Original Article

Impacts of Late School Entry on Children’s Cognitive Development in Rural Northwestern China—Does Preprimary Education Matter?

Qihui Chen

Abstract

This article estimates the causal effect of primary school entry age on children’s cognitive development in rural northwestern China, using data on nearly 1,800 primary school aged children from the Gansu Survey of Children and Families. Instrumental variable estimates, exploiting the discontinuity structure in children’s school entry age around the enrolment cut-off date, indicate that a 1-year delay in school entry reduces children’s scores on a cognitive ability test administered when they were aged 9–12 by 0.11–0.16 standard deviations (of the distribution of test scores). The negative late-school-entry effect is significantly larger in villages with no preprimary schools. It also persists as children advance to higher grades. These findings suggest that delayed school entry, even if it may be rural parents’ rational response to resource constraints, can be harmful for children’s cognitive development in developing areas with underdeveloped preprimary school systems.

Key words: school entry age, cognitive development, preprimary school, rural China

1. Introduction

Despite the fact that almost all countries specify an appropriate age for primary school enrolment (UNESCO 2007), late primary school entry prevails in many developing countries (Lloyd & Blanc 1996; Pradhan 1998; UNESCO 2000; Akyeampong et al. 2007; Nonoyama-Tarumi et al. 2007). Although most existing studies in developing countries agree that children’s late school entry results from their parents’ reaction to resource constraints (e.g. Glewwe & Jacoby 1995; Wils 2004; Chen 2015; Seshie-Nasser & Oduro 2016), these studies have yet to reach a consensus on how late school entry affects children’s subsequent educational outcomes in developing countries. For example, studies conducted in Ghana (Glewwe & Jacoby 1995) and the Philippines (Glewwe et al. 2001) found positive effects of late school entry—late school entry reduces children’s dropout and grade repetition rates in these countries. In contrast, studies done in other developing economies, such as Mozambique (Wils 2004), Uganda (Uganda Bureau of Statistics and ORC Macro 2002), Zambia (Zambia Central Statistical Office and ORC Macro 2003) and rural China (Chen 2015), found negative late-school-entry effects—in particular, these four studies all...
found that late school entry leads to early school departure among school-aged children.¹

In explaining the effects of late school entry, studies finding positive effects, including those done in developed countries,² almost uniformly cited the extra preprimary investments (both educational and nutritional) received by children who entered primary school late as a key factor. In contrast, explanations of negative late-school-entry effects are more diverse. For example, objective causes, such as long distance to school and financial constraints, may themselves cause early school departure (Caldwell 1967). Subjectively, parents who enrol their children in school late may also be less motivated for child education in general, lacking a strong incentive to keep their children in school sufficiently long (Wils 2004). Moreover, compared with their younger classmates, children who entered school late may have more responsibilities for helping their parents with housework, farm work or even wage work, which, in turn, increase the opportunity costs of schooling, rendering schooling less attractive to these children (Chen 2015).

This article contributes to the literature by examining another potential channel through which late school entry may lead to early departure, which may also amplify the effects of the afore-mentioned factors and channels. We hypothesize that late school entry undermines children’s cognitive development in developing areas with underdeveloped preprimary education systems. Without a well-functioning preprimary school system (including daycare, preschool and kindergarten), children who start primary school late are likely to spend a number of years in a mentally unstimulating environment prior to formal schooling, especially when their parents are incapable of providing them with proper and sufficient mental stimulations at home (for example, due to low levels of parental education). The years spent in the mentally unstimulating preprimary environment may, in turn, impose a limit on these children’s potential for cognitive development, rendering schooling more difficult and less attractive to them, eventually pushing some of them out of school earlier than otherwise.

Although thoroughly testing this hypothesis demands more data than we have, we are able to test two related hypotheses that are necessary for it to hold, using data on nearly 1,800 rural children randomly chosen from Gansu province of China. First, we test whether late primary school entry has a negative impact on children’s cognitive ability (measured by scores on two cognitive ability tests, administered respectively in 2000 and 2004). Second, we investigate whether the availability of preprimary education in the village helps to offset some of the negative impact of late school entry (or even to generate a positive late-school-entry effect).

The major identification problem we encounter when estimating the causal effect of late school entry is the potential endogeneity of children’s school entry age. If parents chose to delay their children’s school entry based on factors that are unobserved to the researcher, say, financial constraints or child health problems, simple comparisons of cognitive ability scores among children with different school entry ages may be misleading because such comparisons may pick up the influences of those unobserved factors. Following the recent literature (Fredriksson & Öckert 2005; Bedard & Dhuey 2006; Elder & Lubotsky 2009; Black et al. 2011; Chen 2015), we address this problem by exploiting a source of exogenous variation in school entry age created by China’s Compulsory Education Law of 1994. The law stipulates that children turning age 7 by 31 August in a given year should be enrolled in primary school in that year.³ Given the enrolment cut-off of 31 August, children turning age 7 in September or later in a given year are expected

1. A recent study conducted in the United States (Dobkin & Ferreira 2010) also found that late school entry leads to early school departure.

2. Recent studies conducted in developed countries mostly found beneficial impacts of later school entry on children’s educational outcomes (e.g. Fredriksson & Öckert 2005; Bedard & Dhuey 2006; Cascio & Lewis 2006; Elder & Lubotsky 2009; Smith 2009).

3. However, in areas with difficulty in enrolling all children aged 6 in school, children can be enrolled when they have reached age 7. This is the case in Gansu, where primary school age is defined as 7–12 years old; see http://www.jyxjyj.com/Article/ShowArticle.asp?ArticleID=182 (accessed on 5 April 2014).
to be enrolled in the next school year by law. Thus, among children who comply with the law, September-born children will be almost 1 year older than their August-born classmates upon primary school enrollment, which creates a discrete jump in children’s school entry age around the enrollment cut-off. To the extent that a child’s month of birth is exogenously determined, this discontinuity structure can be exploited to create plausible instrumental variables (IVs) for children’s school entry age.4

Our IV estimation results indicate a significantly negative and persistent late-school-entry effect on children’s cognitive development. A 1-year increase in school entry age lowers children’s scores on both the 2000 and 2004 tests by 0.11–0.16 standard deviations (SDs, of the distribution of test scores). We also found that the availability of preprimary school in the village serves to offset some of these negative effects, which highlights the importance of developing a well-functioning preprimary school system in rural China.

The rest of the article proceeds as follows. The next section describes the study area and data. Section 3 presents the analytical framework underlying the empirical analysis. Section 4 presents and discusses our estimation results. The final section draws conclusions and presents some suggestions for future research.

2. Data

2.1. Study Area

The study area of this article, Gansu province, is located in northwestern China. This province is among the poorest in the country. According to the World Bank (2001), 23% of Gansu’s rural population was characterized as being poor in 2000, approximately four times the national figure of 6.5% (National Bureau of Statistics 2005). Despite the considerable efforts devoted by the Chinese government to developing Western China, investments in basic education have been behind schedule in rural Gansu. The completion rate of the 9-year compulsory education in Gansu was only 55.4% in 2002 (Shen & Wang 2003). Late primary school enrollment is also common. Despite the officially enrollment age set at 7, more than 30% of children in rural Gansu did not start primary school until the age of 8 (Table 1). Dropout rates are also high, especially among secondary school age children. For example, the dropout rates among children aged 15, 16 and 17 years old in rural Gansu were, respectively, 9.1%, 19.9% and 29.4% in 2004 (Zhao & Glewwe 2010). Note that these observations are indeed interrelated. In particular, analysing the data studied by Zhao and Glewwe (2010), Chen (2015) found that late school entry can explain partly children’s dropout behaviour. To serve as a complement to the findings of Zhao and Glewwe (2010) and those of Chen (2015), this article examines whether late school entry undermines children’s cognitive development, a channel through which late school entry may push some children out of school early.

2.2. Survey and Sample

The data analysed in this article are drawn from the Gansu Survey of Children and Families (GSCF), a survey targeting 2,000 rural children aged 9–12 in 2000 in Gansu. A stratified sampling strategy was adopted in 2000 to select 20 counties from all non-urban, non-Tibetan counties in the province. Five villages were then randomly selected from each of these 20 counties. Finally, within each of the 100 selected villages, 20 children were randomly selected from the full cohort of children aged 9–12 as the target children. Separate questionnaires were administered to the target children, their parents, village leaders and teachers of the schools in which these children were enrolled at the time of survey. Follow-up surveys were conducted in 2004, 2007, 2009 and 2015.

Our empirical analysis mainly uses data from the first two rounds, collected in 2000...
and 2004, when the majority of the target children were of school age (9–16). Note that our identification strategy exploits the discontinuity structure of children’s school entry age around the enrolment cut-off (31 August) to achieve identification (see the next section for details about our identification strategy), but the months of birth of children (July 1987–June 1991) were not equally distributed around 31 August. Thus, we impose a number of sample restrictions to ensure that the numbers of birth months on both sides of the enrolment cut-off are balanced, so that in the regression, the contributions of variation from both sides are comparable. More specifically, we exclude children born in November–December 1987, so that those born in July–October 1987 are balanced around 31 August. We also exclude children born after February 1991, so that those born in March 1990 to February 1991 are balanced around 31 August 1990. The resulted analytical sample includes 1,769 children.

2.3. Variables

Table 1 presents the means of the variables used in the analysis, by children’s school entry age (in years).

The outcome variables of primary interest are scores of two cognitive ability tests administered to the GSCF target children. The first test, administered in 2000, was designed by researchers at the Institute for Psychology at the Chinese Academy of Social Sciences in Beijing (Brown 2006). The test assesses three domains of

| School entry age (SA: years) | 5   | 6   | 7   | 8   | 9   | 10+ | Total |
|------------------------------|-----|-----|-----|-----|-----|-----|-------|
| Cognitive ability test scores in 2000 (standardized) | 0.33 | 0.26 | 0.13 | −0.18 | −0.26 | −0.56 | 0.00 |
| Cognitive ability test scores in 2004 (standardized)  | 0.29 | 0.14 | 0.13 | −0.15 | −0.22 | −0.45 | 0.00 |
| Girl, dummy                  | 0.40 | 0.43 | 0.42 | 0.48 | 0.59 | 0.67 | 0.46 |
| Ethnic Han, dummy             | 0.92 | 0.99 | 0.99 | 0.96 | 0.98 | 0.91 | 0.98 |
| Expected school entry age (EA: years)                  | 7.58 | 7.51 | 7.52 | 7.55 | 7.54 | 7.56 | 7.53 |
| Age at test (AT) in 2000 (years)  | 11.06 | 10.91 | 11.23 | 11.33 | 11.37 | 11.57 | 11.24 |
| Age at test (AT) in 2004 (years)  | 15.06 | 14.91 | 15.23 | 15.33 | 15.37 | 15.57 | 15.24 |
| Years of schooling (YS) in 2000 (years)                  | 5.52 | 4.59 | 4.11 | 3.41 | 2.62 | 1.48 | 3.77 |
| Years of schooling (YS) in 2004 (years)                  | 8.44 | 8.00 | 7.53 | 6.78 | 5.59 | 4.47 | 7.10 |
| Father’s education (years)                                         | 7.08 | 6.86 | 6.74 | 6.18 | 5.03 | 5.47 | 6.38 |
| Mother’s education (years)                                        | 4.44 | 4.54 | 4.43 | 3.52 | 2.05 | 2.83 | 3.89 |
| Father is farmer, dummy                                           | 0.80 | 0.75 | 0.80 | 0.77 | 0.77 | 0.82 | 0.79 |
| Mother is farmer, dummy                                           | 1.00 | 0.96 | 0.96 | 0.96 | 0.97 | 0.93 | 0.96 |
| Per capita wealth (Yuan) in log                                 | 7.51 | 7.74 | 7.72 | 7.47 | 7.18 | 7.38 | 7.58 |
| Per capita land (mu) in log                                       | 0.67 | 0.40 | 0.41 | 0.49 | 0.46 | 0.46 | 0.45 |
| Preprimary school availability, dummy                           | 0.64 | 0.71 | 0.83 | 0.75 | 0.51 | 0.49 | 0.75 |
| Observations (%)                                                  | 25 (1.3) | 233 (11.8) | 988 (50.0) | 492 (24.9) | 184 (9.3) | 55 (2.8) | 1,977 (100) |

Note: 1 US dollar = 8.27 Yuan in 2004. 1 mu = 1/15 hectares.

© 2017 The Authors. Asia and the Pacific Policy Studies published by John Wiley & Sons Australia, Ltd and Crawford School of Public Policy at The Australian National University
life skills: (i) prose literacy, which focuses on the knowledge and skills needed to understand and use information from texts that contain extended prose, especially those organized in a typical paragraph structure commonly found in materials such as editorials, news stories, brochures and pamphlets, manuals and fiction; (ii) document literacy, which focuses on knowledge and skills required to locate and use information found in qualitatively different printed materials that contain more abbreviated language and use a variety of structural devices to convey meaning, such as tables, charts, graphs, indices, diagrams, maps and schematics; and (iii) numeracy, which refers to the ability to interpret, apply and communicate mathematical information in commonly encountered situations. Partly due to dropouts, test score information is missing for 225 target children in the 2004 data.6

Note that in each of the two test years, the same test was administered to the target children at the same time, regardless of their age and grade level. Thus, even if a child’s school entry was delayed for one or more years, he/she would still take the same tests at the same time he/she is supposed to take these tests had their school enrolment not been delayed.6 In this sense, scores on these tests measure the stock of cognitive skills a child had developed and accumulated by the time of the tests. For ease of interpretation, test scores are standardized to have zero mean and unit SD by year of birth.7

The key explanatory variable is a child’s primary school entry age, which is constructed from his/her reported school entry age (in integrals of years) and months of birth. Other explanatory variables include characteristics at the child, family and village levels. Child-level covariates include child gender, ethnicity (a dummy for Ethnic Han) and years of schooling. Family-level covariates include parental education, occupation (dummies for full-time farmers, separately for fathers and mothers), household wealth (measured by the log of the value of household durables in Yuan: 1 Yuan ≈ 0.12 US dollars in 2004) and the log of per capita landholding in mu (1 mu = 1/15 hectares).8 Village-level covariates include the full set of village fixed effects. Because there is only one primary school and at most one middle school in each village in rural Gansu (Chen et al. 2017), the village fixed effects effectively control for factors that vary at the village level, such as school quality, prices (e.g. school fees and wage) and local economic conditions.

A number of informative patterns are revealed in Table 1. Notably, there appears to be a strong negative correlation between children’s school entry age and their cognitive ability scores. However, this negative association does not necessarily imply a causal effect operated from school entry age to cognitive ability because this association does not rule out the influences of other factors. For example, Table 1 suggests that late school entry is correlated with the scarcity of familial resources: Children from families with less-wealthy and less-educated parents are more likely to start school older. Therefore, to identify the causal effect of school entry age, one needs to properly control for the influences of potential confounding factors. The next section develops a method for this purpose.

3. Method

3.1. Empirical Model

Conceptually, a child’s cognitive ability tested at time $t$ ($CA_t$) is determined by his/her primary school entry age ($SA$, which captures his/her

5. Among the 225 children with no test scores information in 2004, 108 had dropped out of school by the time of survey in 2004.
6. While the GSCF data contain information on children’s scores on a math test and a reading test, we do not use these tests in the analysis because the contents of these tests are grade-specific. Thus, children with different school entry ages will take different sets of tests, which makes their test scores less comparable.
7. In the analysis, we define ‘year of birth’ to be the year (12 months) centered at the enrollment cut-off date (31 August) to balance the number of children born on either side of the cut-off.
8. Strictly speaking, only variables that were determined before children started primary school should be included in the statistical analysis. For this reason, we use the value of a household’s durable goods in year 2000, rather than its income, as a measure of household wealth.

© 2017 The Authors. Asia and the Pacific Policy Studies published by John Wiley & Sons Australia, Ltd and Crawford School of Public Policy at The Australian National University
cognitive ‘readiness’ for starting formal education at time $s$), test-taking age ($TA_s$, which captures his/her ‘maturity’ for taking tests) and years spent in school by the time of test ($YS_s$, which captures the impact of formal schooling), as well as a set of covariates $X$ that includes child, household and village characteristics discussed in the previous section.\(^9\) More specifically, the statistical relationship of interest is assumed to have the following form:

$$
CA_t = \beta_0 + \beta_1 SA_s + \beta_2 YS_s + \beta_3 TA_t + X\beta_3 + \epsilon
$$

where the error term $\epsilon$ captures the influences of all unobserved factors that may potentially affect a child’s cognitive ability.\(^10\)

Note that as pointed out by Black et al. (2011) and Crawford et al. (2014), among others, for children still in school,\(^11\) the impacts of school entry age ($SA$), years in school ($YS$) and test-taking age ($TA$) cannot be identified separately due to the perfect collinearity among these three variables: $TA = SA + YS$. For example, if test-taking age, $TA$, is held fixed, then a 1-year delay in school entry implies one less year of schooling completed from the time of test. In the regression, one of these three variables will be automatically excluded from Equation 1 if the other two are kept in the model. Thus, decisions must be made with regard to which one to exclude. To the extent that test-taking age mostly captures children’s test-taking skills, which are presumably developed at school, the impact of test-taking age presumably reflects the impact of schooling. Thus, we follow Black et al. (2011) and Crawford et al. (2014) and exclude $TA$ while keeping $YS$ in the model:

$$
CA_t = \beta_0 + \beta_1 SA + \beta_2 YS + X\beta_3 + \epsilon
$$

This setup allows us to identify the impact of late school entry on children’s cognitive ability that does not work through their reduced years of schooling. Ideally, if Equation 2 is correctly specified, such that school entry age, $SA$, is uncorrelated with the error term, $\epsilon$, the causal effect of school entry age, captured by $\beta_1$, can be consistently estimated by using ordinary least squares (OLS) regressions. However, there are a number of problems that may lead to bias in the OLS estimates of $\beta_1$.

3.2. Identification Problems

The most important problem is that children’s school entry age is likely to be endogenous. Because children in rural Gansu are supposed to enter primary school in the year when they turn age 7 by 31 August, one would expect the school entry age of children in rural Gansu to range from 7.00 to 7.99. However, as shown in Table 1, the actual range is much wider, mostly due to the prevalence of delayed school entry. This suggests that parents in rural Gansu may choose to hold their children out of school for a number of years for some reasons. If some of these reasons are unobserved to the researcher, they will enter the error term of Equation 2, thereby creating a correlation between $SA$ and $\epsilon$.

To see the problem and the possible solution more clearly, it is useful to examine the sources of variation in $SA$. In the presence of late school entry, there are three sources of variation in children’s actual school entry age. The first is the usual within-year variation due to children’s differing months of birth that are not around the enrolment cut-off, for example, a

---

9. We also tried to include other factors such as sibship size and birth order. But the results remain similar to those reported in the succeeding texts.

10. Equation 1 should be interpreted as a demand function, which incorporates parents’ behavioral responses to shocks due to changes in $SA$ and other factors, rather than a production function, which reflects the technological input–output relationship. We are unable to estimate a production function because we do not have information on all educational inputs for children’s cognitive development, especially those made prior to their primary education. To facilitate estimation, we substitute these input variables out with their exogenous determinants, such as parental education, family wealth and the prices of these inputs (which are controlled for by village fixed effects). For more discussions on problems one might encounter when estimating an education production relationship, see Glewwe et al. (2015).

11. In 2000, nearly all (96.85%) 2,000 GSCF target children were in school.
child being born in late March and starting school at age 7.42 versus being born in late April and starting school at 7.33 (Figure 1A, Source 1). The second is also a within-year variation, but it is due to some children’s birthdates’ passing the enrolment cut-off that creates a discrete jump in SA: Among children who comply with the enrolment law, those born in late August would start school at age 7.0, while those born in early September would have to start school in the next year at age 7.99 (Figure 1A, Source 2). To the extent that a child’s birth date is exogenously determined, the first two sources of variation in SA are also exogenous.

The third source of variation, however, is potentially endogenous. This source of variation is the cross-year variation that is due to parents’ decisions to hold their child out of school, say, not enrolling a child who is supposed to start school at age 7.5 until this child is aged 8.5 (Figure 1A, Source 3). This variation will be endogenous if parents’ decisions are based on factors that can affect children’s cognitive ability but are unobserved to the researcher, such as children’s cognitive skills developed prior to primary school entry.

If school entry age is the only endogenous variable, a natural solution is to isolate the first two sources of variation from the third one and use them to identify the desired causal impact. Following the recent literature (e.g. Fredriksson & Öckert 2005; Bedard & Dhuey 2006; Elder & Lubotsky 2009; Black et al. 2011; Chen 2015), we construct the expected school entry age (EA) for each child based on his/her month

12. In theory, it is possible that parents, especially those wealthy ones, can manipulate their children’s months of birth. But in a poor area such as rural Gansu, this possibility is minimal. This possibility is checked and ruled out in section 4.2.

13. Unfortunately, we do not have information on children’s exact birthdates. Thus, we are unable to perform analyses that are similar to those of McEwan and Shapiro (2008) and of Smith (2009).
of birth and then use \( EA \) as an IV for \( SA \). Because \( EA \) extracts variations in \( SA \) only from the first two exogenous sources, it is also exogenous. Using \( EA \) to instrument \( SA \), we estimate the following first-stage equation,

\[
SA = \gamma_0 + \gamma_1 EA + \gamma_2 YS + X\gamma_3 + u \quad (3)
\]

jointly with Equation 2 in a two-stage least squares (2SLS) framework.

Figure 1B visualizes the results of the first-stage regression (Equation 3) for children born in August 1987 to August 1990. Note that, exploiting the discontinuity structure in \( EA \) around the enrolment cut-off to achieve identification, our 2SLS strategy is indeed an application of the fuzzy regression-discontinuity method (Lee & Lemieux 2010). The intuition is that children born in August and September in a given year are presumably similar in all aspects, except that the latter are expected to start school almost 1 year older than the former. Because the variation in \( SA \) exacted for identification comes mainly from the contrast in \( EA \) between August-born and September-born children, we can check for robustness of the estimation results by excluding children whose months of birth are ‘far away’ from the enrolment cut-off from the analysis. Thus, some of the regressions discussed blow are performed by using the ‘near cut-off’ sample, which includes only children born 3 months before (June–August) and after the enrolment cut-off (September–November).

Note that the strategy in the preceding texts implicitly assumes that children’s years in school, \( YS \), is exogenous, which is only reasonable if all children advanced normally (i.e. no one repeated or skipped any grade). However, the fact that many children repeated one or more grades in rural Gansu (Chen 2015) casts doubt on this assumption because grade repetition may be correlated with unobserved determinants of children’s cognitive ability. Recognizing that normally older children tend to have spent more years in school, we use children’s age in years, \( AY \), to instrument their years of schooling, \( YS \). Thus, there are two first-stage equations in our 2SLS framework:

\[
SA = \gamma_0 + \gamma_1 EA + \alpha AY + X\gamma_3 + u, \quad (4)
\]

\[
YS = \eta_0 + \eta_1 EA + \delta AY + X\eta_3 + \nu. \quad (5)
\]

Finally, in some of the analyses in the succeeding texts, we allow the impact of late school entry to differ across different subgroups of interest, to help understand the potential drivers of it (if found). This is done by including in Equation 2 interaction terms between \( SA \) and a set of household and village characteristics \( X \), such as the availability of preprimary schools in the village, parental education and family wealth. To identify the effect of the interaction between \( SA \) and a given characteristic \( X \), we use the interaction between \( EA \) and that \( X \) variable as the IV for the \( SA \times X \) interaction.

4. Results

4.1. Validity of Instruments

Before turning to the main IV estimation results, it is important to assess the validity of children’s expected school entry age, \( EA \), as an IV for their actual school entry age, \( SA \). First, to serve as a valid IV for \( SA \), \( EA \) (or more fundamentally, children’s months of birth) should not be subject to parental manipulation, which implies that there should be little correlation between children’s \( EA \) and family resources. To check this, we regress a set of child and family characteristics, including child gender, ethnicity, parental education, occupation, family wealth and landholding, on \( EA \), for both the full and ‘near cut-off’ samples. Table A1 in the Appendix shows that the coefficient on \( EA \) is never statistically significant in any of these regressions, suggesting little correlation between children’s \( EA \) and family resources. While it is impossible to check the correlations between \( EA \) and unobserved characteristics, the lack of correlations between \( EA \) and the many observed characteristics reported in Table A1 suggests that \( EA \) is unlikely to be subject to parental manipulation.
## Table 2 Results of First-Stage Regressions

| Model | IV-1 | IV-2 | IV-3 | IV-4 | IV-5 |
|-------|------|------|------|------|------|
| Dependent variable | School entry age (SA) | School entry age (SA) | Years in school (YS) | School entry age (SA) | Years in school (YS) |
| Sample | Full | Full | Full | Full | Full |
| Expected school entry age (EA) | 1.218*** (0.063) | 1.182*** (0.074) | 0.168** (0.074) | 1.139*** (0.079) | 0.140* (0.081) |
| Age in years (AY) | – | 0.117*** (0.018) | 0.854*** (0.019) | 0.116*** (0.025) | 0.846*** (0.028) |
| Years in school (YS) | –0.258*** (0.021) | – | – | – | – |
| Female | 0.146*** (0.035) | 0.160*** (0.039) | –0.138*** (0.043) | 0.163*** (0.056) | –0.127* (0.067) |
| Ethnic Han | 0.144 (0.156) | 0.117 (0.160) | 0.173 (0.190) | 0.114 (0.220) | –0.192 (0.176) |
| Father’s education | –0.004 (0.006) | –0.013* (0.008) | 0.020** (0.008) | –0.015 (0.010) | 0.013 (0.010) |
| Mother’s education | –0.027*** (0.007) | –0.027*** (0.007) | 0.034*** (0.008) | –0.021*** (0.009) | 0.034*** (0.010) |
| Father is farmer | –0.019 (0.047) | –0.016 (0.051) | 0.074 (0.061) | 0.004 (0.080) | 0.122 (0.083) |
| Mother is farmer | –0.144 (0.106) | –0.145 (0.106) | 0.208* (0.123) | –0.113 (0.167) | 0.246 (0.213) |
| Per capital wealth | –0.005 (0.020) | –0.046** (0.023) | 0.076*** (0.025) | –0.066** (0.032) | 0.085*** (0.031) |
| Per capita landholding | 0.023 (0.023) | 0.003 (0.027) | –0.012 (0.030) | –0.042 (0.027) | 0.032 (0.031) |
| Village/school fixed effects | Yes | Yes | Yes | Yes | Yes |
| Partial R² of IV | 0.1678 | 0.1590 | 0.4441 | 0.2519 | 0.5420 |
| F-stat for significance of IV | 295.512 [0.0000] | 121.009 [0.0000] | 653.724 [0.0000] | 45.653 [0.0000] | 198.504 [0.0000] |
| [p-value] | 1.769 | 1.769 | 1.769 | 1.028 | 1.027 |
| R² | 0.425 | 0.408 | 0.479 | 0.488 | 0.517 |

Note: All regressions include a constant term.
Standard errors are reported in parentheses, adjusted for within-village clustering.

*Significant at 10%.
**Significant at 5%.
***Significant at 1%.

© 2017 The Authors. Asia & the Pacific Policy Studies published by John Wiley & Sons Australia, Ltd and Crawford School of Public Policy at The Australian National University
Second, to be able to identify the impact of SA, EA should have a strong predictive power for SA in the first-stage regression (Equation 4). To verify this, Table 2 reports the first-stage regressions for the three main IV models estimated in this article. Model IV-1 assumes YS to be exogenous. Models IV-2 (for the full sample) and IV-3 (for the ‘near cut-off’ sample) both allow YS to be endogenous. As indicated in the table, the estimated coefficients on EA are highly significant in all three models predicting SA (columns 1, 2 and 4), indicating no sign of the weak-IV problem (Bound et al. 1995). The IV for YS, children’s age in years, Ay, also has a strong predictive power for YS, beyond the impacts of other explanatory variables (columns 3 and 5). These results suggest that the weak-IV problem is hardly a concern here.

### 4.2. Impacts of School Entry Age on Children’s Cognitive Ability

Tables 3–5 present the main results of this article. The first set of results, reported in Table 3,

| Model       | (1)          | (2)          | (3)          | (4)          |
|-------------|--------------|--------------|--------------|--------------|
| OLS         | IV-1         | IV-2         | IV-3         |
| **Endogenous variable** | **Sample** | **Full** | **Full** | **Full** | **‘Near cut-off’** |
| School entry age (SA) | −0.145*** (0.020) | −0.160*** (0.052) | −0.164*** (0.049) | −0.112** (0.055) |
| Years in school (YS) | 0.082*** (0.019) | 0.077*** (0.029) | 0.047* (0.024) | 0.069*** (0.026) |
| Female      | 0.079** (0.037) | 0.081** (0.036) | 0.079** (0.036) | 0.085* (0.050) |
| Ethnic Han  | 0.235 (0.195) | 0.238 (0.189) | 0.243 (0.185) | 0.410 (0.308) |
| Father’s education | 0.013* (0.007) | 0.013* (0.007) | 0.013* (0.007) | 0.003 (0.009) |
| Mother’s education | 0.015** (0.007) | 0.015** (0.007) | 0.015** (0.007) | 0.016* (0.009) |
| Father is farmer | 0.004 (0.060) | 0.004 (0.058) | 0.006 (0.059) | −0.069 (0.068) |
| Mother is farmer | −0.048 (0.083) | −0.050 (0.080) | −0.049 (0.080) | −0.064 (0.135) |
| Per capita wealth in log | 0.023 (0.026) | 0.023 (0.026) | 0.026 (0.026) | 0.032 (0.034) |
| Per capita landholding in log | 0.040 (0.034) | 0.040 (0.033) | 0.040 (0.032) | 0.032 (0.043) |
| School/village fixed effects | Yes | Yes | Yes | Yes |
| Observations | 1,769 | 1,769 | 1,769 | 1,028 |
| R²          | 0.514        | –            | –            | –            |

**Note**: All regressions include a constant term. Standard errors are reported in parentheses, adjusted for within-village clustering.

### Table 4 Impacts of School Entry Age on Children’s Cognitive Ability in 2000 and 2004

| Model       | (1)          | (2)          | (3)          | (4)          |
|-------------|--------------|--------------|--------------|--------------|
| OLS         | IV-1         | IV-2         | IV-3         |
| **Sample** | **Full** | **Full** | **Full** | **‘Near cut-off’** |
| A. Test scores in 2004 | −0.117*** (0.031) | −0.130** (0.065) | −0.133** (0.063) | −0.115 (0.070) |
| Observations (with test scores in 2004) | 1,517 | 1,517 | 1,517 | 871 |
| B. Test scores in 2000 | −0.138*** (0.022) | −0.141** (0.056) | −0.143*** (0.054) | −0.108* (0.062) |
| Observations (with test scores in both waves) | 1,517 | 1,517 | 1,517 | 871 |

**Note**: All regressions include a constant term and the full set of explanatory variables presented in Table 1. Standard errors are reported in parentheses, adjusted for within-village clustering.

*Significant at 10%.
**Significant at 5%.
***Significant at 1%.
is related to the effect of school entry age on children’s (standardized) scores on the cognitive ability test taken in 2000. Column (1) presents the OLS estimates, under the assumption that children’s school entry age and their years in school are both exogenous. The OLS model indicates that, on average, a 1-year increase in school entry age reduces a child’s cognitive ability score by 0.15 SDs of the distribution of test scores. Allowing children’s school entry age, $SA$, to be endogenous, Model IV-1 (column 2) uses their expected school entry age, $EA$, as its instrument. This model indicates that a 1-year delay in school entry reduces a child’s cognitive ability score by 0.16 SDs, an effect significant at the 1% level. Model IV-2 (column 3) further allows children’s years in school, $YS$, to be endogenous and uses their actual age in years, $AY$, to instrument $YS$, which yields a very similar late-school-entry effect. To further check for robustness, Model IV-3 (column 4) restricts the analysis to the ‘near cut-off’ sample. The estimated late-entry effect ($-0.11$ SDs) is, again, significantly negative. Simply put, despite the different assumptions made by the models reported in Table 3, all estimates of the late-school-entry effect are significantly negative and quantitatively similar.

The impacts of other explanatory variables are also informative. In particular, children’s years in school have a significantly positive impact on their cognitive ability. One additional year spent in school increases a child’s cognitive ability by 0.05–0.08 SDs. This suggests that children who entered school late in the study area may have suffered from a ‘double curse’: Not only does late school entry reduce late school starters’ schooling readiness but it also deprives them of some beneficial effect of formal schooling because late starters spent less time in school compared with early starters of the same age. Parents’ years of schooling, especially those of mothers, also positively affect children’s cognitive ability (at least marginally so). Family assets (wealth and landholding) also have positive impacts, but these impacts are not statistically significant. Interestingly, girls outscored boys on the test, perhaps due to biological gender differences in the path of cognitive development.

The second set of results concerns whether the impact of late school entry persists as children advance to higher grades. Specifically, we re-estimate the regressions in the preceding texts, but this time using children’s scores on the ‘life skills’ test administered in 2004 (taken when the target children were 13–16 years old) as the outcome variable. As shown in panel A of Table 4, again, all models yield a negative late-school-entry effect (although the estimates are somewhat smaller in size than their counterparts reported in Table 3). Note that in the 2004 data, test score information is
missing for 225 children (among whom 108 had already left school). One might thus be concerned that such sample attrition may have biased in the estimates reported in panel A of Table 4. Although we are unable to directly address the potential sample-attrition bias, we assess indirectly how sample attrition may have affected the estimates. More specifically, we re-estimate the models by using the 2000 data, restricting the sample to include only children who have valid test scores in both 2000 and 2004. As panel B of Table 4 indicates, the point estimates (of the late-school-entry effect on scores of the 2000 test) are, in general, smaller in magnitude than their counterparts reported in Table 3. This suggests that the negative late-school-entry effects found in the 2004 data (Table 4, panel A) may have been underestimated (i.e. less negative) due to sample attrition because children with relatively low scores are more likely to drop out (which effectively narrows the range of the test scores). Such an underestimation, in turn, suggests that the negative impact of late school entry may persist as children advance to higher grades.

The third set of results concerns the driving channels of the late-entry-age effect. Does late school entry affect children’s cognitive ability through a physiological channel, say, by reducing their physiological potential for cognitive development? Or does it affect children’s cognitive ability through other channels? For example, children who started school late tend to be older than their classmates; thus, they are more likely to be called upon for helping with housework, farm work and wage work, by their parents. These responsibilities may, in turn, distract them from concentrating on school work.14

With information collected on children’s time allocation in both 2000 and 2004, the GSCF data enable us to examine some of these time-related channels. More specifically (given data availability), we estimate the effect of school entry age on children’s time allocated among six activities (i.e. doing homework, housework, farm work and wage work, watching television and playing with friends) in both 2000 and 2004, for the full sample (Model IV-2) and the ‘near cut-off’ sample (Model IV-3). As indicated in Table 5, all but 1 of the 24 estimates are statistically insignificant, suggesting that the negative late-school-entry effects found in the preceding texts (Tables 3 and 4) do not work through these time-related channels. It seems more likely that they work through physiological channels.

4.3. Heterogeneous Impacts of School Entry Age

The finding of a negative late school entry effect on children’s cognitive ability in the preceding texts helps explain why children who entered primary school late tend to leave school early in developing economies such as Mozambique (Wils 2004), Uganda (Uganda Bureau of Statistics and ORC Macro 2002), Zambia (Zambia Central Statistical Office and ORC Macro 2003) and rural China (Chen 2015). However, such a finding is at odds with the findings of most studies conducted in developed countries, which suggest a positive effect of late school entry (Fredriksson & Öckert 2005; Bedard & Dhuey 2006; Elder & Lubotsky 2009; Smith 2009). What, then, is driving the difference?

As hypothesized in the preceding texts, the availability of preprimary school in the community may play a role in driving the direction of the late-school-entry effect. To examine this hypothesis, we include an interaction term between children’s school entry age, SA and the availability of preprimary education in the village in Models IV-2 and IV-3 and then instrument this interaction term by the EA × preprimary school availability interaction. Panel A of Table 6 indicates that the SA × preprimary school availability interaction has a significantly positive impact on children’s cognitive ability, highlighting the importance of developing preprimary education in rural Gansu. However, the availability of preprimary school is unable to completely offset the negative impact of late school entry. This may be due to

14. Also, being older in the class may make late school starters stronger and thus better at sports than their classmates, which may also distract them from concentrating on cognitive development at school. Unfortunately, the GSCF data do not permit us to examine this possibility.
the short length of preprimary education in rural Gansu: The vast majority (82%) of children who ever attended preprimary school spent only 1 year, rather than 3 years (the common practice in urban China), in preprimary school.

Do potential substitutes for preprimary school work, then? After all, if rural parents are able to provide sufficient and proper mental stimulations to their young children at home, the lack of access to preprimary school may not be a serious problem. To explore this possibility, we examine how the effect of late school entry varies across different levels (i.e. above or below median) of parental education and family wealth, under the presumption that better-educated and wealthier parents are better at providing proper mental stimulations to their children. However, as panels B–D of Table 6 indicate, except for one case, the $SA \times$ parental education and $SA \times$ family wealth interactions have virtually no impact on children’s cognitive ability. In other words, potential substitutes for preprimary school do not seem to work in rural Gansu. This, again, highlights the importance of developing a well-functioning preprimary education system in rural China.

Finally, we check whether the impact of delayed school entry varies with the number of years delayed. This is done in a similar fashion to the models reported in panels A–D of Table 6. Specifically, we create a dummy variable for children who started primary school older than age 8. This dummy is indeed endogenous, 15. We thank an anonymous reviewer for suggesting this test.

| Outcome variable | Sample | Test scores in 2000 | Test scores in 2004 |
|------------------|--------|--------------------|--------------------|
|                  |        | IV-2               | IV-3               | IV-2               | IV-3               |
|                  |        | Full               | ‘Near cut-off’      | Full               | ‘Near cut-off’      |
| A. by preprimary school availability | $SA$ | $-0.349^{***}$ (0.104) | $-0.307^{***}$ (0.110) | $-0.359^{***}$ (0.147) | $-0.310^{**}$ (0.143) |
|                  | $SA \times 1$ (preprimary school available in the village) | $0.242^{**}$ (0.117) | $0.246^{**}$ (0.125) | $0.300^{*}$ (0.162) | $0.265$ (0.163) |
| B. by mother’s education | $SA$ | $-0.160^{***}$ (0.048) | $-0.109^{**}$ (0.054) | $-0.136^{**}$ (0.064) | $-0.110$ (0.070) |
|                  | $SA \times 1$ (mother’s education $>$ median) | $-0.008$ (0.009) | $-0.013$ (0.012) | $0.006$ (0.012) | $-0.001$ (0.015) |
| C. by father’s education | $SA$ | $-0.168^{***}$ (0.049) | $-0.120^{**}$ (0.055) | $-0.133^{**}$ (0.063) | $-0.102$ (0.069) |
|                  | $SA \times 1$ (father’s education $>$ median) | $0.009$ (0.008) | $0.008$ (0.012) | $-0.001$ (0.010) | $-0.024$ (0.013) |
| D. by family wealth | $SA$ | $-0.140^{***}$ (0.048) | $-0.117^{**}$ (0.053) | $-0.107^{**}$ (0.061) | $-0.095$ (0.068) |
|                  | $SA \times 1$ (family wealth $>$ median) | $0.010$ (0.007) | $0.018^{**}$ (0.009) | $-0.015$ (0.009) | $-0.013$ (0.012) |
| E. by school entry age | $SA$ | $-0.160^{***}$ (0.060) | $-0.114^{+}$ (0.065) | $-0.124$ (0.076) | $-0.101$ (0.078) |
|                  | $SA \times 1$ ($SA \geq$ age 8) | $-0.004$ (0.012) | $-0.003$ (0.014) | $-0.007$ (0.015) | $-0.012$ (0.015) |
| Observations     | 1,769  | 1,028              | 1,517              | 871                |

**Note:** All regressions include a constant term and the full set of explanatory variables reported in Table 1. $1()$ is the indicator function whose value equals to 1 if its argument is true and 0 otherwise. Standard errors are reported in parentheses adjusted for within-village clustering.

***Significant at 1%.
**Significant at 5%.
*Significant at 10%.
finding a statistically significant interaction between this dummy and SA will indicate possible nonlinearity in the school-entry-age effect, unless the endogeneity in this dummy leads to a substantial overestimation of the size of the school-entry-age effect (i.e. by producing a more negative impact). As indicated in panel E of Table 6, the estimated coefficients on the interaction term are virtually zero in all reported regressions, suggesting that the impact of late school entry is roughly linear in the number of years delayed.

5. Concluding Remarks

Late primary school entry prevails in many developing countries. Yet, relatively little attention has been paid to the consequences of this phenomenon in these countries. This article contributes to the literature by estimating the causal effect of school entry age on children’s cognitive development in rural Gansu, a poor northwestern area with an underdeveloped preprimary education system. Exploiting exogenous variations in children’s school entry age created by the enrolment cut-off set by China’s Compulsory Education Law, we found a negative and persistent effect of late school entry on children’s cognitive ability in rural Gansu. This finding offers an explanation of why late-school-entry effects, such as ‘being relative older among classmates’, cannot be examined by using these data. Future studies conducted in similar regions that employ school administration data to examine these channels may be fruitful. Second, the findings of this article are based on only one poor province in China. More research should be conducted in other areas in China, especially those with better developed preprimary school systems, to provide a fuller picture of the effect of late school entry in China, which may provide more useful implications for other developing countries.

References

Akyeampong K, Djangmah J, Oduro A, Seidu A, Hunt F (2007) Access to Basic Education in Ghana: The Evidence and the Issues. Project Report. Consortium for Research on Educational Access, Transitions and Equity (CREATE), Falmer, UK and Winneba, Ghana.

Bedard K, Dhuey E (2006) The Persistence of Early Childhood Maturity: International Evidence of Long-run Age Effects. *Quarterly Journal of Economics* 121(4), 1,437–1,472.

Black D, Devereux P, Salvanes K (2011) Too Young to Leave the Nest? The Effects of School Starting Age. *Review of Economics and Statistics* 93(2), 455–67.

Bound J, Jaeger DA, Baker R (1995) Problems with Instrumental Variables Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable Is Weak. *Journal of American Statistical Association* 90, 443–50.

Brown P (2006) Parental Education and Investment in Children’s Human Capital in rural China. It is comforting that the Guidelines for Poverty Alleviation and Rural Development (2011–20) recently announced by the State Council (2011) is aiming to achieve basic universalization of preprimary education in China by 2020.

Before closing, a note on two limitations of this study is in order. First, our data are household survey data, which do not contain detailed information on target children’s classmates. Thus, some potential channels that drive late-school-entry effects, such as ‘being relative older among classmates’, cannot be examined by using these data. Future studies conducted in similar regions that employ school administration data to examine these channels may be fruitful. Second, the findings of this article are based on only one poor province in China. More research should be conducted in other areas in China, especially those with better developed preprimary school systems, to provide a fuller picture of the effect of late school entry in China, which may provide more useful implications for other developing countries.

References

Akyeampong K, Djangmah J, Oduro A, Seidu A, Hunt F (2007) Access to Basic Education in Ghana: The Evidence and the Issues. Project Report. Consortium for Research on Educational Access, Transitions and Equity (CREATE), Falmer, UK and Winneba, Ghana.

Bedard K, Dhuey E (2006) The Persistence of Early Childhood Maturity: International Evidence of Long-run Age Effects. *Quarterly Journal of Economics* 121(4), 1,437–1,472.

Black D, Devereux P, Salvanes K (2011) Too Young to Leave the Nest? The Effects of School Starting Age. *Review of Economics and Statistics* 93(2), 455–67.

Bound J, Jaeger DA, Baker R (1995) Problems with Instrumental Variables Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable Is Weak. *Journal of American Statistical Association* 90, 443–50.

Brown P (2006) Parental Education and Investment in Children’s Human Capital in rural China. It is comforting that the Guidelines for Poverty Alleviation and Rural Development (2011–20) recently announced by the State Council (2011) is aiming to achieve basic universalization of preprimary education in China by 2020.
Rural China. *Economic Development & Cultural Change* 54(4), 759–89.
Caldwell JC (1967) Population: General Characteristics. In: Birmingham W, Neustadt I, Omoboe EN (eds) *A Study of Contemporary Ghana: Some Aspects of Social Structure*, Vol. 2. Northwestern University Press, Evanston.
Cascio EU, Lewis EG (2006) Schooling and the Armed Forces Qualifying Test: Evidence from School Entry Laws. *Journal of Human Resources* 41(2), 294–318.
Chen Q (2015) Ready for School? Impacts of Delayed Primary School Enrollment on Children’s Educational Outcomes in Rural China. *International Journal of Educational Development* 45, 112–28.
Chen Q, Xu J, Zhao J, Zhang B (2017) Endogenous Schooling, School Proximity and Returns to Rural Schooling in Northwestern China. *China Agricultural Economic Review* 9(2), 270–86.
Crawford C, Dearden L, Greaves E (2014) The Drivers of Month-of-Birth Differences in Children’s Cognitive and Non-Cognitive Skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 177(4), 829–69.
Glewwe P, Jacoby HG (1995) An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition. *Review of Economics and Statistics* 77(1), 156–69.
Glewwe P, Jacoby HG, King E (2001) Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis. *Journal of Public Economics* 81, 345–68.
Lee D, Lemieux T (2010) Regression Discontinuity Designs in Economics. *Journal of Economic Literature* 48, 281–355.
Lloyd CB, Blanc AK (1996) Children’s Schooling in Sub-Saharan Africa: The Role of Fathers, Mothers, and Others. *Population and Development Review* 22(2), 265–98.
McEwan PJ, Shapiro JS (2008) The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources* 43(1), 1–29.
National Bureau of Statistics (2005) *China Statistical Yearbook*. China Statistical Press, Beijing.
Nonoyama-Tarumi Y, Loaiza E, Engle PL (2010) Late Entry into Primary School in Developing Societies: Findings from Cross-National Household Surveys. *International Review of Education* 56, 103–25.
OECD and Statistics Canada (2000) *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*. OECD Publications, Paris.
Pradhan M (1998) Enrolment and Delayed Enrolment of Secondary School Age Children in Indonesia. *Oxford Bulletin of Economics and Statistics* 60, 413–30.
Seshie-Nasser HA, Oduro AD (2016) Delayed Primary School Enrolment among Boys and Girls in Ghana. *International Journal of Educational Development* 49, 107–14.
Shen B, Wang H (2003) Completion Rates of Compulsory Education and Education Finance Issues in China. *China Education and Research Network*. 14 November, viewed December 2012 <http://www.edu.cn/zong_he_199/20060323/t20060323_83365_1.shtml>.
Smith J (2009) Can Regression Discontinuity Help Answer an Age-Old Question in Education? The Effect of Age on Elementary
and Secondary School Achievement. *The B. E. Journal of Economic Analysis & Policy* 9(1), Article 48.

State Council (2011) *Guidelines on Chinese Children Development (2011–2020)*. No. 24. State Council, Beijing, 8 August, viewed December 2012 <http://www.gov.cn/2011-08/08>.

Uganda Bureau of Statistics and ORC Macro (2002) *Uganda DHS EdData Survey 2001: Education Data for Decision-making*. Uganda Bureau of Statistics and ORC Macro, Calverton.

UNESCO (2000) *Education for All. 2000 Assessment: Global Synthesis*. UNESCO, Paris, 26–28 April, viewed December 2012 <http://www.unesco.org/education/efa/efa_2000_assess/index.shtml>.

Wils A (2004) Late Entrants Leave School Earlier: Evidence from Mozambique. *International Review of Education* 50(1), 17–37.

World Bank (2001) *China: Overcoming Rural Poverty*. World Bank, Washington, DC.

Zambia Central Statistical Office and ORC Macro (2003) *Zambia DHS EdData Survey 2002: Education Data for Decision-making*. Zambia Central Statistical Office and ORC Macro, Calverton.

Zhao M, Glewwe P (2010) Determinants of Education Attainment: Evidence from Rural Area in Northwest China. *Economics of Education Review* 29(3), 451–60.

### Appendix I

#### Table A1 Correlations between Children’s Expected School Entry Age and Observed Characteristics

| Sample                      | (1) Full          | (2) ‘Near cut-off’ |
|-----------------------------|-------------------|--------------------|
| (1) Girl                    | -0.014 (0.039)    | -0.004 (0.042)     |
| (2) Ethnic Han              | 0.010 (0.007)     | 0.016 (0.010)      |
| (3) Father’s education      | 0.025 (0.246)     | 0.030 (0.256)      |
| (4) Mother’s education      | -0.349 (0.244)    | -0.326 (0.253)     |
| (5) Father is farmer        | 0.031 (0.035)     | 0.020 (0.035)      |
| (6) Mother is farmer        | 0.002 (0.018)     | 0.002 (0.019)      |
| (7) Log (per capita wealth) | 0.047 (0.076)     | 0.040 (0.078)      |
| (8) Log (per capita landholding) | -0.109 (0.067) | -0.085 (0.075) |
| Observations                | 1,799             | 1,041              |

*Note*: The table reports the main results of a total of 16 regressions. In each of these regressions, the only explanatory variable is a child’s expected school entry age (EA). The variable in each row is the outcome variable of the regressions reported in that row. Robust standard errors are reported in parentheses, adjusted for intra-village clustering.