Does Education Improve Health? Evidence from Indonesia

RASYAD A. PARINDURI
Nottingham University Business School, The University of Nottingham, Malaysia Campus, Semenyih, Selangor, Malaysia

Online Appendix

A.1. Robustness checks

A.1.1 Using quarter of birth as the assignment variable. I use a more refined assignment variable, quarter of birth (instead of year of birth), as the assignment variable; the results are similar. Figure A1, which illustrates the reduced-form effects of the longer school year, shows that the trend line of the proportion of individuals who completed high school between the 1971 and 1972 cohorts breaks, but we do not see breaks in the trend line of health in general or that of smoking incidence. If anything, there is only a slight break between the 1971 and 1972 cohorts in cognitive capacity measured using the word memorisation exercise. Panel A of Table A1, which presents the third-stage estimates, confirms the trend lines: I do not find evidence that completing high school improves health, health behaviour, or access to modern health facilities; if anything, it leads to a 20 percentage point (40%) increase in cognitive capacity measured by the proportion of words remembered in the memorisation exercise.

A.1.2 Using alternative definitions of the longer school year. I use alternative definitions of the longer school year; overall the results are quite robust. (I present the third-stage estimates of completing high school only for the sake of brevity.) In row (2) of Table A1 I use year of birth as the assignment variable and year of entry to primary school to define the longer school year; in row (3) I use year of entry to primary school to define the longer school year and to be the assignment variable. I do not find evidence that completing high school affects health, health behaviour, or access to modern health facilities, but I find it improves cognitive capacity. The estimate in column (3) and row (2) is large, about 20 percentage points and statistically significant at the 1 per cent level; the estimate in row (3) is also large though it is significant at the 10 per cent level, which is probably caused by inaccuracies of year of entry to primary school in the data.

A.1.3 Analyses by subgroup. It is possible that education matters for some subgroups of people in the population, but overall the results of analyses by subsample are robust (Table A2). I find a few interesting details, however. (The table presents the results by the location of individuals when they were 12 years old, gender, and level of school they were attending when they experienced the longer school year.) One, the effects of education on cognitive capacity is mostly driven by the effects on people who grew up in rural areas: the estimate of the effects on people who grew up in rural areas is large and statistically significant; that of people who grew up in urban areas is small and statistically insignificant. (We should cautiously interpret the estimates in Table A2, however, because I do multiple comparisons, the standard errors are large, and the difference between statistically significant and insignificant estimates may be
Complete senior high school

Somewhat or very healthy in general

Able to memorize words

Currently smoke

**Figure A1.** Using quarter of birth as the assignment variable.

| Table A1. | Using alternative assignment variables and definitions of longer school year |
|-----------|-------------------------------------------------------------------------|
| Dependent variable | Healthy in general (1) | Hypertension (2) | Able to memorise words (3) | Currently smoke (4) | Take vitamins (5) | Have private health insurance (6) | Treated by shamans (7) |
| **A. Assignment variable: quarter of birth** | 0.16 | -0.02 | 0.21** | 0.04 | 0.14 | 0.19 | 0.0004 |
| Longer school year (using year of birth) | (0.14) | (0.13) | (0.07) | (0.20) | (0.10) | (0.14) | (0.26) |
| **B. Assignment variable: year of birth** | 0.16 | -0.01 | 0.19** | 0.09 | 0.13 | 0.18 | -0.07 |
| Longer school year (using year of entry) | (0.14) | (0.13) | (0.07) | (0.20) | (0.10) | (0.15) | (0.23) |
| **C. Assignment variable: year of entry** | 0.25 | -0.32 | 0.25 | 0.50 | 0.03 | 0.42 | -0.33 |
| Longer school year (using year of entry) | (0.28) | (0.30) | (0.15) | (0.45) | (0.20) | (0.30) | (0.51) |

**Notes:** The number in each cell is the equation-by-equation 2SLS estimate of the effects of completing high school. The dependent variables are listed on the top row. The figures in parentheses are standard errors. One and two stars indicate statistical significance at the levels of 5 and 1 per cent, respectively.
Two, the effects of education on cognitive capacity may be driven by the effects on females: the estimate of the effects on females is large and statistically significant; that on males is also large, but statistically significant at the 10 per cent level only. Three, education seems to improve access to private health insurance for individuals who grew up in rural areas: The estimate is large and statistically significant at the 1 per cent level; that of urban is statistically insignificant.

Four, focusing on the subsample of people who were in primary school when they experienced the longer school year, and on another subsample of people who were in primary school or junior high school, education seems to affect health in general: the estimates are large and statistically significant at the 5 per cent level.

A.1.4 Two-stage least-square estimations of the effects of education. I estimate two-stage instrumental-variable regressions of the effects of education on health in which I use older cohorts as an instrument for education; I find similar results (Table A3). Columns (1–2) present the second-stage estimates of the effects of educational attainment and completing high school, respectively: the magnitude, sign, and statistical significance of the estimates of the effects of education on cognitive capacity are similar to those of the basic results; all others estimates are statistically insignificant.

A.1.5 False treatments. I check whether false treatments (that is, incorrectly defined older cohort) affect education and health to find out whether some aggregate time trend drives the basic results but I do not find such evidence. I use the subperiod to the left of the cut-off (1960–1971) from the 1960–1961 to 1970–1971 cut-offs (Table A4). (For the 1961–1962 cut-off, for example, the false treatment equals one if a person was born in 1961 or earlier and zero otherwise.) The estimate of the first stage is statistically significant only for the 1965–1966 cut-off (column (1) and row (6)); the corresponding estimate of the second stage in which educational attainment is the outcome variable is also statistically significant (column (2)) but that of the second stage in which completing high school is the statistically insignificant.) Two, the effects of education on cognitive capacity may be driven by the effects on females: the estimate of the effects on females is large and statistically significant; that on males is also large, but statistically significant at the 10 per cent level only. Three, education seems to improve access to private health insurance for individuals who grew up in rural areas: The estimate is large and statistically significant at the 1 per cent level; that of urban is statistically insignificant. Four, focusing on the subsample of people who were in primary school when they experienced the longer school year, and on another subsample of people who were in primary school or junior high school, education seems to affect health in general: the estimates are large and statistically significant at the 5 per cent level.

A.1.4 Two-stage least-square estimations of the effects of education. I estimate two-stage instrumental-variable regressions of the effects of education on health in which I use older cohorts as an instrument for education; I find similar results (Table A3). Columns (1–2) present the second-stage estimates of the effects of educational attainment and completing high school, respectively: the magnitude, sign, and statistical significance of the estimates of the effects of education on cognitive capacity are similar to those of the basic results; all others estimates are statistically insignificant.

A.1.5 False treatments. I check whether false treatments (that is, incorrectly defined older cohort) affect education and health to find out whether some aggregate time trend drives the basic results but I do not find such evidence. I use the subperiod to the left of the cut-off (1960–1971) from the 1960–1961 to 1970–1971 cut-offs (Table A4). (For the 1961–1962 cut-off, for example, the false treatment equals one if a person was born in 1961 or earlier and zero otherwise.) The estimate of the first stage is statistically significant only for the 1965–1966 cut-off (column (1) and row (6)); the corresponding estimate of the second stage in which educational attainment is the outcome variable is also statistically significant (column (2)) but that of the second stage in which completing high school is the statistically insignificant.) Two, the effects of education on cognitive capacity may be driven by the effects on females: the estimate of the effects on females is large and statistically significant; that on males is also large, but statistically significant at the 10 per cent level only. Three, education seems to improve access to private health insurance for individuals who grew up in rural areas: The estimate is large and statistically significant at the 1 per cent level; that of urban is statistically insignificant. Four, focusing on the subsample of people who were in primary school when they experienced the longer school year, and on another subsample of people who were in primary school or junior high school, education seems to affect health in general: the estimates are large and statistically significant at the 5 per cent level.

A.1.4 Two-stage least-square estimations of the effects of education. I estimate two-stage instrumental-variable regressions of the effects of education on health in which I use older cohorts as an instrument for education; I find similar results (Table A3). Columns (1–2) present the second-stage estimates of the effects of educational attainment and completing high school, respectively: the magnitude, sign, and statistical significance of the estimates of the effects of education on cognitive capacity are similar to those of the basic results; all others estimates are statistically insignificant.

A.1.5 False treatments. I check whether false treatments (that is, incorrectly defined older cohort) affect education and health to find out whether some aggregate time trend drives the basic results but I do not find such evidence. I use the subperiod to the left of the cut-off (1960–1971) from the 1960–1961 to 1970–1971 cut-offs (Table A4). (For the 1961–1962 cut-off, for example, the false treatment equals one if a person was born in 1961 or earlier and zero otherwise.) The estimate of the first stage is statistically significant only for the 1965–1966 cut-off (column (1) and row (6)); the corresponding estimate of the second stage in which educational attainment is the outcome variable is also statistically significant (column (2)) but that of the second stage in which completing high school is the statistically insignificant.) Two, the effects of education on cognitive capacity may be driven by the effects on females: the estimate of the effects on females is large and statistically significant; that on males is also large, but statistically significant at the 10 per cent level only. Three, education seems to improve access to private health insurance for individuals who grew up in rural areas: The estimate is large and statistically significant at the 1 per cent level; that of urban is statistically insignificant. Four, focusing on the subsample of people who were in primary school when they experienced the longer school year, and on another subsample of people who were in primary school or junior high school, education seems to affect health in general: the estimates are large and statistically significant at the 5 per cent level.

A.1.4 Two-stage least-square estimations of the effects of education. I estimate two-stage instrumental-variable regressions of the effects of education on health in which I use older cohorts as an instrument for education; I find similar results (Table A3). Columns (1–2) present the second-stage estimates of the effects of educational attainment and completing high school, respectively: the magnitude, sign, and statistical significance of the estimates of the effects of education on cognitive capacity are similar to those of the basic results; all others estimates are statistically insignificant.

A.1.5 False treatments. I check whether false treatments (that is, incorrectly defined older cohort) affect education and health to find out whether some aggregate time trend drives the basic results but I do not find such evidence. I use the subperiod to the left of the cut-off (1960–1971) from the 1960–1961 to 1970–1971 cut-offs (Table A4). (For the 1961–1962 cut-off, for example, the false treatment equals one if a person was born in 1961 or earlier and zero otherwise.) The estimate of the first stage is statistically significant only for the 1965–1966 cut-off (column (1) and row (6)); the corresponding estimate of the second stage in which educational attainment is the outcome variable is also statistically significant (column (2)) but that of the second stage in which completing high school is the statistically insignificant.) Two, the effects of education on cognitive capacity may be driven by the effects on females: the estimate of the effects on females is large and statistically significant; that on males is also large, but statistically significant at the 10 per cent level only. Three, education seems to improve access to private health insurance for individuals who grew up in rural areas: The estimate is large and statistically significant at the 1 per cent level; that of urban is statistically insignificant. Four, focusing on the subsample of people who were in primary school when they experienced the longer school year, and on another subsample of people who were in primary school or junior high school, education seems to affect health in general: the estimates are large and statistically significant at the 5 per cent level.
outcome is not (column (3)). (The estimate of the second stage in which educational attainment is the
dependent variable for the 1962–1963 cut-off is also statistically significant but the corresponding
estimate of the first stage is not.) All other estimates are statistically insignificant. Using subperiod to
the right of the cut-off (1972–1987), the first-stage estimates of false treatments from the 1972–1973 to
1986–1987 cut-offs are zero; people who were born after 1972 did not experience the longer school
year. The estimates in row (6) and columns (1–2) are statistically significant probably because the 1965
and later cohorts were exposed to the *SD Inpres* construction programme initiated in 1972. In any case,
policies that might induce the discontinuity do not compromise the basic results because they
happened long before 1979 when the government extended the longer school year. I estimate the
effects of the ‘exposure’ to the *SD Inpres* programme on health but none of the estimates is statistically
significant (Table A5).

### A.1.6 Continuity of control variables at the 1971–1972 cut-off.
I present a formal test on the
continuity of some control variables at the 1971–1972 cut-off in Table 1 and discuss them in
Subsection 2.2; I present the trend lines of the averages (or proportions) in Figure A2.

### A.1.7 Continuity of the density of birth year.
I create histograms of birth year using one-, two-, three-
and four-year bin-widths; I do not see an obvious discontinuity of the density at the 1971–1972 cut-off
(Figure A3).

I check whether the density of birth year at the 1971–1972 cut-off discontinues using Cattaneo,
Jansson, and Ma’s (2016a, 2016b) manipulation tests; I do not find evidence to reject the null
hypothesis of the continuity of the density (the robust bias-corrected t-stat and p-value are 0.96 and
0.34, respectively). Therefore, I do not find evidence that parents perfectly manipulate the years their
children were born or entered primary school to take advantage or avoid the longer school year, which
is unsurprising given that the government announced the longer school year in the middle of the 1978
academic year and implemented it immediately, long after the children who entered primary schools in
1978–1979 were born.
Using father’s education as an instrument. Measurement errors in the estimation of the effects of education on health are a concern but, because the outcome variable in the last-stage regressions is health outcomes, it may be difficult to find a good instrument to address the problem. Some papers such as Card (2001) use parents’ education as an instrument for education to estimate returns to education, but, in estimations of the effects of education on health, parents’ education does not satisfy the exclusionary restriction assumption that the instrument affects health only through education. Parent’s educational attainment correlates with, for example, income and wealth, which affect children’s health directly.

I estimate two-stage least-square regressions of the effects of education on health in which I instrument education using father’s educational attainment in the first stage. Table A6 presents the first-stage estimates, which show that father’s educational attainment ‘increases’ education; Table A7 presents the second-stage estimates, which show education ‘improves’ health outcomes and health behaviour. However, because the exclusionary restriction assumption does not hold, the estimates in Table A2 are unlikely to be the estimates of the effects of education on health.

A full set of control variables. I use cubic polynomial function of age and sets of ethnicity, gender and location when 12 years old dummies as additional control variables. I present some results in Table 7; I present the results on all outcome variables in Table A8.
| Dependent variable                                 | Equation-by-equation 2SLS estimate of the effects of educational attainment |
|---------------------------------------------------|--------------------------------------------------------------------------|
| **A. Health outcomes**                            |                                                                          |
| Healthy in general                                 | 0.02                                                                     |
| Healthy compared to peers                         | 0.05                                                                     |
| Had bed rest in the past four weeks               | 0.02                                                                     |
| Hypertension                                      | 0.09                                                                     |
| Overweight                                        | -0.12                                                                    |
| Obese                                             | 0.03                                                                     |
| Ever diagnosed chronic illness                    | (0.07)                                                                  |
| Have disabilities that limit work                 | -0.05                                                                   |
| Remember today's date                             | -0.02                                                                   |
| Able to memorise words                           | 0.004                                                                   |
| Able to memorise the words a few minutes later    | 0.02                                                                    |
| **B. Health behaviour**                           |                                                                          |
| Ever smoke                                        | 0.11                                                                    |
| Currently smoke                                   | 0.15                                                                    |
| Take vitamins and supplements                     | 0.001                                                                   |
| Eat vegetables daily                              | -0.03                                                                   |
| Eat fruits daily                                  | -0.05                                                                   |
| **C. Access to insurance and modern health facilities** |                                                                  |
| Have health insurance                             | 0.05                                                                    |
| Have private health insurance                     | 0.05                                                                    |
| Do check up                                       | -0.03                                                                   |
| Treated by shamans                                | -0.07                                                                   |
| Receive inpatient care at hospitals or clinics    | -0.02                                                                   |

*Notes:* The number in each cell is the equation-by-equation 2SLS estimate of the effects of educational attainment. The dependent variables are listed on the left column. The figures in parentheses are standard errors. One and two stars indicate statistical significance at the levels of 5 and 1 per cent, respectively.
Figure A2. The averages of control variables by birth year.
Figure A3. The density of birth year.

Notes: The horizontal axis in each graph is the year of birth; the red vertical line is the 1971–1972 threshold. The sizes of bins from top-right clockwise are 1, 2, 3 and 4 years.

Table A6. The first-stage estimates of a two-stage least-square regressions of the effects of educational attainment on health in which father’s educational attainment is the instrument

|                          | (1)     | (2)     | (3)     |
|--------------------------|---------|---------|---------|
| Father’s educational attainment | 0.51**  | 0.51**  | 0.46**  |
|                          | (0.01)  | (0.01)  | (2.49)  |
| Adjusted R²              | 0.24    | 0.25    | 0.29    |
| Number of observations   | 12,836  | 12,836  | 12,799  |
| Controls                 |         | ✓       | ✓       |
| Age cubic polynomial     | ✓       | ✓       |         |
| Gender, ethnicity, and location when twelve years old dummies | ✓ | ✓ | |

Notes: The number in each cell is the estimate of father's educational attainment from the first-stage of a two-stage least-square estimation of the effects of educational attainment on health. The figures in parentheses are robust standard errors clustered by year of birth. One and two stars indicate statistical significance at the levels of 5 and 1 per cent, respectively.
| Dependent variable                        | (1)       | (2)       |
|------------------------------------------|-----------|-----------|
| Healthy in general                       | 0.005**   | 0.04**    |
|                                         | (0.001)   | (0.01)    |
| Hypertension                             | -0.001    | -0.010    |
|                                         | (0.001)   | (0.01)    |
| Able to memorise words                  | 0.02**    | 0.16**    |
|                                         | (0.001)   | (0.01)    |
| Currently smoke                          | -0.01**   | -0.09**   |
|                                         | (0.002)   | (0.02)    |
| Take vitamins and supplements            | 0.01**    | 0.11**    |
|                                         | (0.002)   | (0.01)    |
| Have private health insurance            | 0.05**    | 0.42**    |
|                                         | (0.003)   | (0.02)    |
| Treated by shamans                       | 0.002     | 0.02      |
|                                         | (0.004)   | (0.04)    |

Notes: The number in each cell in column (1) is the 2SLS estimate of the effects of educational attainment on the dependent variable listed on the left column; each number in column (2) is that of the effects of completing high school. The figures in parentheses are robust standard errors clustered by year of birth. All regressions include gender, ethnicity, and location when twelve years old dummies as well as a cubic polynomial function of age. One and two stars indicate statistical significance at the levels of 5 and 1 per cent, respectively.

| Dependent variable                        | (1)       | (2)       |
|------------------------------------------|-----------|-----------|
| Healthy in general                       | 0.02      | 0.18      |
|                                         | (0.02)    | (0.14)    |
| Healthy compared to peers                | 0.02      | 0.16      |
|                                         | (0.01)    | (0.10)    |
| Had bed rest in the past four weeks      | -0.01     | -0.05     |
|                                         | (0.01)    | (0.11)    |
| Hypertension                             | -0.0002   | -0.001    |
|                                         | (0.02)    | (0.14)    |
| Overweight                               | -0.04     | -0.27     |
|                                         | (0.05)    | (0.36)    |
| Obese                                    | -0.01     | -0.10     |
|                                         | (0.02)    | (0.17)    |
| Ever diagnosed chronic illness           | 0.003     | 0.02      |
|                                         | (0.01)    | (0.11)    |
| Have disabilities that limit work        | -0.01     | -0.07     |
|                                         | (0.01)    | (0.07)    |
| Remember today's date                    | 0.01      | 0.05      |
|                                         | (0.01)    | (0.07)    |
| Able to memorise words                  | 0.03**    | 0.21**    |
|                                         | (0.01)    | (0.07)    |
| Able to memorise the words a few minutes later | 0.02*  | 0.16*  |
|                                         | (continued) |
Table A8. (Continued)

| Dependent variable | (1) | (2) |
|--------------------|-----|-----|
|                    | (0.01) | (0.08) |
| **B. Health behaviour** | |
| Ever smoke         | 0.02  | 0.18 |
|                    | (0.02) | (0.15) |
| Currently smoke    | 0.03  | 0.20 |
|                    | (0.02) | (0.15) |
| Take vitamins and supplements | 0.02  | 0.12 |
|                    | (0.01) | (0.11) |
| Eat vegetables daily | -0.003 | -0.02 |
|                    | (0.03) | (0.21) |
| Eat fruits daily   | -0.01 | -0.05 |
|                    | (0.02) | (0.15) |
| **C. Access to insurance and modern health facilities** | |
| Have health insurance | 0.03  | 0.20 |
|                    | (0.03) | (0.19) |
| Have private health insurance | 0.03  | 0.20 |
|                    | (0.02) | (0.15) |
| Do check up        | -0.04 | -0.30 |
|                    | (0.02) | (0.17) |
| Treated by shamans | 0.003 | 0.02 |
|                    | (0.05) | (0.24) |
| Receive inpatient care at hospitals or clinics | 0.001 | 0.004 |
|                    | (0.01) | (0.08) |

Notes: The number in each cell in column (1) is the equation-by-equation 2SLS estimate of the effects of educational attainment; that in column (2) is the equation-by-equation 2SLS estimate of the effects of completing high school. The dependent variables are listed on the left column. The figures in parentheses are standard errors. One and two stars indicate statistical significance at the levels of 5 and 1 per cent, respectively.

References

Card, D. (2011). Estimating the returns to schooling: progress on some persistent econometric problems. *Econometrica*, 69, 1127–1160.

Cattaneo, M. D., Jansson, M., & Ma, X. (2016a). Simple local regression distribution estimators with an application to manipulation testing. *University of Michigan Working Paper*.

Cattaneo, M. D., Jansson, M., Ma, X. (2016b). rddensity: Manipulation testing based on density discontinuity. *Stata Journal*, forthcoming.