education years.

Lastly, children in an urban area might benefit from better education and other infrastructure, e.g. the number of schools, electricity, roads, water and sanitation facilities, which can reduce the direct costs of education, which in turn positively affects the rate of primary school attendance ([Nauges & Strand 2013](#)). We measure urbanization as the percentage of the population living in urban areas in each district.

Ample evidence suggests that children who lack access to safe drinking water and improved sanitation facilities have a bigger risk of getting diarrhoea ([Wolf *et al.* 2014](#); [Komarulzaman *et al.* 2017](#)). [Clasen & Haller \(2008\)](#), for example, found that about 94% of the disease burden of diarrhoea is attributed to unsafe water, lack of sanitation, and poor hygiene. Some of the most important indices of water which can show that the sources of water are harmful for drinking includes biological oxygen demand (BOD), chemical oxygen demand (COD), NH_4^+ , NO_2^- , NO_3^- and the faecal coliform contamination ([Sepehri & Sarrafzadeh 2018](#)). Furthermore, [UNICEF & WHO \(2009\)](#) found that the incidences of diarrhoea could be decreased by one-third to one-fourth through improved living conditions such as access to safe water, improved sanitation, and better hygiene. These effects are represented by two variables: (1) the access to improved sanitation is computed by the percentage of households that have a private toilet with septic tank and (2) the coverage of health facilities in a district as measured by the percentage of births assisted by a skilled healthcare worker such as doctor, midwife, or paramedic.

Table 1 presents the descriptive characteristics of the variables for 259 districts in Indonesia from 1994 to 2014. The average level of primary school enrolment is 90%,

Table 1 | Descriptive statistics of selected variables for 259 districts in Indonesia, 1994–2014

| Variables | Mean (level) | Mean (change) |
|--|--------------|---------------|
| Primary school enrolment (%) | 90.10 | 0.42 |
| Student absent from school (%) | 12.91 | 0.16 |
| Diarrhoea prevalence (7–15 years; %) | 0.88 | 0.01 |
| Household using piped and bottled water (%) | 25.69 | 1.03 |
| Household with private water facility (%) | 56.84 | 1.05 |
| Household with improved sanitation (%) | 36.06 | 1.66 |
| Health facilities coverage (%) | 66.26 | 1.88 |
| Average food expenditure (000,000 IDR) | 0.70 | 0.07 |
| Mean years of parents' education | 6.78 | 0.13 |
| Relative difference of mother and father education years | 0.89 | 0.01 |
| Household living in urban area (%) | 39.75 | 0.61 |

with an average annual change of 0.4 percentage points. On average, 13% of these students were absent from school due to health complaints. Over the period, almost 1% of the primary school students suffered from diarrhoea. Overall, about 26% of the households use piped and bottled water (good quality water source) as their main source for drinking water. Moreover, 57% of households have access to a nearby water source. The hygiene level is quite low as indicated by the low access to improved sanitation (36%) with a low rate of improvement (1.7 percentage points per year). The average coverage of health facilities is 66% and increases each year by almost 2 percentage points. Even though the average primary school enrolment is quite high, the average education level of the parents is still low, namely 6.8 years. On average, mothers' number of education years is 0.89 of fathers' number of education years. This indicates that the education level of mothers is lower than that of fathers, although the difference is, on average, not large. Lastly, about 40% of households are living in urban areas.

METHOD

Our model is based on the assumption that school attendance is a linear (we also experimented with quadratic

terms, but these proved to be insignificant) function of access to a nearby water source, diarrhoea prevalence, control variables, time trend, and district-level fixed effects. Moreover, we assume that diarrhoea prevalence is determined by the quality of water, the location of the water source, and a set of control variables. To explore how the effect of the main variables varies over time, we computed interactions between the main explanatory variables and the time trend.

To test our theoretical expectations, several fixed effects regression models are estimated (see Verbeek (2004), Ch. 10 and Carter Hill *et al.* (2012), Ch. 15). The first model takes diarrhoea prevalence in the district as the dependent variable and tests whether this prevalence is influenced by characteristics of the water sources in the district. The second model takes school enrolment in the district as the dependent variable and tests whether this enrolment is influenced by characteristics of the water sources and diarrhoea prevalence in the district. The third model is similar to the second model but with school absenteeism as the dependent variable. All models have two versions. The first contains only the main effects of the water and health variables, whereas the second version adds the interactions to the main effects.

Fixed effects models are used because we are interested in how changes in educational enrolment and absenteeism are related to changes in the quality of water, the location of water sources, and diarrhoea prevalence. Fixed effects models have the advantage over alternatives like random effects models that they make it possible to study these relations while completely controlling for all (measured and unmeasured) regional characteristics that are stable over time. One could argue that districts are not random drawings from a population but 'one of kind' and then fixed effects are preferred over random effects (Verbeek (2004), p. 351).

Two versions of fixed effects models are estimated: a static and a dynamic one. The static one implicitly assumes that adjustments take at most one period, in this case a year. In many cases that is not realistic. We, therefore, estimate a dynamic fixed effects model too. This model assumes that the actual observations gradually move towards the long-run equilibrium. An error correction specification represents this dynamic model (see, e.g. Carter Hill *et al.* (2012), Ch. 12 and 13). The absolute value of the so-called error correction

coefficient represents the speed towards the equilibrium; the closer the absolute value is to one, the faster the adjustment takes place. This model explicitly takes the development over time into account and assumes that its coefficients are constant. It can be estimated if the time series behave well, that is that they do not show regularities such as a time trend, and when the time series of the various variables remain in each other's neighbourhood. The first condition is known as the stationarity condition and the second as the cointegration condition. We test for both and find that the data confirm these conditions.

RESULTS

We first report the result of the static panel fixed effect model. Thereafter, those of the dynamic model are presented as a check on the robustness of the long-run relations. Note that only the effect of the variables of main interest (percentage of households using piped and bottled water, percentage of households with private water facility, and diarrhoea prevalence) is presented in the tables. The table with the full results can be found in the Appendix, available with the online version of this paper.

Static panel analysis

Table 2 presents the panel fixed effect analysis for diarrhoea (D1 and D2), school absenteeism (A1), and school enrolment (E1 and E2). Models D1, A1, and E1 include only the main effects, and models D2 and E2 include both the main and interaction effects.

The prevalence of diarrhoea (D1) is reduced if the percentage of households using piped and bottled water is higher in the district (coefficient estimated at -0.004 , sig $p < 0.05$). A district's effort to increase the percentage of households using piped and bottled water by, for example, 10% is potentially reducing diarrhoea prevalence in the given district by 0.04% – four times the average annual change of diarrhoea prevalence in Indonesia (0.01%, see Table 1). The estimated coefficient of the percentage of households with a private water facility is not significant. In addition, the model with the interaction terms (model D2) reveals that the protective effect of piped and bottled

Table 2 | Estimating the effect of access to water on diarrhoea prevalence (D1 and D2), school absenteeism (A1), and school enrolment (E1 and E2), 259 districts in Indonesia, 1994–2014

| | D1 B/SE | D2 B/SE | A1 B/SE | E1 B/SE | E2 B/SE |
|---|---------------------|----------------------|---------------------|---------------------|----------------------|
| <i>Main effects</i> | | | | | |
| Household (HH) using piped and bottled water (%) | −0.004 (0.002)* | −0.003 (0.002)* | | | |
| Diarrhoea prevalence (%) | | | 2.334 (0.117)*** | −0.202 (0.072)** | −0.263 (0.063)*** |
| HH with private water facility (%) | −0.000 (0.001) | 0.000 (0.001) | −0.027 (0.008)** | 0.025 (0.009)** | 0.012 (0.008) |
| Time trend | 0.052 (0.009)*** | 0.047 (0.009)*** | 0.375 (0.039)*** | 0.419 (0.052)*** | 0.298 (0.043)*** |
| <i>Interaction term</i> | | | | | |
| HH using piped and bottled water (%) × time trend | | −0.000 (0.000)*** | | | |
| Diarrhoea prevalence (%) × time trend | | | | | 0.035 (0.015)* |
| HH with private water facility (%) × time trend | | | | | −0.008 (0.001)*** |
| R^2 | 0.0395 | 0.0426 | 0.212 | 0.397 | 0.468 |

Notes: A dependent variable in models D1 and D2 is diarrhoea prevalence (7–15 years; %), model A1 is student absent from school (%), and models E1 and E2 are primary school enrolment (%). Control variables in models D1 and D2 are HH with improved sanitation (%), health facility coverage (%), food expenditure (000,000 IDR), parents' education (year), relative difference of mothers' and fathers' education years, and household living in an urban area (%). Control variables in models A1, E1, and E2 are food expenditure (000,000 IDR), parents' education (year), relative difference of mothers' and fathers' education years, and households living in an urban area (%). All models are estimated using the (static) panel fixed effect. Full results are listed in the Appendix, Table A-1 (available with the online version of this paper). Significance level at 0.1% ***, 1% **, and 5% *.

water on diarrhoea prevalence becomes stronger over time. This stronger effect might be caused by an increase in the quality of water used by households over the period (sig $p < 0.01$). This quality change is driven by an increase in the share of bottled water in the mixture of piped and bottled water over the period (see Figure 4). All in all, these results support the existence of an indirect effect of water quality on school attendance (models A1, E1, and E2).

School absenteeism in a district is positively related to diarrhoea prevalence and negatively related to the percentage of households with a private water facility (Table 2, A1). The estimated coefficients are 2.334 (sig $p < 0.001$) and −0.027 (sig $p < 0.01$), respectively. The first coefficient means that a 1% reduction in diarrhoea prevalence (7–15 years) in a given district would reduce school absenteeism of primary school students in that district by 2.3%. The latter coefficient indicates that an increase in the percentage of households with a private water facility in the district by 10% is associated with a reduction of primary school absenteeism by 0.27%. The fact that diarrhoea prevalence is

positively related with school absenteeism and that good quality water is negatively related with diarrhoea provides evidence for the existence of an indirect effect of water quality on the reduction of school absenteeism. No interaction effects with the time trend are found to be significant, which indicates that the effects do not change over time.

Lastly, school enrolment is significantly related to diarrhoea prevalence and the percentage of households with a private water facility (Table 2, E1). The estimated coefficients are −0.202 (sig $p < 0.01$) and 0.025 (sig $p < 0.01$), respectively. When the prevalence of diarrhoea in a district decreases by 1%, the percentage of primary school enrolment would increase by 0.2%. In addition, an increase in the percentage of households with a private water facility in a district by 10% would increase the district's percentage of primary school enrolment by 0.25%. One might argue that the effect is rather small, but this effect is actually more than half of the average annual change of primary school enrolment in Indonesia (0.42, see Table 1). In addition, given the latest primary school enrolment in

Indonesia, which reached 35 million children (Indonesia MoEC 2018), the 0.25% increase in primary school enrolment is equal to an additional primary school enrolment of 88,000 children. This effect is still important, given that the districts in Indonesia already reached, on average, a relatively high rate of primary enrolment (90%, see Table 1) and, hence, they might be facing the last mile problem. The coefficients of the interaction terms (model E2) suggest that both effects are diminishing over time. These diminishing effects indicate a convergence process; primary school enrolment in Indonesia is coming closer to full enrolment and, thus, the room for factors like water supply to influence enrolment is decreasing. Diarrhoea prevalence and private water facilities, therefore, become less important for explaining differences between districts in school enrolment. Or, conversely, this interaction effect suggests that improvements in water and sanitation are particularly beneficial in the districts with poorest water and sanitation infrastructures.

Robustness check: dynamic fixed effect

The static models neglect the dynamics and thus estimate a long-term relation only. As a robustness check, we estimate the long run relation as part of a dynamic model, a so-called dynamic fixed effect (DFE) method (Blackburne & Frank 2007). According to Pesaran & Shin (1999), a DFE model delivers consistent and efficient estimates of the parameters in a long-run relationship even when variables with different orders of integration, $I(0)$ or $I(1)$ or a mix of these two, are included in the regression. We performed two panel unit root tests, i.e. Harris & Tzavalis (1999), and Im *et al.* (2003) tests, to be sure that for no time series did the degree of integration exceed 1. The unit root test results, presented in Table A-2 in the Appendix (available online), indicate that the series of the percentage of households with piped and bottled water, the percentage of households with improved sanitation, food expenditure, and the level of parents' education are $I(1)$. Hence, all variables have a degree of integration equal to or below 1, so that the estimated coefficients are consistent and efficiently estimated.

Table 3 reports the results of the DFE analysis for diarrhoea (D3 and D4), school absenteeism (A2), and school enrolment (E3 and E4). In general, results from DFE

analysis show that the sign of the long-run coefficient of our main variables remains similar to those obtained by the static panel fixed effect analysis. The percentage of households with piped and bottled water is supportive for reducing diarrhoea prevalence (model D3), and this effect is stronger over time (model D4). The direct and indirect effect of water on school absenteeism is reconfirmed (model A2). However, the direct effect of (nearby) water on school enrolment becomes insignificant (model E3). The analysis with interaction terms (model E4) again demonstrates that the direct and indirect effects of water on school enrolment become less important over time. Lastly, for all models, the error correction coefficient is negative and highly significant, indicating the existence of convergence effects towards the long-run equilibrium.

DISCUSSION

The current study explores the direct and indirect effects of access to good quality water and a nearby water source on primary school absenteeism and school enrolment. School absenteeism is included because many enrolled pupils appear to be absent from school due to illnesses, child labour, and household responsibilities (Neuzil *et al.* 2002; Azor-Martínez *et al.* 2014).

The percentage of households with a private water facility in the district, the proxy for a nearby water facility, is (directly) associated with lower school absenteeism and higher school enrolment. These results are in line with those of previous studies on water and schooling in developing countries (Dreibelbis *et al.* 2013; Nauges & Strand 2013; Hunter *et al.* 2014; Zhang & Xu 2016). The provision of drinking water in close proximity to households reduces the burden of water fetching, which is often borne by the children. The provision of a nearby water source would reduce school absenteeism (Dreibelbis *et al.* 2013; Hunter *et al.* 2014) and increase school enrolment (Nauges & Strand 2013). The direct effect of a nearby water source on school enrolment should be interpreted cautiously as it is not confirmed in the robustness check.

The results also point to the relevance of the indirect effect. The percentage of households using piped and bottled water – our proxy for access to good quality water – is

Table 3 | Estimating the effect of access to water on diarrhoea prevalence (D3 and D4), school absenteeism (A2), and school enrolment (E3 and E4) based on DFE analysis, 259 districts in Indonesia, 1994–2014

| | D3 B/SE | D4 B/SE | A2 B/SE | E3 B/SE | E4 B/SE |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| <i>Main effects</i> | | | | | |
| HH using piped and bottled water (%) | −0.006 (0.002)** | −0.004 (0.002)* | | | |
| Diarrhoea prevalence (%) | | | 2.710 (0.182)*** | −0.426 (0.155)** | −0.519 (0.137)*** |
| HH with private water facility (%) | −0.000 (0.002) | 0.001 (0.002) | −0.032 (0.011)** | 0.013 (0.012) | 0.011 (0.011) |
| Time trend | 0.026 (0.013)* | 0.022 (0.013) | 0.339 (0.053)*** | 0.356 (0.080)*** | 0.222 (0.067)*** |
| <i>Interaction term</i> | | | | | |
| HH using piped and bottled water (%) × time trend | | −0.000 (0.000)*** | | | |
| Diarrhoea prevalence (%) × time trend | | | | | 0.041 (0.019)* |
| HH with private water facility (%) × time trend | | | | | −0.008 (0.001)*** |
| Error correction coefficient | −0.855 (0.022)*** | −0.858 (0.022)*** | −0.814 (0.018)*** | −0.485 (0.024)*** | −0.550 (0.023)*** |

Notes: A dependent variable in models D3 and D4 is Δ diarrhoea prevalence (7–15 years; %), model A2 is Δ student absent from school (%), and models E3 and E4 are Δ primary school enrolment (%). Control variables in models D3 and D4 are, both the level and the difference of, HH with improved sanitation (%), health facility coverage (%), food expenditure (000,000 IDR), parents' education (years), relative difference of mothers' and fathers' education years, and households living in an urban area (%). Control variables in models A2, E3 and E4 are, both the level and the difference of, food expenditure (000,000 IDR), parents' education (years), relative difference of mothers' and fathers' education years, and households living in an urban area (%). All models are estimated using the DFE method. Full results are listed in the Appendix, Table A-3 (available with the online version of this paper). Significance level at 0.1% ***, 1% **, and 5% *. The xtpmg routine in Stata does not provide any goodness-of-fit measures.

associated with lower diarrhoea prevalence in the region, which in turn is associated with lower absenteeism and higher primary school enrolment. This finding is a confirmation of Dreibelbis *et al.*'s (2013) suggestion that improved health is a pathway to an explanation of the association between access to clean water and better school attendance. We argue that good quality drinking water can improve children's health outcomes (Wolf *et al.* 2014; Komarulzaman *et al.* 2017), which increases their educational attendance through a reduction in absenteeism (Dreibelbis *et al.* 2013; Hunter *et al.* 2014). In addition to this short-term benefit, access to drinking water of higher quality can (indirectly) benefit children's education in the longer term by improving school enrolment in the region (Zhang & Xu 2016).

The interaction analysis shows that the direct effect of having a private water facility and the indirect effect of access to piped and bottled water on school enrolment are

diminishing over time. Given the increasing trend of school enrolment in Indonesia, this diminishing effect is understandable. The role of water access saturates as school enrolment get close to full coverage.

We could not confirm the effect of having a private water facility on diarrhoea prevalence and, hence, an indirect effect of private water on schooling could not be inferred. This could be caused by the common household practice of storing water. Hence, the benefits of having a private water facility are not guaranteed, as the cleanliness of the water is often not maintained during storage (Komarulzaman *et al.* 2017). Another explanation might be that the effects of living at a large distance from the water source are not well captured by the private water variable, as the distance to the source might be small for many households using a public source. The location of the source would be better measured by fetching time, as it includes both the walking and queuing time.

Our study has three unique characteristics. First, besides the direct effect of water access on school attendance, we include the indirect effect of water access on schooling through diarrhoea. Second, unlike previous studies which analyse the effect of water on either school absenteeism (Dreibelbis *et al.* 2013) or school enrolment (Nauges & Strand 2013; Zhang & Xu 2016), the current study explores the benefits of water access for school attendance measures. Lastly, the panel structure of the data set with a large number of districts over an extended period of time is exploited by estimating panel models. This is an important extension of earlier studies (Dreibelbis *et al.* 2013) of water and schooling, which have focused on the cross-section variation within a single period. Panel analysis is powerful enough to capture the variation among observations both across districts and over time. The latter is important since it takes time to switch towards another resource and for an intervention to take effect. When an infrastructure, e.g. a safe water source, is established in the vicinity of households, the benefit gained from this facility might not immediately be enjoyed by all of these households. Moreover, the model allows us to test whether the magnitude of the effect of water on schooling changes over time. Hence, exploring both the variation across sections and over time could give a better understanding of the dynamics of water supply on school attendance.

This study has some limitations, many of which are related to the data used. We constructed time series at the district level from the Susenas data set. As such, this was already quite an exercise as, in some cases, the boundaries of the districts changed over time. Moreover, it is the only possibility to generate time series. However, ideally, one would like to have longitudinal data at the household level, as at the district level, one can only investigate the relations for the average household. Consequently, the conclusions are very general. In practice, households within a district differ, and it is of interest how these differences between the households interact with their environment. The present data do not allow for such an analysis over time. Another disadvantage relates to the way fetching water is measured. In the current study, we use the variable of whether households have a private water facility. If not, then we assume that children have to fetch water outside of the premises. We do not know the distance between

this source and their parents' house. Moreover, we do not know who is fetching the water. Adding this information to the data set would help in finding more specific policy recommendations.

CONCLUDING REMARKS

The current study extends the understanding of the impact of water on schooling in two ways. Firstly, while previous studies have focused on the direct effect, our study provides evidence of both direct and indirect effects of water on schooling. Secondly, we study the effect of water access on education outcomes as measured by both school absenteeism and school enrolment. Our results suggest that the access to safe water at the premises can potentially improve education outcomes at the district level. The availability of good-quality and nearby water facilities is negatively related to school absenteeism and positively related to school enrolment. This result highlights the importance of safe water provision for children's health and schooling. Consequently, this benefit should be taken into account by the policy-makers in identifying and quantifying the potential benefits of developing water infrastructure in developing countries.

ACKNOWLEDGEMENT

This work was supported by a scholarship from the Directorate General of Higher Education (DIKTI), the Ministry of Education and Culture, Republic of Indonesia.

REFERENCES

- Azor-Martínez, E., Gonzalez-Jimenez, Y., Seijas-Vazquez, M. L., Cobos-Carrascosa, E., Santisteban-Martínez, J., Martínez-López, J. M., Jimenez-Noguera, E., Galan-Requena, M. D. M., Garrido-Fernández, P., Strizzi, J. M. & Gimenez-Sanchez, F. 2014 *The impact of common infections on school absenteeism during an academic year. American Journal of Infection Control* **42** (6), 632–637.
- Blackburne, E. F. & Frank, M. W. 2007 *Estimation of non-stationary heterogenous panels. Stata Journal* **7**, 197–208.
- Carter Hill, R., Griffiths, W. E. & Lim, G. C. 2012 *Principle of Econometrics*, 4th edn. John Wiley & Sons, Hoboken, NJ.

- Checchi, D. 2006 *The Economics of Education: Human Capital, Family Background and Inequality*. Cambridge University Press, New York.
- Clasen, T. F. & Haller, L. 2008 *Water Quality Interventions to Prevent Diarrhoea: Cost and Cost-Effectiveness*. World Health Organization, Geneva.
- Dreibelbis, R., Greene, L. E., Freeman, M. C., Saboori, S., Chase, R. P. & Rheingans, R. 2013 *Water, sanitation, and primary school attendance: a multi-level assessment of determinants of household-reported absence in Kenya*. *International Journal of Educational Development* **33** (5), 457–465.
- Eide, E. R. & Showalter, M. H. 2010 Human capital. In: *Economics of Education* (D. J. Brewer & P. J. McEwan, eds). Elsevier Ltd, Oxford, UK, pp. 27–32.
- Glick, P. & Sahn, D. E. 2000 *Schooling of girls and boys in a West African country: the effects of parental education, income, and household structure*. *Economics of Education Review* **19** (1), 63–87.
- Harris, R. D. F. & Tzavalis, E. 1999 *Inference for unit roots in dynamic panels where the time dimension is fixed*. *Journal of Economics* **91**, 201–226. doi:10.1016/S0304-4076(98)00076-1.
- Huisman, J. & Smits, J. 2015 *Keeping children in school*. *SAGE Open* **5** (4), 1–16.
- Hunter, P. R., Risebro, H., Yen, M., Lefebvre, H., Lo, C., Hartemann, P., Longuet, C. & Jaquenoud, F. 2014 *Impact of the provision of safe drinking water on school absence rates in Cambodia: a quasi-experimental study*. *PLoS ONE* **9** (3), e91847.
- Im, K. S., Hashem Pesaran, M., Shin, Y. & Lin, L. 2003 *Testing for unit roots in heterogeneous panels*. *Journal of Economics* **115**, 53–74.
- Indonesia Ministry of Education and Culture 2018 *Neraca Pendidikan Nasional 2018*, Jakarta, Indonesia. Available from: <http://npd.kemdikbud.go.id/?appid=download>.
- Jalan, J. & Ravallion, M. 2003 *Does piped water reduce diarrhea for children in rural India?* *Journal of Econometrics* **112** (1), 153–173.
- Komarulzaman, A., Smits, J. & de Jong, E. 2017 *Clean water, sanitation and Diarrhoea in Indonesia: effects of household and community factors*. *Global Public Health* **12**, 1141–1155. doi:10.1080/17441692.2015.1127985.
- Mangyo, E. 2008 *The effect of water accessibility on child health in China*. *Journal of Health Economics* **27** (5), 1343–1356.
- Nauges, C. & Strand, J. 2013 *Water Hauling and Girls' School Attendance Some New Evidence From Ghana, Policy Research Working Paper (No. 6443)*, The World Bank.
- Neuzil, K. M., Hohlbein, C. & Zhu, Y. 2002 *Illness among schoolchildren during influenza season: effect on school absenteeism, parental absenteeism from work, and secondary illness in families*. *Archives of Pediatrics and Adolescent Medicine* **156** (10), 986–991.
- Pesaran, M. H. & Shin, Y. 1999 *An autoregressive distributed lag modeling approach to cointegration analysis*. In: *Econometrics and Economic Theory in the 20th Century: The Ragnar Frish Centennial Symposium* (S. Strom, ed.). Cambridge University Press, Cambridge, pp. 371–413.
- Sepehri, A. & Sarrafzadeh, M. H. 2018 *Effect of nitrifiers community on fouling mitigation and nitrification efficiency in a membrane bioreactor*. *Chemical Engineering and Processing: Process Intensification* **128**, 10–18.
- Sorenson, S. B., Morssink, C. & Campos, P. A. 2011 *Safe access to safe water in low income countries: water fetching in current times*. *Social Science & Medicine* **72** (9), 1522–1526.
- Statistics Indonesia 2014 *Survei Sosial Ekonomi Nasional (Susenas) 2014*. Jakarta, Indonesia.
- Todaro, M. P. & Smith, S. C. 2012 *Economic Development*, 11th edn., Addison-Wesley, Boston, MA.
- UNICEF 2012 *Issue Brief. Water, Sanitation & Hygiene, Indonesia*, Jakarta, Indonesia.
- UNICEF & WHO 2009 *Diarrhoea: Why Children are Still Dying and What can be Done*, Geneva.
- Verbeek, M. 2004, *A Guide to Modern Econometrics*, John Wiley & Sons, Ltd, Chichester.
- World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) (2017). *Progress on Drinking Water, Sanitation and Hygiene Update and SDG Baselines*, Geneva.
- Wolf, J., Prüss-Ustün, A., Cumming, O., Bartram, J., Bonjour, S., Cairncross, S., Clasen, T., Colford, J. M., Curtis, V., De France, J., Fewtrell, L., Freeman, M. C., Gordon, B., Hunter, P. R., Jeandron, A., Johnston, R. B., Mäusezahl, D., Mathers, C., Neira, M. & Higgins, J. P. T. 2014 *Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression*. *Tropical Medicine & International Health* **19** (8), 928–942.
- Zhang, J. & Xu, L. C. 2016 *The long-run effects of treated water on education: the rural drinking water program in China*. *Journal of Development Economics* **122**, 1–15.

First received 1 March 2019; accepted in revised form 2 May 2019. Available online 21 May 2019