The Effect of School Interruption on Child Human Capital Accumulation

Magda Tsaneva**

October 20, 2013

Abstract

This paper examines human capital accumulation during the economic crisis in Indonesia in late 1990s. The study focuses on the effect of school interruption on child cognitive outcomes. I use several different empirical approaches, including estimation of a value-added and an IV model, as well as an endogenous switching regression model, to account for the potential endogeneity in the decision to stay in school. Results indicate economically and statistically significant effects of school dropout on both math scores (obtained through formal training) and picture scores (a general measure of capabilities). This suggests that permanent school interruption during the crisis affected children who had the potential to do well in school. Sensitivity analyses show that learning takes place in different environments and children who work experience a smaller reduction in their picture scores. In addition, some suggestive evidence is presented on the fact that even temporary school interruptions matter for child cognitive development.

*I am very grateful to Kenneth Leonard, Richard Just, Sergio Urzua, Dana Andersen, Joe Maher, and participants at the AREC Student Workshop for helpful comments and suggestions.
**Department of Agricultural and Resource Economics, University of Maryland, College Park, MD.
e-mail: mtsaneva@arec.umd.edu

1. Introduction

Poverty has been shown to have a big impact on child cognitive development. Ghuman et al. (2005) find that Filipino children of 0-36 months of age have higher language scores when the household has higher physical and human capital assets, while Paxson and Schady (2007) present evidence on the positive association between socio-economic status and cognitive ability of pre-school children in Ecuador. Using data from the US, Guo and Harris (2000) find that the relationship between poverty and cognitive development acts through five different pathways: cognitive stimulation, parenting style, physical environment, child's ill health at birth, and ill health in childhood. In addition, resource constraints may be associated with inability to finance formal education and higher vulnerability to income shocks. For example, Jacoby and Skoufias (1997) show that due to financial market imperfections in rural India and parental inability to borrow against future income, school attendance fluctuates during periods of low income. Macroeconomic crises in low-income countries have also been associated with lower enrollment and attendance rates (Ferreira & Schady, 2009). Less is known, however, about the direct effects of school dropout on child cognitive development.

This paper uses data from the panel Indonesia Family Life Survey (IFLS) to examine the impacts of shocks to schooling during the economic crisis in late 1990s on child human capital accumulation. The IFLS is one of the very few surveys in developing countries that administers a couple of cognitive tests to all individuals between the ages of 7 and 25. This allows me to study the consequences of school interruption on mathematics scores as well as general cognitive skills (measured by Raven's Progressive Matrices assessment). If the decision whether to stay in school is correlated with the potential outcomes, the ordinary least squares (OLS) estimates of the effect of school interruption will be biased. Therefore, I use a value-added model to account for unobserved ability and a more general instrumental variable model with schooling cost shifters as instruments for the decision to drop out of school. In addition, I estimate an endogenous switching regression model to account for heterogeneity in unobservable ability and household constraints that may affect both the decision to interrupt school and the final cognitive outcomes. I find economically and statistically significant effects of school interruption on both measures of skills. The results suggest that children who dropped out of school during the crisis were not all low ability children but were poor children who had the potential to do well in school had they stayed. Sensitivity analyses further show that children who dropped out of school to work experienced a lower decrease in their cognitive scores, suggesting that learning occurs in different environments. While the study focuses on permanent school dropout, some suggestive evidence is presented on the fact that even temporary school interruptions matter for child cognitive development.

Whether school dropout affects cognitive outcomes is a priori uncertain. A recent paper on the US finds that, on average, high school dropout would have a large negative effect on math scores (Li et al., 2004). However, dropouts themselves do not suffer a reduction in cognitive outcomes because of the low potential they have. Thus, dropouts select out of schooling based on the marginal benefits of schooling which differ between students. In the human capital literature, many studies on selection have been inspired by the Roy model and the theory of worker comparative advantage in a world where different jobs require different skills (Roy, 1951). In a seminal contribution to the literature, Willis and Rosen (1979) show for the US that the decision whether to attend college depends on the expected lifetime earnings under the two options (i.e., college or no college). The paper also finds evidence of comparative advantage, showing that college graduates would have been worse off if they hadn't gone to college compared to people who did not go to college in the first place, and vice versa. In the Indonesian setting, however, it is unclear whether dropout has a negative effect or no effect. If dropout is induced by low income rather than low skills, dropout may be expected to have a large negative effect on the cognitive outcomes of the children who do not complete their education, unlike in the US.

If resource constraints matter and school dropout affects potential cognitive outcomes, it would be important to prevent school interruptions for low income children, especially during income shocks when the proportion of high ability but low income children among dropouts increases. Keeping these children in school would be an effective policy to improve labor market outcomes of future generations. For example, Duflo (2001) finds that there are significant returns to education in Indonesia, ranging from 6.5% to 10.8%. Dropout can therefore be expected to have important labor market consequences for these children if they wouldn't have dropped out otherwise. In addition, the literature on labor markets in the US, as well as developing countries (Glewwe, 2002; Hanushek & Woessmann, 2008), suggests that schooling quality and cognitive outcomes have independent effects on labor outcomes. In Indonesia, too, Suryadarma (2010) finds that a 1 SD increase in picture scores increases monthly wages by 7% on average, although he does not find an effect of math scores on wages, conditional on education. Thus, children who dropped out of school during the economic crisis in Indonesia may have lower expected earnings as adults both because of less education and lower cognitive achievements.

Previous research has documented the effect of income shocks on enrollment rates and years of education, as well as the effect of socio-economic conditions on cognitive skill development for young children. My study contributes to the literature by 1) studying cognitive development in school-age children (between the ages of 7 and 18), and 2) examining the effects of school interruptions during a time of a significant income shock. The next section discusses the schooling system in Indonesia, provides some background on the economic crisis of late 1990s, and describes the data used for this analysis. In section 3, I present the conceptual model and its empirical specification. Next, section 4 contains information on the sample selection and variable definitions. Section 5 discusses results on determinants of school dropout, while section 6 presents the findings from the analyses on the effect of school interruption. Section 7 concludes.

2. The Indonesian context

In Indonesia, there are three educational levels – primary school for ages 7 to 13, then three years each for junior high school and senior high school. Nine years of schooling are compulsory. However, while primary school enrolment is universal, junior high school enrolment in 1997 was 72.2%, while senior high school enrolment was only 46.5% (Lanjouw et al., 2001). Repetition, especially during primary school, is fairly common with 14.2% of students in grade one and 4.5% in grade five of primary school repeating the grade in 1993 (Jones & Hagul, 2001). A considerable heterogeneity in types of schools exists at all levels. At the primary school level, students choose between public and private schools, which can be secular or religious. At the junior high and senior high school level, schools are further classified into general curriculum schools or vocational schools. After each level, students take state exams, EBTANAS, and their performance determines placement in higher level schools. School attendance usually requires an annual registration fee, as well as monthly fees. Even in public schools, where annual fees have been abolished, parents are expected to pay monthly fees (Suryadarma et al., 2006).

In the two decades prior to the Asian financial crisis, the proportion of the population living under poverty in Indonesia fell from 40.1% in 1976 to 11.3% in 1996 (Lanjouw et al., 2001). During that time the government invested heavily in education, increasing the number of schools and improving enrollment rates of primary school students from 69% in 1973 to 83% in 1978 (Duflo, 2001). By 1986, universal primary school enrolment was reached. Then, in late 1997, the Asian Financial Crisis hit. Inflation reached 80% in 1998, real wages fell by 40%, and total per capita expenditure declined by 23% (Frankenberg et al., 2003). Using data from the IFLS, Thomas et al. (2004) find that during the crisis parents pulled children out of school as a way to cope with the income shock. They estimate that the proportion of 8-13 year olds not enrolled in school in 1998 was close to 20% higher than what a linear trend would have predicted. Parental inability to pay school fees or the need for child labor may lead to school dropout, school interruption between school years, or school interruption during the school year and grade repetition later on.

This paper's analysis on the effects of school interruption is based on two waves of the longitudinal IFLS dataset, which provides information both before and after the crisis. The second wave of the survey, IFLS2, took place between July and November 1997, while the third wave, IFLS3, took place between June and December 2000. The survey follows households and individuals over time. It documents all children in the household and provides child-specific information on education attainment and parental education spending for the academic years 1997/1998 and 1999/2000, including annual registration fees, monthly fees, and other expenses on food, travel and extracurricular activities. Further, it administers cognitive tests during the time of the survey to all individuals aged between 7 and 25. The cognitive tests in 1997 include forty questions each on Indonesian language and mathematics. The tests in 2000 include a Raven's Progressive Matrices assessment and five questions on mathematics.[1]

3. Conceptual framework

In order to estimate the effect of school interruption on cognitive outcomes, a production function for human capital accumulation could be estimated. Since current outcomes are a function of all past investments, direct production function estimation is complicated by missing and endogenous inputs. In order to deal with this issue, a popular approach in the literature has been the estimation of value-added models, where current achievement is presented as a function of current investment and past stocks. Thus, past test scores can be used to account for missing inputs, as well as unobserved ability. This model, however, yields unbiased results only under very restrictive assumptions about the degree of correlation of the error terms in test scores over time (Todd & Wolpin, 2003). In addition, the value-added model assumes a non-age varying technology of human capital production. Cunha and Heckman (2007) suggest that the production of skills and abilities is characterized by self-productivity, where skills from one period are carried over to another period, as well as dynamic complementarity, where skills in one period raise the efficiency of investment in future periods. They argue that because of these features of the human capital production functions, there are sensitive periods of investment. Cunha et al., (2010) show empirically for the case of the US that production functions do vary by age. The evidence accumulated on the long-lasting effects of early life shocks also serves to support this argument (e.g., Currie & Hyson, 1999; Almond & Currie, 2011). Thus, while I present results using a value-added representation of the production function, the main results of the paper are based on a reduced-form estimation of cognitive achievements.

If parents are utility maximizers, choosing inputs in the production of human capital of their child, the first-order conditions of their optimization problem yield demand functions for the various inputs, governed by prices and income. By replacing inputs in the production function with these expressions, I can write the reduced-form demand functions for human capital as a function of prices and income, as well (Rosenzweig & Schultz, 1983). If prices are constant across observations, their effect is

[1] Raven's picture test is believed to provide a general measure of intelligence and other than in studies using the IFLS data (such as Sim et al, 2012, and Cas, 2011), it has recently been used in studies by Malamud et al (2011) in Romania, and Powers (2010) in Mexico.

absorbed in the intercept term (Todd & Wolpin, 2007). Thus, I estimate cognitive test scores as a function of income to account for missing inputs, as well as other variables that may affect unobserved parental preferences for inputs and household constraints such as child gender, mother years of education, household composition (i.e., presence of older sibling in the household), urban/rural area of residence, and province fixed effects. In addition, in order to account for the fact that children in the sample are of different schooling levels, I control for years of education in 2000. Since the survey took part over a six-month period, I also include dummy variables for month of interview.

The main model I attempt to estimate is: $Y_i = \beta X_i + \alpha D_i + U_i$, where Y_i is the measure of cognitive outcome of child *i* in the year 2000, X_i is a function of child and household characteristics as discussed above, and D_i is a dummy variable that takes the value of 1 if the child interrupted schooling, and is 0 otherwise. If Y_1 is the potential cognitive score when individuals are "treated" (i.e., have a school interruption) and Y_0 is the potential cognitive score when individuals are not "treated" (i.e., stay in school), then the potential outcome can be written as a function of the observable characteristics, *X*, and the unobservable error terms $U_{i,1}$ and $U_{i,0}$, where $Y_{i,1} = X_i\beta_1 + U_{i,1}$ and $Y_{i,0} = X_i\beta_0 + U_{i,0}$. If the latent propensity to be treated, explained by some exogenous shifters *Z* and an error term *V*, is $I_i = Z_i\gamma + V_i$, then I define the observed binary outcome of school interruption as $D_i = 1$ if $I_i \geq 0$ and $D_i = 0$ if $I_i < 0$. The observed outcome can then be re-written as:

$$Y_i = (1 - D_i) * Y_{i,0} + D_i * Y_{i,1} = X_i\beta_0 + D_i(\beta_1 - \beta_0)X_i + U_{i,0} + D_i\big(U_{i,1} - U_{i,0}\big).$$

If there is no heterogeneity in unobservables ($U_{i,1} = U_{i,0} = U_i$) and the decision to stay in school is not correlated with the potential outcome ($Cov(V_i, U_i) = 0$), this model can be estimated using OLS. However, if the error term U_i contains unobservable characteristics such as child ability, parental preference for education, insurance capacity, resource constraints, then these are likely to also affect the decision whether to pull the child out of school as a coping mechanism during an income shock (i.e., $Cov(V, U_i) \neq 0$). In this case, even when $U_{i,1} = U_{i,0} = U_i$, the variable D_i and the error term U_i in the equation above will be correlated, and the OLS estimator will be biased. An instrumental variable approach can be used in this case to account for the endogeneity.[2]

[2] If the unobservables are defined at the household level, i.e., the error term is comprised of a random error and a family error component only, then a siblings fixed effects (FE) estimation would also yield unbiased results. The siblings FE estimation, however, would identify the effect of schooling based on variation between siblings, which is not large. In addition, this method assumes away any individual-level error component. Child FE could be used in the case with both a family and an individual error component. The data used in the analysis, however, is not suitable for applying this approach because of the difference in the cognitive assessments in 1997 and 2000 in terms of both type and number of questions. Child fixed effects would also assume away any time-varying unobservables.

If $U_{i,1} \neq U_{i,0}$, any instrument which is a good predictor of the propensity to interrupt school will by definition be correlated with D_i and thus correlated with the error term in the observed outcomes equation, $U_{i,0} + D_i\left(U_{i,1} - U_{i,0}\right)$. When children are heterogeneous in terms of ability and family effects, and when the decision of staying in school is endogenous and related to these ability and family effects, an endogenous regime-switching regression model can be used to estimate the effect of dropout on child cognitive scores. Using the notation above, this model is estimated by maximizing the following likelihood function:

$$L = \prod_{i=1}^{n} [f(Y_i|D_i = 1) * P(D_i = 1)]^{D_i} [f(Y_i|D_i = 0) * P(D_i = 0)]^{1-D_i}$$

where $P(D_i = 1)$ stands for the probability of dropping out of school and $f(Y_i|D_i = 1)$ and $f(Y_i|D_i = 0)$ are the conditional distributions of the observed outcomes. Under the assumption of joint normality of the error terms, the first term in the likelihood function can be represented as:

$$f(Y_i|D_i = 1) * P(D_i = 1) = f\left(Y_{i,1}\middle|I_i \geq 0\right) * P(I_i \geq 0) = f\left(Y_{i,1}, I_i\middle|I_i > 0\right) = \int_{-\infty}^{-Z_i\gamma} f\left(U_{i,1}, V_i\right)dV_i =$$

$$f\left(U_{i,1}\right) \int_{-\infty}^{-Z_i\gamma} f\left(V_i\middle|U_{i,1}\right)dV_i = \frac{1}{\sigma_1} * \phi\left(\frac{U_{i,1}}{\sigma_1}\right) * \int_{-\infty}^{-Z_i\gamma} \left(\frac{1}{\sqrt{1-\rho_1^2}} * \phi\left(\frac{V_i - \rho_1\left(\frac{U_{i,1}}{\sigma_1}\right)}{\sqrt{1-\rho_1^2}}\right)\right),$$

where $\sigma_1$is the standard deviation of the error term $U_{i,1}$, ρ_1 is the correlation coefficient between V_i and $U_{i,1}$, and $\phi(.)$ is the standard normal probability distribution function. Similar manipulations are applied to the second part of the likelihood function to yield a final log-likelihood function

$$lnL = \sum_i^n D_i * \left[ln\phi\left(\frac{U_{i,1}}{\sigma_1}\right) - ln\sigma_1 + ln\Phi\left(\frac{\left(Z_i\gamma + \frac{\rho_1 U_{i,1}}{\sigma_1}\right)}{\sqrt{1-\rho_1^2}}\right)\right] +$$

$$+(1 - D_i) * \left[ln\phi\left(\frac{U_{i,0}}{\sigma_0}\right) - ln\sigma_0 + \ln\left(1 - \Phi\left(\frac{\left(Z_i\gamma + \frac{\rho_0 U_{i,0}}{\sigma_0}\right)}{\sqrt{1-\rho_0^2}}\right)\right)\right].$$

This likelihood function accounts for the fact that the error terms are possibly not independent. In the case when both correlation coefficients, ρ_1 and ρ_0, are zero (i.e., $Cov(U_1, V) = 0$ and $Cov(U_0, V) = 0$), OLS estimation of the effect of dropout would be consistent. If at least one of the coefficients is found to be statistically different from zero, then an endogenous switching model is appropriate. If the two coefficients have opposite signs and $\rho_1 > 0$ while $\rho_0 < 0$, the results provide evidence for comparative advantage. The reason is that individuals who choose to be in regime 1 (i.e., have high values of the unobservable V term) also have high outcomes in regime 1 (high V implies high U_1 when $\rho_1 > 0$). Similar logic follows for individuals in regime 0. A model of comparative advantage is plausible when the final outcome of interest is wage and wages could be earned in jobs requiring different skills. In this study, however, the outcome of interest is cognitive skill. There could be comparative advantage if some children learn best at school, while other children learn best in other settings, e.g., household work. Work, however, is unlikely to completely replace the formal training in mathematics, for example. In that case, we may expect that children who do well in school would do well out of school, too, and children who do poorly in school, would do no worse out of school if they drop out (but still worse than the average).

Various studies on selection based on unobservable differences in ability, tastes, expectations, capacity to finance education and other family effects have examined the effect of schooling choices on earnings, allowing for differences in the returns to education across groups (e.g., Averett & Burton, 1996; Borjas, 1987; Cameron & Heckman, 2001; Willis & Rosen, 1979). When the outcome of interest is child cognitive development (rather than wages), selection into schooling based on unobservables can be analyzed in the context of children having different human capital production functions (rather than earnings functions). This is implied by the potential outcomes model above, which does not constrain the coefficients under the two treatments to be equal. While it is important to examine the different determinants of cognitive development for the two groups of children (those with interrupted vs. continuous schooling), the main focus of this study is quantifying the effect of school interruption. To do so, I estimate several treatment parameters of interest. Since the coefficients resulting from the maximum likelihood estimation (mle) are corrected for the selection process, the average treatment effect of school interruption on cognitive outcomes can be calculated as: $ATE = E[Y_1 - Y_0] = \left(\beta_1^{mle} X_1 - \beta_0^{mle} X_0\right)$.[3] Next, the effect of treatment on the treated is the difference in potential outcomes for the individuals who had a school interruption, predicting their expected counterfactual cognitive scores (if they had not dropped out of school) based on observed characteristics and the returns to those characteristics in the other regime. It can be expressed as: $TT = E[Y_1 - Y_0|D = 1] = (\beta_1^{mle} - \beta_0^{mle}) * X_1$. Finally, the treatment on the untreated is the difference in predicted potential outcomes for those individuals who stayed in school: $TUT = E[Y_1 - Y_0|D = 0] = \left(\beta_1^{mle} - \beta_0^{mle}\right) * X_0$. If people select into schooling based on comparative advantage so that low-ability students drop out while high ability students stay in school, then the TT effect should be small as was found for the US by Li et al (2004). In the context of school dropout during the Indonesian crisis, however, it may not be comparative advantage but rather resource constraints which govern school decisions. For example, the marginal cost of schooling (relative to the

[3] The standard errors for all three treatment parameters, ATE, TT, and TUT are estimated by bootstrapping the maximum likelihood estimation with 100 replications.

marginal benefit from consumption) may be high for poorer families and these parents may withdraw children from school when hit by a shock. These children may initially be of average rather than low ability. Further, they may have a hard time learning outside of school because of lack of parental compensatory investment and cognitive stimulation (since they come from poor families). In that case, the TT effect may be large.[4]

4. Sample selection

I restrict the sample to nuclear households from 1997 only, excluding extended families or families where young members (under 25) are not children of the household head or spouse. The reason for this restriction is to make the sample of households more homogeneous, to ensure that parents have decision-making ability over child investments, and to avoid issues of resource allocation among cousins.[5] At the individual level, I restrict the sample to children of school age (between 7 and 18 years of age) in both 1997 and 2000, who attend school in 1997. In addition, individuals who are not interviewed in the academic calendar year 2000/2001, provide inconsistent answers (for example, reporting "still in school" in 2000, but also reporting having dropped out of school permanently the year before, or reporting years of schooling in 1997 higher than years of schooling in 2000), or have missing values for any of the variables used in the analysis are excluded.

I estimate the effect of school interruption on two cognitive outcomes at the time of the survey in 2000: Raven's progressive matrices score (a measure of general intelligence) and math scores. In order to avoid the problem of having a limited dependent variable (since the number of correct and wrong answers is predetermined), I standardize both test scores, as is common practice in the literature (e.g., Malamud & Pop-Eleches, 2011; Paxson & Schady, 2007). Since the sample of children includes children between the ages of 7 and 18, I standardize the scores by age (with a mean 0 and a standard deviation 1) in order to account for the non-linear relationship between scores and ages. Age-normed scores have previously been used by Todd and Wolpin (2007), Paxson and Schady (2007) and others. I control for gender in all regressions to account for the potentially different distribution of scores by gender.

I classify children who were in school in 1997 but were not in school in the academic year 1999/2000 as permanent school dropouts. Their outcomes are compared to children with continuous enrollment between the two survey waves, the comparison group (2,020 observations). In the sample comprised of the children with continuous schooling and children who drop out, the proportion of school

[4] Note that IV estimation assumes no heterogeneity in unobservables. Therefore, with IV, ATE=TT=TUT.
[5] In a related study, I look at parental education investment in this period and find that, on average, parental resource allocation among siblings is not affected by child ability. This may not be the case, however, for extended family members.

dropouts is 17.18% (347).[6] This fraction of permanent dropouts may not be representative of the population because of the various sample restrictions imposed on the data, as previously discussed. Yet, my findings are consistent with Thomas et al. (2004) who show that in 1998 the proportion of male children not enrolled in school at age 12 is 8.2%, which increases to 21.5% for 14-year old males, and is higher for older males and for females at all ages. Table 1 presents descriptive statistics of the characteristics of the dropout vs. the comparison group. Dropouts are on average 2.9 years older and have 2.35 more years of education. Yet, they have significantly lower math scores in 1997, which deteriorate by 2000, while the scores of the comparison group increase over time. Dropouts are significantly poorer in 1997 than the comparison group, although the average per capita expenditure for the two groups is not significantly different in 2000.

Thomas et al. (2004) report that the proportion of 8-13 year olds not enrolled in school in 1998 is close to 20% higher than what a linear trend would have predicted: a change, which could be attributed to the crisis. One caveat of my study is that the effect of school dropout on cognitive outcomes is identified from two groups of children: 1) children who would not have interrupted schooling had the crisis not hit, as well as 2) children who would have interrupted anyway. If a crisis-related instrument for the decision to drop out is found, the effect on the population of "compliers" (the first group of children) can be isolated. However, potential instruments for the size of the crisis, such as inflation rates at the province level or the average change in wage rates for the household head's occupation, are all found to be directly correlated to pre-crisis cognitive outcomes or to fail the exogeneity test for instruments. In addition, the instruments do not have a monotonic effect on school interruption because while the high inflation increased the direct costs of schooling, it also decreased the opportunity cost of schooling, as real wages fell and child labor became less profitable. As a result, the instruments I use for school interruption are not direct measures of the crisis (more on this below). This implies that the effect I will be identifying is an average for the two groups of children. I believe that the estimates are likely to be downward biased if the children in the second group select out of schooling due to low ability and potential, while the children in the first group are forced into school interruption, even if they have good chances of benefitting from further education. As argued earlier in the paper, however, even children in the second group may be dropping out because of resource constraints, rather than low ability, in which case the downward bias will be small.

---

[6] This sample does not include children who had a temporary school interruption between 1997 and 2000. More details on this group of children are provided below.

5. Determinants of school interruption

The previous sections argued that school interruption is potentially endogenous in a model of cognitive achievements. As explained earlier, the cognitive test scores are estimated as a function of child gender and years of schooling, and various characteristics at the household level. In this framework, identifying instruments include variables that affect the local labor market for child labor and the costs of schooling. Since many children may engage in informal labor and receive non-market wages, Gunnarsson et al. (2006) identify child age as a suitable instrument for the decision whether to drop out of school. Child age is a proxy for child marginal productivity of labor. Higher age should also increase the probability of working in the presence of laws or social norms for child labor. Based on the previous literature, the effect of the crisis is expected to vary by age and therefore, I use age as an identifying instrument.[7] Other factors that are expected to affect the schooling decision are the per capita expenditure of the household in 1997, an indicator for non-zero school registration fees in 1997, and distance to school in 1997. Household expenditure prior to the crisis can serve as a proxy for the capacity to finance education. The crisis has been shown to affect disproportionately households in the bottom of the expenditure distribution (Thomas et al., 2004) . Thus, this variable serves as a good proxy for the effect of the shock. Per capita expenditures (pce) are a function of household size, prices, and income and changes in pce are not necessarily a predictor of changes in income (Thomas et al., 2004). This could explain why conditional on current expenditures, past expenditures are not significant predictors of current test scores (while past education investments are). Next, I use registration fees and distance to school as determinants of the cost of schooling. I find that unlike monthly fees, registration fees do not affect school quality (when measured by test scores). While the validity of the instruments is based on the theoretical considerations described above, the instruments (both individually and in groups) pass the empirical exogeneity tests (which rely on the assumption that at least one of the instruments is valid).

The first-stage regression of school dropout provides evidence for the relevance of the excluded instruments. The R-squared for the first stage equation is 0.374 and the excluded instruments have a high predictive power with F-statistics higher than 10 ($F(4, 1687)=180$, p-value<0.001). As expected, age is a significant positive predictor of permanent dropout. One year increase in age is associated with an increase in the probability of school interruption of 0.151 percentage points, or 88 percent. Females are more likely to drop out than males and having an older sibling who lives in the same household is associated with a lower probability of dropping out. This finding is consistent with Thomas et al. (2004), who find that older females would drop out of school to help support the family. As expected, school

[7] Since the scores are standardized by age, age does not enter the main outcome equation directly, as in Todd and Wolpin (2007).

distance reduces the probability of staying in school. Further, the log of per capita household expenditures in 1997 is found to be a significant predictor of permanent dropout, while registration fees have no effect. At the same time, sensitivity analyses described below show that non-zero registration fees have a significantly negative effect on temporary school interruptions, while per capita expenditures do not. One potential explanation of these findings is that long-term constraints (proxied by expenditure levels) matter for permanent school dropout, while short-term constraints matter for temporary interruptions.

6. Effect of school interruption on cognitive outcomes

6.1. Basic model

The effect of permanent school interruption on cognitive outcomes is first estimated using OLS. The results are presented in Table 2. The effect of dropout is large for both math (0.44 SD) and picture scores (0.26 SD). Thus, while school interruption has higher effect on school performance, it also affects general cognitive skills. These estimates are comparable to the findings in the related literature. Studying child labor in Latin America, Gunnarsson et al. (2006) find that a child working 1SD more than the mean will have 0.408 SD lower math scores. In the US, Li et al. (2004) find an average treatment effect of school dropout on math scores of -1 SD. The effect on picture scores can be compared to other papers that have used the Raven progressive matrices assessment. Malamud and Pop-Eleches (2011) find that computer use in Romania increases picture scores by 0.3 SD. Using data from the IFLS, Cas (2011) studies the effect of the safe motherhood program on child cognitive development. She finds that an increase in prenatal care through the safe motherhood program in Indonesia increases child cognitive scores between 0.15 to 0.36 SD depending on the intensity of exposure to the program.

While OLS results are informative, the estimates are not causal if dropout is endogenous. If the decision whether to stay in school is a function of the difference between marginal benefits (MB) and marginal costs (MC), lower MB and higher MC have the same effect of raising the probability of dropping out. If a child dropped out of school because of low skills (low MB) then under certain conditions, controlling for past scores may help eliminate the bias, as discussed previously. The value-added models in columns (2) and (5) yield support for this hypothesis as the estimates of the dropout effect are lower than (but close to) the OLS estimates. Similarly, if the endogeneity in the schooling decision was based only on omitted ability bias, the IV estimates could be expected to be lower than OLS estimates. If some children dropped out because of resource constraints (high MC), however, then the IV estimates may be larger. The second-stage results are consistent with this second alternative: the IV estimates are about two times larger than the OLS estimates for math scores and about three times larger for picture scores. At the same time, if there is heterogeneity in the MB and MC in the population, the IV estimate of the average treatment effect will not be consistent. In the IV estimation, some observations may be given disproportionately larger weights, accounting for the large size of the effect.

The next section presents results using an endogenous regime switching model which allows for differences in unobservable characteristics and enables calculation of different treatment effects: the average treatment effect (ATE), the treatment on the treated (TT) and treatment on the untreated (TUT). It also addresses the fact that the OLS regression models restrict the coefficients for both groups of children to be equal, which may introduce an interaction bias if the human capital production functions of children who drop out are different than the production functions of those who stay in school. One reason why this may be the case is that, as shown in Table 1, children who drop out are significantly older that children who don't and production technologies may vary over age, as discussed earlier.

6.2. Endogenous Switching Model

Table 3 presents the outcome equation results from the maximum likelihood estimation of the regime switching models where "regime 1" denotes outcomes for children who dropped out of school and "regime 0" – for those who stayed in school. In regime 0, female gender is associated with 0.09 SD higher math scores than male and each additional year of mother's education increases scores by 0.03 SD. Higher income (as proxied by expenditures) translates into better math outcomes only for those who stay in school. In contrast, income is not a significant predictor of picture scores. Further, while having an older sibling in the household does not affect math performance, it does increase picture scores for children who stay in school, suggesting that children (especially younger ones) may learn from siblings. In addition, unlike the case of math scores, females who stay in school do worse in picture scores than males.

For both math and picture scores, years of education are significant determinants of cognitive outcomes in both regimes. Schooling becomes more important for those who drop out. The effect of an additional year on math scores for dropouts is 2.3 times larger than for non-dropouts (0.145 SD vs. 0.064 SD). For picture scores, that ratio is 1.9. Consistent with age-varying production functions, this result implies that accumulated years of schooling are more important for older children, while current investments (proxied by mother's education and per capita expenditure in 2000) may be more important for the younger children. The stronger link between years of education and cognitive skills in the older, dropout, group may also be due to any selection into schooling that has already happened and is not modeled here. For example, many of the younger children in the comparison group may not continue education because of poor skill or lack of investment, while children who have reached higher grades may already be self-selected, especially since children need to pass state exams at the end of each schooling level (e.g., primary and junior high school) in order to continue to the next level.

As discussed earlier, the term ρ_1 stands for the correlation coefficient between the error term in outcome equation 0 (U_0) and the error term in the choice model (V), and similarly for the term ρ_1. In both the math and picture score models, only ρ_0 is significant. Still, this yields support for an endogenous switching regression model. A positive correlation coefficient ρ_0 suggests that children who do not interrupt schooling have higher cognitive outcomes in the interrupted regime than average. On the other hand, children who interrupt, have average outcomes in their regime (because ρ_1 is not significantly different from 0). This result suggests that high ability students remained in school, while both low ability as well as average students dropped out. This finding implies that endogeneity of schooling is not based on unobserved ability only.

In order to understand better the effect of school interruption on the two groups of children, I next examine three different treatment effect parameters (in Table 4). If the decision to drop out of school is not correlated with unobservable characteristics (i.e., exogenous switch) the treatment effects on treated and untreated populations should be similar. Any difference in parameter estimates would be due to differences in observable characteristics for the two groups of children. With an endogenous switch, the treatment parameters may differ due to differences in unobservable characteristics as well. Table 4 shows that both TT and TUT are negative and significantly different from zero with a TT of -0.68 SD and TUT of -0.49 SD for math scores, and a TT of -0.67 SD and TUT of -0.54 SD for picture scores. Thus, unlike the case of US dropouts examined by Li et al. (2004), the negative effect of school dropout for those who drop out is large and almost identical for both school-acquired skills (math score) and general skills (picture score). Interestingly, the TT effect is even larger than the TUT effect. This suggests that if the children who stayed in school were forced to drop out, then the loss in cognitive scores they experience would be smaller. The result has several possible explanations. First, based on the value-added estimates, it is plausible that some children drop out because of low unobserved ability (but not all, or else, the effect of dropout would be closer to zero, as in the Li et al. (2004) study). Yet, if there is selection into higher levels of schooling, the comparison group of children who stay in school may include children with even lower unobserved ability who may drop later on at an even earlier schooling level. A second possibility is that some of the children who drop out do so because of resource constraints rather than low ability. Had they stayed in school, they would have done well. Even if the two groups of children had the same unobserved ability, those who had continuous schooling would have lost less in the counterfactual world. These children do not face unobserved resource constraints and if forced to drop out, they may still be able to continue learning, e.g., they may have more books at home or their parents may be more educated and may be able to provide more cognitive stimulation.

6.3. Dropping out of school to work

In the analysis presented so far, school dropout is found to have a large negative effect on cognitive outcomes. This effect, however, may be an underestimate if skills can be acquired not only in school but also outside of school. About 54% of the sample children who dropped out of school report having worked in the past year (compared to 2.4% of the comparison group). If working increases cognitive outcomes and dropping out of school is positively correlated with the probability of working, then the estimates of the negative effect of dropping out of school will be downward biased. In order to test for this hypothesis, I include an additional control in the main outcome equation: a dummy variable indicating working status. While the effect of working on math scores seems to be little (TT increases from -0.68 SD to -0.69 SD), Table 4 shows that, as expected, the TT effect for picture scores increases greatly: from -0.67 SD to -0.93 SD. This yields support to the hypothesis that child learning occurs in different environments (Hull & Schultz, 2001; Saxe, 1988). Working, however, is not an alternative to formal schooling, as it has little impact on improving math skills.

6.4. Interrupting schooling temporarily

Previous analyses of the effect of the Indonesian crisis on enrollment rates do not distinguish between students who drop out permanently vs. those who drop out temporarily. If there are sensitive periods for investment in children as suggested by previous literature (Almond & Currie, 2011a; Cunha et al., 2010; Currie & Hyson, 1999), children may find it harder to catch up in school after a schooling shock, even if that shock was temporary. The cost of their interruption may therefore be large. To test this hypothesis, I study the effect of temporary school interruptions on math scores. While the IFLS survey does not explicitly ask about temporary school interruption, I construct this variable based on years of schooling attained between the two survey waves in 1997 and 2000 for all children who attend school in both survey years. If in 2000 the child lags behind in predicted years of schooling based on school attainment by 1997, I assume the child has had a school interruption. This school interruption could be due to repetitions of a school year (e.g., 1997/1998) or a failure to enroll in the next school year (e.g., in 1998/1999). Using this measure, I find that 655 of school-age children are lagging behind in school attainment. This is 21% of the total sample of 3020 children, including the 2020 children with continuous enrollment and the 347 children who dropped out permanently. One caveat of using this definition is that it may suffer from measurement error as years of schooling could be misreported. A measurement error in this binary variable would not only bias the OLS estimates downwards but would also yield the IV estimator invalid because the outcome and measurement error will be negatively correlated (i.e., non-classical measurement error). The survey question, however, minimizes the measurement error by asking which grade the child completed in each schooling level (as opposed to simply asking for the total years of education). The additional sample restrictions, described earlier, are aimed at reducing this measurement error, as well.

Table 5 presents a summary of the results. An OLS regression model suggests that temporary shocks to schooling reduce math scores by 0.09 SD. Controlling for unobserved ability in the value-added model yields very similar results. Das et al. (2007) show that an increase of one day per month in teacher absences in Zambia is associated with a decrease in math scores of 0.017SD. This would suggest that my results are similar to an increase of about 6 days per month in teacher absences. Since the children in my sample likely experienced a more than marginal increase in absences (for them to repeat the grade or not enroll), this effect is relatively small. Accounting for the endogeneity of the schooling decision, using the same cost-of-schooling instruments as for permanent dropout, I find that the IV estimate and the estimate from an endogenous switching regression model are very similar. They suggest that the average treatment effect is about ten times larger than the simple OLS result. Again, since temporary schooling interruptions are not well researched, and since the length of school interruption in my data is unknown (children may lag behind in years of schooling because they repeated the grade due to some absences during the year, or because they failed to enroll and skipped one whole year of classes), it is difficult to compare this estimate to previous work. The large size of the effect, however, would be plausible under an age-varying production function, where children cannot catch up even after a temporary shock to their cognitive development. The fact that the ATE of temporary school interruption is larger than that of permanent school interruption in my data could be consistent with that explanation since children who interrupt temporarily are on average 3.6 years younger than those who drop out.

7. Conclusion

The Asian financial crisis of late 1990s had a major impact on many aspects of Indonesian life, interrupting and reversing the progress Indonesia had made over previous decades in economic and social development. One of the potentially long-lasting impacts of the crisis was the shock to children's human capital accumulation, as parents and the government failed to fully insure children against the resulting shocks to household income. Previous studies have documented the effect of the crisis on enrollment rates and years of schooling. Studying child cognitive development, this paper adds to the literature in a fundamental way that has not been previously explored. I examine the impact of school dropout on child cognitive outcomes in the context of an economic crisis as well as more generally in the context of low income.

I use several different approaches to estimate the effect of school dropout, all of which yield consistent estimates of the large effect of school dropout on cognitive outcomes. Under the assumption of no endogeneity of the dropout decision, the ordinary least squares (OLS) regression models show a significant decrease in both math and picture scores. Accounting for endogeneity due to omitted ability bias, I then use a value-added model, which yields effects which are smaller but similar to the previous

OLS estimates. Next, I use age, per capita expenditures, nonzero registration fees, and school distance as the exogenous schooling cost shifters and quantify the effect of school interruption in a more general instrumental variable model. The IV estimates are large, suggesting that the endogeneity of the decision to drop out may not only be due to unobserved ability but also to unobserved resource constraints. Finally, I estimate an endogenous switching regression model to account for heterogeneity in unobservables. I find that the effect of treatment on the treated for permanent school dropout is large and significant at about 0.5 standard deviations decrease in both math and picture scores. The results suggest that children who dropped out during the crisis would have done well at school had they stayed. In light of previous findings on the high returns to years of schooling as well as cognitive outcomes in Indonesia, the findings imply that children who dropped out of school due to resource constraints rather than low ability may have lower expected earnings as adults compared to what they could have earned.

Bibliography

Almond, D., & Currie, J. (2011a). Chapter 15 – Human Capital Development before Age Five. *Handbook of Labor economics, Vol. 4B* (Vol. 4, pp. 1315–1486). Elsevier Inc. doi:10.1016/S0169-7218(11)02413-0

Almond, D., & Currie, J. (2011b). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–172. doi:10.1257/jep.25.3.153

Aughinbaugh, A., & Gittleman, M. (2003). Does Money Matter ? A Comparison of the Effect of Income on Child Development in the United States and Great Britain. *Journal of Human Resources*, *38*(2), 416–440.

Averett, S. L., & Burton, M. L. (1996). College Attendance and the College Wage Premium : Differences by Gender. *Economics of Education Review*, *15*(1), 37–49. doi:10.1016/0272-7757(95)00027-5

Borjas, G. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, *77*(4), 531–553.

Cameron, S. V., & Heckman, J. J. (2001). The Dynamics of Educational Attainment for Black, Hispanic, and White Males. *Journal of Political Economy*, *109*(3), 455–499. doi:10.1086/321014

Cas, A. G. (2011). Long-run Cognitive and Education Impacts of Early Life Public Health Intervention : Evidence from Safe Motherhood Program in Indonesia.

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., Heckman, J., & Schennach, S. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica : Journal of the Econometric Society*, *78*(3), 883–931. doi:10.3982/ECTA6551

Currie, J., & Hyson, R. (1999). Is the Impact of Health Shocks Cushioned by Socioeconomic Status ? The Case of Low Birthweight. *The American Economic Review*, *89*(2), 245–250.

Das, J., Dercon, S., Habyarimana, J., & Krishnan, P. (2007). Teacher Shocks and Student Learning : Evidence from Zambia. *Journal of Human Resources*, *42*(4), 820–862.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia : Evidence from an Unusual Policy Experiment. *The American Economic Review*, *91*(4), 795–813.

Duncan, G. J., Yeung, W. J., Brooks-gunn, J., & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review*, *63*(3), 406–423.

Ferreira, F. H. G., & Schady, N. (2009). Aggregate Economic Shocks, Child Schooling, and Child Health. *The World Bank Research Observer*, *24*(2), 147–181. doi:10.1093/wbro/lkp006

Frankenberg, E., Smith, J. P., & Thomas, D. (2003). Economic Shocks , Wealth , and Welfare. *The Journal of Human Resources*, *38*(2).

Ghuman, S., Behrman, J. R., Borja, J. B., Gultiano, S., & King, E. M. (2005). Family Background , Service Providers , and Early Childhood Development in the Philippines: Proxies and Interactions. *Economic*, *54*(1), 129–164.

Glewwe, P. (2002). Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes. *Journal of Economic literature*, *40*(2), 436–482.

Gunnarsson, V., Orazem, P. F., Sanchez, M., & Sa, M. A. (2006). Child Labor and School Achievement in Latin America. *The World Bank Economic Review*, *20*(1), 31–54. doi:10.1093/wber/lhj003

Guo, G., & Harris, K. M. (2000). The mechanism mediating the effects of poverty on children's intellectual development. *Demography*, *37*(4), 431–447.

Hanushek, E. a, & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, *46*(3), 607–668. doi:10.1257/jel.46.3.607

Heckman, J. J., & Rubinstein, Y. (2001). The Importance of Noncognitive Skills : Lessons from the GED Testing Program. *American Economic Review*, *91*(2), 145–149.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of labor economics*, *24*(3), 411–482.

Hull, G., & Schultz, K. (2001). Literacy and Learning Out of School : A Review of Theory and Research. *Review of Educational Research*, *71*(4), 575–611.

Jacoby, H. G., & Skoufias, E. (1997). Risk, Financial markets, and Human capital in a Developing country. *Review of Economic Studies*, *64*(3), 311–335.

Jones, G. W., & Hagul, P. (2001). Schooling in Indonesia: Crisis-Related and Longer-Term Issues. *Bulletin of Indonesian Economic Studies*, *37*(2), 207–231. doi:10.1080/00074910152390892

Lanjouw, P., Pradhan, M., Saadah, F., Sayed, H., & Sparrow, R. (2001). Poverty , Education and Health in Indonesia : Who Benefits from Public Spending ?

Li, M., Poirier, D. J., & Tobias, J. L. (2004). Do dropouts suffer from dropping out? Estimation and prediction of outcome gains in generalized selection models. *Journal of Applied Econometrics*, *19*, 203–225.

Malamud, O., & Pop-Eleches, C. (2011). Home Computer Use and the Development of Human Capital. *The Quarterly Journal of Economics*, *126*(2), 987–1027. doi:10.1093/qje/qjr008

Paxson, C., & Schady, N. (2007). Cognitive Development among Young Children in Ecuador The Roles of Wealth, Health, and Parenting. *Journal of Human Resources*, *42*(1), 49–84.

Rosenzweig, M. R., & Schultz, T. P. (1983). Estimating a Household Production Function : Heterogeneity , the Demand for Health Inputs , and Their Effects on Birth Weight. *Journal of Political Economy*, *91*(5), 723–746.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Saxe, B. (1988). Candy Selling and Math Learning. *Educational Researcher*, *17*(6), 14–21.

Suryadarma, D. (2010). The Merits of Ability in Developing and Developed Countries.

Suryadarma, D., Suryahadi, A., Sumarto, S., & Rogers, F. H. (2006). Improving Student Performance in Public Primary Schools in Developing Countries: Evidence from Indonesia. *Education Economics*, *14*(4), 401–429. doi:10.1080/09645290600854110

Thomas, D., Beegle, K., Frankenberg, E., Sikoki, B., Strauss, J., & Teruel, G. (2004). Education in a crisis. *Journal of Development Economics*, *74*(1), 53–85. doi:10.1016/j.jdeveco.2003.12.004

Todd, P. E., & Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children : Home , School , and Racial Test Score Gaps. *Journal of Human Capital*, *1*(1), 91–136.

Willis, R. J., & Rosen, S. (1979). Education and Self-Selection Sherwin Rosen. *Journal of Political Economy*, *87*(5), S7–S36.

Table 1: Characteristics of children who drop out vs. the comparison group

Variable	Comparison, mean (sd)	Dropout, mean (sd)	P-value
Age in 1997	10.59	13.50	<0.001
	(2.41)	(1.97)	
Female	0.51	0.53	0.507
	(0.5)	(0.5)	
Years of education in 1997	4.66	7.01	<0.001
	(2.35)	(2.11)	
Years of education in 2000	7.39	8.16	<0.001
	(2.2)	(2.26)	
Math score in 1997	0.22	-0.15	<0.001
	(0.94)	(0.79)	
Math score in 2000	0.25	-0.23	<0.001
	(0.95)	(0.93)	
Picture score in 2000	0.17	-0.13	<0.001
	(0.9)	(1.02)	
Urban area of residence	0.47	0.43	0.126
	(0.5)	(0.5)	
Real per capita expenditure in	287.61	213.37	0.003
1997, '000 Rps	(466.5)	(178.75)	
Real per capita expenditure in	271.20	246.70	0.144
2000, '000 Rps	(296.91)	(233.62)	
Has an older sibling living in the	0.54	0.26	<0.001
household	(0.5)	(0.44)	
Number of Children	2020	347	

Notes:
[1] P-values based on a t-test.

Table 2: The effect of school dropout on cognitive outcomes

Variable	Math score OLS	Math score OLS-VA	Math score 2nd stage	Picture score OLS	Picture score OLS-VA	Picture score 2nd stage
	(1)	(2)	(3)	(4)	(5)	(6)
Dropout	-0.4358**	-0.3732**	-0.8145**	-0.2577**	-0.2214**	-0.7534**
	(0.0527)	(0.0543)	(0.0975)	(0.0592)	(0.0631)	(0.1045)
Math score in 1997		0.1671**			0.1428**	
		(0.0243)			(0.0232)	
Years of education in 2000	0.0679**	0.0604**	0.0766**	0.0612**	0.0579**	0.0722**
	(0.0088)	(0.0093)	(0.0091)	(0.0084)	(0.0090)	(0.0088)
Female	0.0737**	0.0785**	0.0735**	-0.1688**	-0.1628**	-0.1682**
	(0.0363)	(0.0380)	(0.0362)	(0.0350)	(0.0369)	(0.0357)
Mother years of education	0.0355**	0.0315**	0.0291**	0.0265**	0.0226**	0.0180**
	(0.0055)	(0.0059)	(0.0057)	(0.0053)	(0.0057)	(0.0056)
Has older sibling who lives	0.0616*	0.0314	0.0163	0.1271**	0.1229**	0.0672*
in household	(0.0365)	(0.0387)	(0.0375)	(0.0364)	(0.0388)	(0.0388)
Log of per capita	0.0828**	0.0641**	0.0768**	0.0652*	0.0349	0.0586*
expenditure in 2000	(0.0311)	(0.0310)	(0.0309)	(0.0337)	(0.0345)	(0.0341)
Urban area of residence	0.1847**	0.1400**	0.1849**	0.1843**	0.1097**	0.1863**
	(0.0429)	(0.0456)	(0.0428)	(0.0412)	(0.0439)	(0.0418)
Month of interview fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.8460**	-1.4836**	-1.7475**	-1.7282**	-1.3306**	-1.6226**
	(0.3790)	(0.3832)	(0.3762)	(0.4072)	(0.4225)	(0.4116)
Observations	2367	2095	2367	2367	2095	2367
Number of clusters	1688	1688	1688	1688	1688	1688

Notes:

[1] Standard errors in parentheses, clustered at the household level.

[2] * significant at the 10% level, ** significant at the 5% level

Table 3 - Switching regression for school dropout

	Permantent dropout			
	Math scores		Picture scores	
Variable	Regime 1	Regime 0	Regime 1	Regime 0
Years of education in 2000	0.145**	0.064**	0.124**	0.066**
	(0.025)	(0.010)	(0.025)	(0.010)
Female	-0.039	0.097**	-0.367**	-0.139**
	(0.093)	(0.039)	(0.102)	(0.037)
Urban area of residence	0.264**	0.177**	0.210	0.176**
	(0.101)	(0.047)	(0.113)	(0.044)
Log of per capita expenditure in 2000	0.036	0.082**	0.062	0.050
	(0.071)	(0.034)	(0.078)	(0.036)
Mother years of education	0.021	0.031**	0.003	0.022**
	(0.016)	(0.006)	(0.018)	(0.006)
Has older sibling who lives in household	0.192	0.021	-0.114	0.117**
	(0.105)	(0.040)	(0.129)	(0.041)
Month of interview fixed effects	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Constant	-2.180**	-1.774**	-2.261**	-1.523**
	(0.867)	(0.417)	(0.975)	(0.435)
ρ_1	-0.012		0.144	
	(0.089)		(0.091)	
ρ_0	0.285**		0.507**	
	(0.109)		(0.169)	

Notes:
[1] Standard errors in parentheses, clustered at the household level.
[2] * significant at the 10% level, ** significant at the 5% level

Table 4: Treatment effects of school dropout

Variable	ATE	TT	TUT
Permanent school dropout			
Math scores	-0.52**	-0.68**	-0.49**
	(0.101)	(0.108)	(0.117)
Picture scores	-0.56**	-0.67**	-0.54**
	(0.119)	(0.126)	(0.134)
Math scores controlling for work	-0.45**	-0.69**	-0.41**
	(0.145)	(0.159)	(0.167)
Picture scores controlling for work	-0.65**	-0.93**	-0.60**
	(0.167)	(0.211)	(0.192)

Notes:

[1] Average Treatment Effect (ATE), Treatment on the Treated (TT), and Treatement on the Untreated (TUT) are calculated based on selection-corrected coefficients from a regime-switching regression.

[2] Standard errors are calculated using the bootstrap method with 100 replications.

[3] ** denotes significance at the 5% confidence level, * denotes significance at the 10% level.

Table 5: Effects of temporary school interruption on math scores

Model	Effect
OLS	-0.0914**
	(0.0429)
OLS-VA	-0.0806*
	(0.0455)
IV	-1.1055**
	(0.2526)
ATE	-1.06**
	(0.167)

Notes:
[1] Average treatment effect (ATE) calculated based on selection-corrected coefficients from a regime-switching regression. Standard errors are calculated using the bootstrap method with 100 replications.
[2] ** denotes significance at the 5% confidence level,
* denotes significance at the 10% level.