UNITED NATIONS
UNIVERSITY
UNU-WIDER

This paper is a draft submission to the

WIDER Development Conference

Human capital and growth

6-7 June 2016 Helsinki, Finland

Nature or Nurture: Evidence from Indonesia

Cara Ebert, University of Göttingen
Erik Plug, University of Amsterdam

January 28, 2016

Abstract

This paper examines the intergenerational persistence of education in Indonesia. Specifically, it explores the family environmental factors that affect children's education. First, I use data from the Indonesian Family Life Survey, distinguishing between adopted and biological children, to disentangle nature from nurture components in the intergenerational schooling persistence. Second, I follow an instrumental variables strategy using data from a school construction program to estimate the causal effect of parental education on child education. I find no evidence that nature plays an important role in explaining the intergenerational transmission of schooling outcomes. Additionally, parental education contributes a major part to the nurture component. This is an encouraging result for policy makers. It shows that schooling is effective in improving nurturing skills and that intergenerational returns to schooling exist. Thus, there exist long-term effects of increasing schooling today on human capital accumulation. Because my results differ from evidence on developed countries, I hypothesize that the effect of an additional year of parental education has larger effects on nurture in Indonesia - where education levels and child rearing standards are low - compared to more advanced countries.

Keywords: Education, Development, Human Capital, Human Development

JEL Classification: I24, I25, O15

1 Introduction

More educated parents have more educated children. The higher the intergenerational persistence in education is, the lower the equality of opportunity for new generations. Understanding the intergenerational persistence in education goes back to an old question: is it nature, nurture or a combination of both that drives persistence? Nature refers to the transmission of education that is driven by genetic endowments - simply put, more able parents bear more able children. Nurture by contrast, refers to the transmission of education that is driven by family environmental factors. Parental education affects parent's child rearing skills, family income, time spent with a child, their attitude towards health and education, among many other facets that shape the home environment. Because the term nurture defines a broad spectrum of family environmental factors, it is worthwhile to break it down into more explicit components. The causal effect of parental education on the child's education can be thought of as such a component. More specifically, the causal effect represents the pure contribution of parental education to the nurture factor - it is the change in child schooling caused by an increase in parental education while the ability distribution and all other family characteristics stay constant.

Disentangling the mechanisms at work in the intergenerational schooling persistence has important policy implications. If nurture determines the intergenerational relationship, then increasing educational attainment of one generation will have social returns in the sense that it increases the educational attainment of future generations. However, if persistence is predetermined by nature, policies should focus on improving the schooling environment for disadvantaged children despite the absence of intergenerational returns.

A number of papers attempt to estimate the size of nature and nurture components or the causal effect of parents' education on children's education. Three types of identification strategies have evolved in the literature. The first strategy uses parent twin pairs to difference out innate characteristics. The second strategy is to estimate the persistence for adopted children who do not carry the genes of their adoptive parents. Both, the twin and the adoption strategy eliminate nature components and thus identify the nurture factor in the intergenerational schooling persistence. The third strategy relies on instrumental variables (IV) to estimate the causal effect of parents' schooling where educational reforms are the most commonly used instrument for parental education.

To the best of my knowledge, there exists only one study that applies one of the three introduced identification strategies to a developing country.[1] The novelty of this paper is that it fills this gap using data from Indonesia. Based on longitudinal data from the Indonesian Family Life Survey (IFLS) I apply two of the three estimation techniques. First, I identify the role of nurture (and nature) using the adoption approach. Second, I

[1] De Walque (2005) applies the adoption strategy to Rwandan data.

measure the causal effect of parental education on the child's education using instrumental variables. To instrument parents' education I closely follow Duflo (2001) and make use of a large scale primary school construction program, the Sekolah Dasar INPRES Program, which took place in Indonesia between 1973-1974 and 1978-1979. Combining the two estimation strategies gives me a deeper insight into the nurture production function in Indonesia and lends credence to my results. For both the adoption and IV strategy I apply two measures of educational attainment. One is years of education and the other one is proceeding to secondary school. In anticipation of possible imprecise IV estimates, I pool mother's and father's education together and thus estimate the increase in either parent's education on child education. This approach is very similar to the one taken by Oreopoulos et al. (2006). If the effects of fathers and mothers are identical, this approach has the advantage that it controls for assortative mating, avoids multicollinearity, and delivers more precise estimates.

The matter in question is how family environmental factors in developing countries differ from developed countries in their effect on the child's educational attainment. Developing countries are characterized by relatively low education levels and child-rearing standards. Two paths of thought evolve from this observation. On the one hand, nature may be particularly important to overcome disadvantages. Therefore, only the more able children of more able parents possess the faculty to cope with unfavorable environments and reach higher education levels. In this case, one would expect nurture to play an inferior role in developing countries. And therefore, the causal impact of parents' education is also low compared to estimations of the causal effect for developed countries. On the contrary, at lower education levels another year of schooling is potentially very effective in terms of gaining nurturing skills. This would imply decreasing marginal returns to a country's average level of education. In this case, the causal effect of parents' education on child education would be larger in a developing country compared to developed countries where education levels are higher on average.

In the existing literature for the developed world the relevance of nurture factors is low to non-existent. For example Behrman & Rosenzweig (2002) use pairs of identical twin mothers in Minnesota to difference out inherent mother characteristics and address assortative mating induced biases by controlling for fathers' education and earnings. Despite positive conventional intergenerational schooling correlations, they find no effect of mothers' schooling on child schooling but a positive effect of fathers' schooling. In support of Behrman & Rosenzweig (2002), Plug (2004) finds that for adopted children in Wisconsin the positive relationship that existed for biological children disappears for mothers while that for fathers remains greater than 0.2. Sacerdote (2007) uses a sample of Korean-born American adoptees and finds a significantly positive relationship between parental education and the child's schooling outcome. This relationship is however much lower than that for biological children. For most of the estimates from the United States

the simple intergenerational schooling correlation ranges between 0.30 and 0.45.

Black et al. (2005) apply an instrumental variables methodology to identify the causal effect of parental education on the child's educational attainment. To instrument parental education they use an increase in compulsory schooling from 7 to 9 years in Norway. They find no causal effect of parental education except for that of mothers on their sons. Similarly, Chevalier (2004) uses a change in compulsory schooling laws in the United Kingdom as an instrument for parental education and finds no causal effect of fathers' schooling but a positive effect of mothers' schooling on the child's educational attainment. Holmlund et al. (2011) compare all three estimation strategies - twins, adoption and IV - based on a comprehensive Swedish data set and suggest that the estimates differ systematically across identification strategies. For Scandinavian countries they conclude an intergenerational association of about 0.2 to 0.3 and a causal effect estimate of about 0.1 which explains a large part of the nurture component.

In summary, the majority of the literature is based on data from the United States or Scandinavia and they all find that the nurture component in the intergenerational persistence is fairly small.

The developing country literature is limited and predominantly focused on simple intergenerational schooling correlations. Behrman et al. (2001) find that the intergenerational schooling persistence between the most educated parent and her child in Mexico, Peru, Colombia and Brazil is between 0.5 and 0.7.[2] Emran & Sun (2011) measure the intergenerational transmission of education and occupation for rural China and compare estimated coefficients for 1988 and 2002. To estimate the causal effect of parents schooling on the childs schooling, the authors follow Klein & Vella (2009). Their method exploits heteroskedasticity in parents education to generate an instrument for parents education. They find that the schooling mobility remained static for daughters and decreased for sons. However, the size of the coefficients varies greatly between specifications such that it is difficult to draw conclusions from their analysis other than regarding time trends. To my knowledge, De Walque (2005) is the only study that applies one of the three proposed identification strategies to a developing country. He applies the adoption strategy to Rwandan data where the definition of adopted children is used in a broader sense as children living absent from their parents. He finds estimates for biological children to range between 0.14 and 0.31, and estimates for adopted children to be lower, ranging between 0.08 and 0.25 depending on the specification. However, when interacting the 'adoptive' parents' education with indicators for whether the 'adoptive' parent is a relative, only coefficients for children living with relatives stay large and significant. This questions the importance of nurture in the educational persistence in Rwanda.

The results of this paper indicate great importance of nurture in the intergenerational

[2]These estimations for Latin America do not take into account assortative mating effects.

transmission of education in Indonesia. I find that the intergenerational schooling persistence is between 0.23 and 0.29 for years of education and between 0.10 and 0.27 for proceeding to secondary school. For neither variable nor identification strategy I find support that nature matters. Additionally, I find that parental education contributes a major part to the nurture factor. Therefore, I hypothesize that an additional year of parental education has larger effects on child education in Indonesia where education levels and child rearing standards are low compared to more advanced countries. Nevertheless, my analysis is prone to violations of identifying assumptions wherefore no definite conclusions can be drawn.

The thesis is structured as follows. Section 2 introduces the methodology. Section 3 describes the data. Section 4.1 presents the results of the adoption strategy and section 4.2 presents the results of the IV estimation. Section 5 concludes.

2 Methodology

This section introduces two methodological approaches to investigate the nurture production function. The first approach separates nurture's role in the intergenerational schooling persistence from that of nature by making use of an adoption strategy. The second approach estimates the causal effect of parental education on the child's education using instrumental variables.

For both of the proposed methodologies I pool fathers' and mothers' educational attainment together. This way I am able to control for assortative mating, avoid multicollineraity, but most importantly increase the precision of my IV estimates. When doing so, I assume that the father's and mother's influence on the child's schooling is of equal size. This is an arguable assumption as for instance Black et al. (2005) and Chevalier (2004) have found that the influence of mothers' schooling on child schooling is greater than that of fathers. In contrast, Behrman & Rosenzweig (2002) and Plug (2004) have found no effects of mothers' schooling on child schooling but positive effects for fathers. However, as suggested by Oreopoulos et al. (2006) pooling parental education together delivers greater precision in the IV estimates. Oreopoulos et al. (2006) argues that the instruments for fathers' and mothers' schooling are too highly correlated. Including their education separately would result in insignificant coefficients. Furthermore, if policy makers introduce schooling reforms that affect both genders the effect of fathers' and mothers' schooling together is of interest for the child's outcome.

The first methodological approach is a comparison of OLS estimates between a sample of biological children and adopted children. The idea is that adopted children do not share the same genetics with their adoptive parents but they are assumed to grow up in similar environments as biological children. Two underlying assumptions follow from this: (1) the adoption process is random such that innate characteristics of adopted

children and their adoptive parents are uncorrelated, and (2) adoptive parents are similar to biological parents. Then adoption can be considered an experimental intervention that randomly assigns children to families. Hence, the difference between the estimates for biological and adopted children can be attributed to intergenerational persistence caused by common genes (nature). I will separately estimate the following regression model for biological and adopted children:

$$S_i^c = \beta_0 + (S_i^f + S_i^m)\beta_1 + X_i\beta_2 + \epsilon_i. \tag{1}$$

S_i^c is the educational attainment of child i and similarly S_i^f and S_i^m describe the educational attainment of child i's father and mother respectively. X_i is a vector of the child's and its parents' background characteristics and ϵ_i is a child specific error term. In the estimation for adopted children β_1 can be interpreted as the nurture component in the intergenerational schooling persistence. I discuss possible violations to the identifying assumptions in detail in the result section.

The second approach identifies the causal effect of parents' educational attainment on that of their children by creating an exogenous variation in parents' schooling through instrumentation. Same as Duflo (2001) I will make use of a large scale primary school construction program, the Sekolah Dasar INPRES Program, that took place in Indonesia between 1973-1974 and 1978-1979. During this time the schooling infrastructure was massively extended through the construction of primary schools in disadvantaged areas. The validity of this estimation strategy relies on two assumptions. First, the instruments must have a jointly significant effect on parental education (relevance), and second, the instruments must have no effect on child schooling other than the effect through parental schooling, or said differently, the instruments need to be uncorrelated with the error term (exogeneity). The instruments used here are a series of interaction terms of parent's year of birth indicators and the program intensity in the parent's region of birth. The exposure of a parent to the program is determined through her region of birth and year of birth. Individuals that are older than primary school age at the time of program implementation (12 or older) as well as individuals from no program areas should not have been affected by the program. For children younger than 12 in program areas the effect depends on the age of the child, meaning it depends on how many more years it will still attend primary school, and the program intensity in the region of birth.[3] Therefore, the instruments create variation through timing and location. The instruments' relevance and enxogeneity are discussed in detail in the results section.

The two-stage least squares (2SLS) model follows closely Duflo (2001). The first stage

[3] Duflo (2001) finds no program effects for old cohorts and increasing program effects for younger cohorts.

takes the following form:

$$S^f_{ikm} + S^m_{iln} = \alpha_0 + \sum_{k=1}^{K}(Birth^f_{ik} * Intensity_m)\alpha_{1k} + \sum_{l=1}^{L}(Birth^m_{il} * Intensity_n)\alpha_{1l} + X_i\delta + u_k + p_l + v_m + s_n + \epsilon_{ikm} + \varepsilon_{iln}. \quad (2)$$

$S^f_{ikm} + S^m_{iln}$ stands for child i's parents' educational attainment, that is the sum of child i's father's educational attainment born in year k and region m and child i's mother's educational attainment born in year l in region n. The first sum on the right hand side represents a series of interaction terms of child i's father's birth of year dummies and the program intensity in the father's region of birth. The second sum on the right hand side represents the equivalent for mothers. Further, I control for region and year of birth fixed effects for fathers and mothers. X_i is a vector of the child's and its parents' background characteristics and ϵ_{ikm} and ε_{iln} present father and mother specific error terms. The fact that program regions differ from non-program regions is taken account of by the region of birth fixed effects. Subsequently, the second stage takes the following structural form:

$$S^C_{iklmn} = \beta_0 + (S^f_{ikm} + S^m_{iln})\beta_1 + X_i\gamma + \mu_k + \pi_l + \omega_m + \rho_n + \vartheta_{iklmn} \quad (3)$$

where S^C_{iklmn} is the educational attainment of child i with her father born in year k and region m and her mother born in year l and region n. Again it is controlled for background characteristics and each parent's year and region of birth fixed effects. ϑ_{iklmn} illustrates a child specific error term. β_1 can be interpreted as the causal effect of an additional year of either parent's schooling on the child's years of schooling (for S measured in years of schooling). This causal effect can be thought of as the extent to which parental education contributes to nurture factors explaining child education.

3 Data

3.1 Adoption practices in Indonesia

Indonesia has one of the lowest adoption rates in the world. In 1999 there was only one adoption per 100,000 births and only 0.1 adoptions under age five per 100,000 children under age five. The total number of domestic adoptions in 1999 was 33.[4] Compared to other Asian and development countries Indonesia has relatively low levels of desired fertility with about three or less children per family. Information about domestic adoption practices and parents motives to adopt are sparse. However, based on requirements of adoption regulations from 1976 and 1982 indicative information is available. Adoption is

[4]See. http://www.un.org/esa/population/publications/adoption2010/child_adoption.pdf, p.68.

said to be only implemented in the best interest of the child. The examination of motives of adoptive parents, including their sincerity, seriousness and awareness of consequences of adoption is a substantial element of the adoption process. This has to be expressed in a letter of motivation and in a statement that the adopted child will be registered to have an education and health insurance. Additionally, references from close relatives, a letter of income statement and a letter of good conduct from the police headquarter are required. Prospective parents must be in the age range of 30 to 55 years and the couple must have been married for at least 5 years. The couple must be either childless or have no more than one adopted or biological child. Prospective parents are to believe in god and have to have the same religion as the child. Furthermore, adoptive parents have to pass medical check-ups at an Indonesian government hospital and a psychological study. Also the involuntary infertility of the couple needs to be proven by a gynecologist. After a child is released to prospective parents there is a fostering period of six month in which a social worker will do a home visit and an interview.[5]

These adoption standards and requirements ensure that adoptions are undertaken out of good intentions and that adoptive parents come from backgrounds that make it possible to provide for the child. However, it also indicates that adoptive families may differ from biological families with regard to family size. In the samples I use in this analysis, the average family size is very similar for both groups, with 2.6 children in biological families and 2.8 children in adoptive families.[6] Unfortunately, no information about the matching process of prospective parents and children are available to me. Though, the randomness of the adoption process may be limited for a few reasons. In the adoption process parents and their children are sorted by religion. Furthermore, the adoption is initiated through orphanages such that prospective parents can "go and look" at the children and may find out about the child's background. 56% of children in orphanages have both biological parents still alive but their parents were unable to take care of them. In these cases biological parents have to provide consent to the adoption. This carries the risk of non-exogenous adoption placement. However, in any case it can be assumed that almost all children at an orphanage are disadvantaged and, additionally, young children in urgent situations are given adoption preference.[7] Hence, when all adopted children come from disadvantaged backgrounds, in some way, the adoption process can be viewed random again.

[5] See http://www.sayapibujakarta.org/eng/adoption.html and http://www.expat.or.id/info/adoptingchildreninindonesia.html.

[6] The number of children in adoptive families being greater than two indicates that the requirement of adoptive parents having no more than one child may be loosely implemented. However, the majority of adoptive families has two or less children and the rest may be attributable to differing rules for step child adoptions. Therefore, family size does not constitute a great concern for this analysis.

[7] See http://adoption.state.gov/country_information/country_specific_info.php?country-select=indonesia.

3.2 The Sekolah Dasar INPRES Program

In consequence of high revenues from an oil boom the Indonesian government decided to allocate the acquired funds to a massive primary school construction program, the Sekolah Dasar INPRES Program. Schools were planned to be constructed disproportional to a regions primary school enrollment rate and, hence, targeted at disadvantaged areas with the aim of equity across regions. At the same time teacher training and recruitment were extended such that the occurring teaching demand could be fulfilled. Overall 61,807 primary schools were built between 1973-1974 and 1978-1979. Thereby, more than one school was built per 500 students aged 5 to 14 at that time. The program constituted a drastic change in Indonesian education policy since stagnating education expenditure preceded the Sekolah Dasar INPRES Program.[8]

3.3 The Indonesian Family Life Survey

The analysis is based on the RAND Indonesian Family Life Survey (IFLS), which is an ongoing panel survey including individual, household and community level data. More specifically, the data is taken from the first (IFLS1 in 1993) and most recent (IFLS4 in 2007) survey wave. I linked parents and children from two parent families living in one household in 1993. In 1993 I observed household and parent characteristics, whereas I observed child education characteristics 14 years later in 2007. However, in 2007 children were not required to live with their parents anymore.

Because some children were just born in 1993 they potentially have not finished school in 2007. To solve issues arising from censored data I eliminate all children that were of age 22 or younger in 2007.[9] At the age of 23 each child had the chance to at least have finished an undergraduate university degree or even a graduate degree.[10]

For the adoption sample, I did not impose any restrictions on parents in order to reach the maximum amount of observations possible. Therefore, the father's year of birth ranges between 1922 and 1974, and the mother's year of birth ranges between 1936 and 1977. For the IV estimation, the sample size decreases compared to the adoption estimation sample because not for every observation there is region of birth information available. Further, I removed parents born before 1945 to avoid having children in the sample that were of age six before the primary school construction program was fully implemented.[11] Therefore, in the IV sample, father's year of birth ranges from 1945 to

[8] See Duflo (2001) for more information.

[9] This restriction only applies when educational attainment is measured in years of schooling. For the measure of proceeding to secondary school the restriction is redundant because at age 14 every child had the chance to proceed to secondary school.

[10] Black et al. (2005) use children's year of birth indicators to control for censoring. To check the robustness of my estimates I also estimate the model with child year of birth dummies.

[11] With this restriction the oldest child in the sample is born in 1975.

1974 and mother's year of birth ranges from 1945 to 1977.

Each household member was asked to describe the relationship to the household head. This variable distinguishes between biological and adopted children where adoption includes cases of step child adoption.

Information whether a certain school was part of the INPRES program was taken from the second IFLS survey (IFLS2 in 1997). The program intensity in a region was calculated based on the share of schools that were part of the INPRES program in one region and this information was then linked to the parents region of birth.

Educational attainment is measured in two variables: years of education and whether the individual proceeded to secondary school. The variables were derived from information about an individual's highest level of education (indicated as primary education, junior secondary and so on) and information about the completion or years of education at that level. For children the proceeding-to-secondary-school variable is an indicator whereas for parents the variable can take three values: zero if no parent, one if either parent and two if both parents proceeded to secondary school.

Table 1 presents the descriptive statistics. Columns (1) and (2) show background characteristics for biological and adopted children older than 22. On average the two samples are very similar. Adopted children are about one year older than biological children whereas their fathers are about three years younger and their mothers are about two and a half years younger than those of biological children. This is surprising since adoptive parents have to go through the process of realizing to be infertile and are therefore usually older than biological parents. The estimations in section 4 control for the child's and parents' age. Overall the level of education is lower in adoptive families. In view of the fact that in Western societies adoptive parents are higher educated and wealthier on average this observation is somewhat surprising. Also De Walque (2005) reports in his study from Rwanda that adoptive parents have more years of education than biological parents. In order to make the sample of biological and adopted children more comparable in terms of parents' education, I create samples with parents having no more than 12 years of education. Thereby, the restricted sample of biological children becomes more comparable to the full sample of adopted children with respect to education.

Columns (3) and (4) of table 1 show background characteristics of the IV estimation sample separated in non-reform prone parents - which are those parents that were either too old to benefit from the reform or that were born in non-reform areas - and reform prone parents - which are parents that were of reform-according age and were born in reform areas.[12] The samples of non-reform prone and reform prone parents differ in some respect. First children with reform prone parents are more likely to grow up in rural areas. This makes sense because rural areas are typically worse off in terms of primary

[12]The sample is restricted to children being older than 22.

Table 1: Descriptive statistics for adopted and biological children as well as for non-reform and reform prone parents (with all children being older than 22 in 2007).

	Adoption sample		IV sample	
	(1) Biological	(2) Adopted	(3) Non-reform prone	(4) Reform prone
Child characteristics				
Rural area	0.40	0.41	0.50	0.58
Gender (Male=1)	0.51	0.52	0.47	0.52
Age in 2007	29.64	30.42	26.14	24.93
Years of schooling	11.33	9.63	11.35	10.01
Proceeded to secondary school	0.85	0.69	0.85	0.78
Father characteristics				
Age in 2007	60.51	57.43	54.98	49.34
Years of schooling	7.51	6.57	7.34	5.72
Proceeded to secondary school	0.44	0.26	0.42	0.26
Mother characteristics				
Age in 2007	54.93	52.56	51.05	43.07
Years of schooling	6.10	5.82	6.11	4.64
Proceeded to secondary school	0.30	0.28	0.30	0.12
Sum of father and mother characteristics				
Years of schooling	13.61	12.40	13.45	10.36
Proceeded to secondary school	0.75	0.54	0.72	0.38
Observations	4594	96	1096	340

school enrollment rates. Further, it is understandable that non-reform prone parents and their children are older because only younger parents had a chance to benefit from the school construction project. It is surprising that the reform prone sample reveals considerably lower educational attainment levels than the non-reform prone sample. However, with respect to the reform regions being disadvantaged regions in terms of school enrollment rates it seems reasonable that the educational attainment is lower despite the school construction project. When checking the educational attainment of a sample with older parents from program regions, I find that the sample is not as low educated as the younger sample with parents from program regions, which stands in contradiction with my argument. Because also the sample with younger parents in non program regions has lower educational attainment than the sample with older parents in non program regions, I suggest that some exogenous event affected educational attainment in the younger cohort.[13]

The non-reform prone sample is overall very similar to the sample of biological children in terms of educational attainment. To account for the differences in background characteristics between samples, I control in the regression specification for parent and child age as well as child gender and whether the household lives in a rural area.

4 Results

4.1 OLS results: Biological versus adopted children

Table 2 shows OLS results of the intergenerational schooling persistence in Indonesia for biological children that are older than 22. Persistence is measured in years of schooling. Column (1) indicates that an additional year of either parent's schooling increases the child's educational attainment by 0.277 years. The estimate is robust to the inclusion of child cohort dummies in column (2) which allow for an educational attainment trend over time and other cohort specific characteristics. The size of the intergenerational persistence is fairly low compared to other studies on developing countries. Behrman et al. (2001) find that the intergenerational schooling persistence in Mexico, Peru, Colombia and Brazil is between 0.5 and 0.7.[14] Instead, the estimates are comparable to OLS estimates from Scandinavia (see Black et al. (2005), Holmlund et al. (2011)).

The summary statistics revealed that adoptive parents are lower educated than biological parents. To correct for this difference in education columns (3) and (4) present estimates for parents with no more than 12 years of education. The intergenerational persistence for less educated parents is higher than it is for the full sample. This suggests

[13] Censoring cannot explain the lower educational attainment in the program prone sample because all parents have finished schooling in 1993 and only children older than 22 were considered.

[14] However, the estimates by Behrman et al. (2001) do not account for assortative mating effects and are therefore most likely upward biased compared to estimates from this study.

that the return to education is higher among lower educated parents and implies that the marginal return to education falls as education increases. This observation is in line with other studies which find larger persistence for less educated samples (see Oreopoulos et al. (2006)). However, the difference in persistence between the two samples is not statistically significant.

Table 3 shows OLS results of the intergenerational schooling persistence for adopted children that are older than 22. The structure of the table is analog to table 2. An additional year of schooling of either parent increases the adopted child's educational attainment by about 0.23 years. As expected, the persistence for adopted children is lower than for biological children in comparison to both the unrestricted sample in columns (1) and (2) and the more comparable sample in columns (3) and (4) of table 2. This suggests that only a small part of the intergenerational schooling persistence can be explained by the transmission of innate characteristics that support or limit a child's educational success. However, the difference between the biological and adoptive children's coefficients is insignificant.

The fact that the estimates for the restricted sample of adoptees are larger than the estimates for the restricted sample of biological children may seem counterintuitive. However, adoptive parents are less educated on average and in addition the sample in columns (3) and (4) is restricted to parents with no more than 12 years of schooling. This depresses average educational attainment further. Decreasing marginal returns to education are likely to explain this phenomenon as well as the fact that the estimates in columns (3) and (4) exceed the estimates in columns (1) and (2) of table 3. However, non of these differences is statistically significant. An alternative explanation could be that low educated parents adopt out of need and therefore send their children to school no longer than mandatory. For instance they adopt a child in order to have an additional worker to provide for the family. If parents and children are both low educated, the intergenerational schooling correlation may be relatively high. However, the information about adoption practices in Indonesia do not support the hypothesis that parents adopt out of need.

The evidence presented in this section does not indicate that nature constitutes a decisive factor in the intergenerational schooling persistence. In the following I discuss potential violations of identifying assumptions and subsequently deliver additional evidence by conducting the same analysis for a different measure of educational attainment.

4.1.1 Concerns regarding the identifying assumptions

Five general concerns arise when comparing OLS estimates between biological and adopted children. The first is sample selection in the sample of adoptive parents, i.e. adoptive parents differ from biological parents. If this is the case estimates for the two groups are not comparable. For instance, if adoptive parents care less about the edu-

Table 2: OLS results of the intergenerational schooling persistence measured in years of schooling for biological children that are older than 22.

	Full sample		Parent's years of education <12	
	(1)	(2)	(3)	(4)
	Years of schooling	Years of schooling	Years of schooling	Years of schooling
	β / SE	β / SE	β / SE	β / SE
Parents' years of schooling	0.277***	0.274***	0.291***	0.287***
	(0.007)	(0.007)	(0.009)	(0.009)
Gender child (Male=1)	−0.033	−0.044	−0.026	−0.035
	(0.086)	(0.086)	(0.091)	(0.091)
Age child	−0.085		−0.076	
	(0.063)		(0.066)	
Age parent	0.074	0.084*	0.071	0.080
	(0.051)	(0.051)	(0.052)	(0.053)
Age spouse	0.018	−0.001	0.019	0.000
	(0.042)	(0.043)	(0.044)	(0.044)
Age squared child	0.001		0.000	
	(0.001)		(0.001)	
Age squared parent	−0.001	−0.001	−0.000	−0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Age squared spouse	−0.000	0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Rural area (when child)	−1.138***	−1.138***	−1.150***	−1.149***
	(0.092)	(0.092)	(0.097)	(0.096)
Observations	4594	4594	4224	4224
R^2	0.344	0.357	0.280	0.296

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note: Columns (1) and (3) include covariates as indicated in the table. In columns (2) and (4) child cohort fixed effects are added and child age variables are excluded.

Table 3: OLS results of the intergenerational schooling persistence measured in years of schooling for adopted children that are older than 22.

	Full sample		Parent's years of education <12	
	(1)	(2)	(3)	(4)
	Years of schooling	Years of schooling	Years of schooling	Years of schooling
	β / SE	β / SE	β / SE	β / SE
Parents' years of schooling	0.229***	0.220***	0.335***	0.304***
	(0.051)	(0.055)	(0.072)	(0.078)
Gender child (Male=1)	−1.423**	−2.065***	−1.415**	−2.118***
	(0.658)	(0.730)	(0.679)	(0.767)
Age child	−0.088		−0.244	
	(0.624)		(0.632)	
Age parent	−0.435	−0.522	−0.373	−0.442
	(0.315)	(0.336)	(0.321)	(0.347)
Age spouse	−0.133	−0.060	0.005	−0.014
	(0.346)	(0.418)	(0.358)	(0.423)
Age squared child	0.002		0.004	
	(0.010)		(0.010)	
Age squared parent	0.005	0.006	0.004	0.005
	(0.003)	(0.004)	(0.004)	(0.004)
Age squared spouse	0.001	0.001	−0.000	0.000
	(0.004)	(0.005)	(0.004)	(0.005)
Rural area (when child)	−1.351*	−1.086	−1.143	−0.980
	(0.705)	(0.822)	(0.716)	(0.877)
Observations	96	96	89	89
R^2	0.338	0.476	0.356	0.500

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note: Columns (1) and (3) include covariates as indicated in the table. In columns (2) and (4) child cohort fixed effects are added and child age variables are excluded.

cation or health conditions of their child than biological parents do, differences between the two estimates cannot be attributed to genetic effects. The same consequence would result if adoptive parents were more motivated to adopt out of need than biological parents are to have another child.

The second concern is that the adoption allocation process is non-random. Randomness implies that child and parent pre-adoption characteristics are not correlated. However, if children are allocated to their adoptive parents based on background characteristics the randomness assumption is violated and the estimate will be upward biased (in case of positive correlation). The reasoning behind this is simply that the genetic effects were not successfully eliminated. This is likely to be the case for the sample of adoptees I use. For one thing, parents can visit orphanages to search for a child to adopt. While doing so they are likely to learn about the child's background and in cases where the child's biological parents are still alive adoptive parents may even meet the biological parents as they have to approve to the adoption. For another thing, the sample includes cases of adoption of step children who carry half of the genes of their parents. Therefore, I expect the estimates for adopted children to be upward biased due to non-randomness.

The third concern is that adoption takes place at a too old age. Children under the age of six are given absolute preference for adoption in Indonesia. However, when children are not being adopted very early in their lives their previous environment such as their biological parents' nurture or the orphanage may have a big influence on their educational attainment. It is plausible that children that spent several years in an orphanage have a clear disadvantage compared to biological children which may not be offset by good parenting after adoption. This will put downward pressure on the OLS estimate for adopted children because their educational attainment will be less correlated with their adoptive parents' education. Unfortunately, there is no information about the timing of adoption available and therefore I cannot test for too-late-adoption.

The last concern is that the sample of adoptees is too small to identify any differences in the estimates from the biological sample. The estimates in table 3 come with large standard errors due to the small sample size. Therefore, even if those estimates represent the true intergenerational schooling persistence net of genetic effects I cannot conclude a statistically significant difference between the coefficients of biological and adopted children.

In sum it is difficult to judge whether the upwards directed selection bias or downward directed 'late-adoption' bias dominates. Hence, the presented evidence so far is inconclusive about the importance of nature in the intergenerational schooling persistence. To bolster the results so far the following section investigates estimates for a different measure of educational attainment.

Table 4: OLS results of the intergenerational schooling persistence measured in proceeding to secondary school for biological and adopted children.

	Biological children		Adopted children	
	Full sample	Parent edu <12	Full sample	Parent edu <12
	(1)	(2)	(3)	(4)
	Secondary school	Secondary school	Secondary school	Secondary school
	β / SE	β / SE	β / SE	β / SE
Parents proceeded to secondary school	0.103***	0.104***	0.161***	0.184***
	(0.005)	(0.005)	(0.045)	(0.053)
Gender child (Male=1)	−0.003	−0.003	−0.130*	−0.139*
	(0.007)	(0.008)	(0.071)	(0.076)
Age child	−0.001	−0.002	−0.024	−0.030
	(0.003)	(0.003)	(0.034)	(0.037)
Age parent	0.001	0.001	−0.038	−0.035
	(0.004)	(0.004)	(0.033)	(0.035)
Age spouse	−0.000	0.000	0.019	0.024
	(0.003)	(0.004)	(0.033)	(0.035)
Age squared child	−0.000	−0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.001)
Age squared parent	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Age squared spouse	0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Rural area (when child)	−0.089***	−0.096***	−0.188**	−0.166**
	(0.008)	(0.009)	(0.076)	(0.080)
Observations	7533	6893	133	122
R^2	0.107	0.097	0.225	0.217

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

4.1.2 OLS robustness check: Proceeding to secondary school

While Indonesia introduced compulsory primary schooling in the 1980s the percentage of students continuing with secondary education remained relatively low. Therefore, the intergenerational persistence in terms of proceeding to secondary school is of high policy relevance and can be used to reinforce the results from the previous section. Columns (1) and (2) of table 4 present estimates for biological children and columns (3) and (4) for adopted children.[15] By the time a child's educational attainment is observed every child had the chance to proceed to secondary school. Therefore, these estimates do not underly censoring and no age limitations are imposed on children in table 4.[16]

Because ability correlations should be eliminated in adoptive families, it is surprising that the estimates for adopted children (0.184) are higher than for biological children (0.103). The same concerns regarding the identifying assumptions with respect to years of education hold for proceeding to secondary school. However, sample selection is the only concern that is capable of explaining these results. Therefore, differences between adoptive and biological parents are likely to explain why the intergenerational persistence in proceeding to secondary school is higher for adoptees than it is for biological children. However, the difference between the estimates of the two groups is not statistically significant, wherefore no definite conclusions can be drawn.

In summary, from the comparison of estimates of biological and adopted children, I was unable to identify intergenerational persistence in education that is driven by the transmission of genetic endowments. So far, neither persistence in years of schooling nor persistence in proceeding to secondary school indicate that innate characteristics matter in the intergenerational process. The empirical strategy applied is far from perfect because identifying assumptions are likely to be violated. Therefore, I apply an instrumental variables technique in the subsequent sections to shed further light on the family environmental factors that affect children's educational attainment.

4.2 Instrumental variables estimation

4.2.1 2SLS: Years of schooling

To examine the causal relationship between parental and child education this section applies an instrumental variables strategy. I use a large scale school construction program from the 1970s to instrument the sum of parents education. Educational attainment is measured in years of schooling and children younger than 22 are excluded from the sample due to censoring. For the IV estimates to be comparable to the baseline OLS

[15] The coefficients are robust to inclusion of child cohort fixed effects.

[16] I estimated the intergenerational persistence in proceeding to secondary school also for the sample where children younger than 23 were excluded. These estimates are slightly higher and the difference to the sample including all children is insignificant.

results I reestimate table 2 with the same sample that is used for the IV estimation. The OLS estimates in columns (1) and (3) of table 5 are somewhat higher than the ones in table 2 but the difference is not statistically significant. Therefore, I conclude that no sample selection took place when the IV sample was created.[17]

As expected, the IV estimates are lower than the OLS estimates. Column (2) suggests that an additional year of either parent's schooling increases the educational attainment of the child by 0.285 years whereas the OLS estimation suggests a coefficient of 0.294. However, the difference between the two estimates is fairly small.[18]

There are many reasons that potentially explain the similarity between the OLS and IV coefficients. One explanation is that the IV estimate represents a local average treatment effect (LATE). Because the INPRES program targeted primary school students the program effect on low educated people is higher than the effect on highly educated people who are assumed to have attended primary school anyways.[19] Also the program was designed in a way to support disadvantaged areas that have low primary school enrollment rates. Therefore, the group of compliers comprises rather low educated individuals. As argued in section 4.1 heterogeneous treatment effects are likely such that low educated people potentially have higher social returns to education.

An alternative explanation is that the instruments are too weak and therefore the first stage's predictive power is not sufficient. The F-tests for the IV estimates in columns (2) and (4) are 4.12 and 4.95 respectively. Thus, the first stage variables are jointly significant, although, the relationship is not very strong. Consequently, weak instruments may lead to an upward bias of the impact of parents' education.[20]

To test whether the IV estimation underlies weak instruments or represents a LATE one can look at a sample that is restricted to lower educated parents because they are expected to be more responsive to the program.[21] Going from column (2) to (4) the first stage F-test increases by 0.82 and the coefficient increases to 0.308. I also tested the relevance of the first stage for a more restricted sample with parents having less than 9 years of education, but the instruments power did not increase further.[22] A conclusion about the weakness of the instrument and the resulting bias is difficult. However, the

[17] For the IV estimation information about parents' region of birth was required which reduced the sample size.

[18] The results are robust to the inclusion of child cohort effects.

[19] Primary education was not mandatory at the time the parents generation went to school.

[20] The bias is assumed to be positive because the instruments and parents years of schooling are positively correlated.

[21] Testing whether the school construction program did not affect more educated parents by restricting the sample to highly educated parents or by including interaction terms of the instruments and a 'higher-education' dummy in the first stage regression was impossible. This is because after restricting the sample there exists not sufficient variation in the first stages dependent variable, parents education.

[22] The estimate drops to 0.13 for parents with less than 9 years of schooling. However, this is likely to be due to the downward pressure induced by the strong sample restriction (see Holmlund et al. (2011)).

Table 5: OLS and IV results of the intergenerational schooling persistence measured in years of schooling for children older than 22.

	Full sample		Parent's edu <12	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
	Years of schooling	Years of schooling	Years of schooling	Years of schooling
	β / SE	β / SE	β / SE	β / SE
Parents' years of schooling	0.294***	0.285***	0.316***	0.308***
	(0.012)	(0.048)	(0.014)	(0.049)
Gender child (Male=1)	0.036	0.189	0.023	0.172
	(0.152)	(0.135)	(0.158)	(0.140)
Age child	−0.314	−0.192	−0.456	−0.484
	(0.921)	(0.825)	(0.955)	(0.837)
Age parent	0.305	−2.579	0.246	−3.884*
	(0.249)	(2.219)	(0.256)	(2.247)
Age spouse	0.057	−0.124	0.066	−0.051
	(0.057)	(0.254)	(0.058)	(0.241)
Age squared child	0.004	0.002	0.007	0.008
	(0.017)	(0.016)	(0.018)	(0.016)
Age squared parent	−0.004	0.026	−0.003	0.042
	(0.003)	(0.028)	(0.003)	(0.028)
Age squared spouse	−0.000	0.001	−0.000	0.001
	(0.000)	(0.002)	(0.000)	(0.002)
Rural area (when child)	−0.746***	−0.941***	−0.762***	−0.972***
	(0.156)	(0.275)	(0.161)	(0.269)
Observations	1436	1436	1357	1357
R^2	0.341	0.593	0.305	0.580
F-test (first stage)	/	4.12	/	4.95

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

increase in the IV coefficient supports the LATE hypothesis.

Another explanation for the similarity in OLS and IV coefficients is that the OLS estimate is prone to error-in-variables bias whereas the IV estimate is not. Hence, the coefficients in columns (1) and (3) may suffer from downward bias.

At last I discuss the exogeneity of the instruments. An instrument is exogenous if it has no effect on children's schooling other than the indirect effect through parents' schooling. One obvious violation is the direct effect on children's schooling, that is, the constructed schools will most certainly still function as schools at the time the child generation goes to primary school. Hence, if the household did not move away from the parents region of birth the child will benefit from the INPRES program directly. However, compulsory primary schooling was introduced in the 1980s such that the child would have completed primary school in any case. Nevertheless, indirect effects like reduced travel time to school potentially have an impact on educational attainment. Another threat to exogeneity is that teacher quality decreased or that the teacher-child ratio decreased since the constructed schools had to be staffed with new teachers. According to Duflo (2001) the proportion of teachers meeting the minimum qualification requirements did not worsen significantly and the stock of teachers grew by 43 percent whereas the stock of primary schools doubled. Therefore, impacts working through this channel are not considered severe. It is important to recall that children of program prone parents attended primary school 20 years after the INPRES program was implemented. Therefore it is unlikely that the program had a large impact on the child generation other than through parents' educational attainment.

Having said that, it is likely that there exists a larger difference between OLS and IV results which I am unable to identify. Nevertheless, the IV results are consistent with evidence from the comparison of biological and adopted children. Both identification strategies do not indicate that genetic transmissions are of relevance for the intergenerational persistence in educational attainment. Instead, nurture seems to play an important factor. And furthermore, the evidnce from this section indicates that a large part of the nurture component can be attributed to parental education. In the following section I will apply the IV approach to the intergenerational persistence with respect to proceeding to secondary school and deliver further support for the preceding results.

4.2.2 2SLS: Proceeding to secondary school

Table 6 presents OLS and IV estimates for the intergenerational persistence in terms of proceeding to secondary school. Because the variable proceeding to secondary school is not prone to censoring I only show results for the full sample of children. In table 6 the OLS estimates are based on the same sample as the IV estimation is and they are very similar to the OLS coefficients for biological children in table 4. Therefore, I conclude

that no sample selection took place when the IV sample was created.

The IV estimate in column (2) indicates a causal intergenerational relationship of 0.217 which is much higher than the OLS estimate of 0.100 in column (1). This is surprising because the IV estimate represents the causal effect which is thought to constitute a share of the overall schooling persistence. Same as in the previous section these estimates are potentially prone to various shortcomings or biases. First, a comparison of columns (1), (3) and (5) shows that the IV coefficients cannot reflect local average treatment effects because the OLS estimates almost stay constant when I restrict the sample to parents with lower education. Second, the OLS estimates are not likely to suffer from error-in-variables bias because every individual with more than primary education continued to secondary school irrespective of the final educational attainment. Therefore, the proceeded-to-secondary-school variable is more robust to measurement error than years of education is and I expect the downward bias to be small. Third, the instrument's exogeneity is arguable for the same reasons as stated above, and fourth, the instruments may not be relevant. The first stage F-test increases from 3.89 (column (2)) to 4.24 (column(4)) to 9.10 (column (6)) the more I restrict the sample to parents with lower education. The increase in F-tests indicates that the INPRES program mainly affected low educated people in regard to proceeding to secondary school.[23] This is in line with the weak instrument hypothesis. The difference between the OLS and IV estimates increases for the sample with the strongest first stage to 0.272. Hence, while the estimations in columns (2) and (4) potentially underlie weak instruments, it does not explain why the OLS coefficients are lower than the IV coefficients. In fact, that the IV coefficient is lager than the OLS coefficient, is consistent with the finding that the estimates for adopted children are larger than the estimates for biological children in table 4. Nevertheless, it is difficult to plausibly reason this phenomennon.

Taken as a whole the two identification strategies deliver consistent estimates for each measure of educational attainment. For years of education the intergenerational persistence ranges between 0.23 and 0.29 whereas the persistence for proceeding to secondary school can only be narrowed down to range between 0.10 and 0.27. Across all estimation techniques and measures of educational attainment I find no indication of genetic transmission to play a major role in explaining the intergenerational schooling persistence. Therefore, I hypothesize that the intergenerational schooling persistence in Indonesia is mostly driven by nurture. And furthermore, the evidnce from this section indicates that a large part of the nurture component can be attributed to parental education.

[23] Testing whether the school construction program did not affect more educated parents by restricting the sample to more educated parents or by including interaction terms of the instruments and a 'more-education' dummy in the first stage regression was impossible because then there exists not sufficient variation in the first stages dependent variable, parents education, as this would take the value 2 in all cases.

Table 6: OLS and IV results of the intergenerational schooling persistence measured in proceeding to secondary school.

	Full sample		Parent's edu <12		Parent's edu <9	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
	Secondary school	Secondary school	Secondary school	Secondary school	Secondary school	Secondary school
	β / SE	β / SE	β / SE	β / SE	β / SE	β / SE
Parents proceeded to secondary school	0.100***	0.217***	0.101***	0.208***	0.107***	0.272***
	(0.006)	(0.035)	(0.006)	(0.036)	(0.021)	(0.099)
Gender child (Male=1)	−0.016	−0.009	−0.016	−0.008	−0.022	−0.002
	(0.011)	(0.011)	(0.012)	(0.012)	(0.018)	(0.016)
Age child	0.006	0.011	0.006	0.009	0.017	0.008
	(0.013)	(0.013)	(0.014)	(0.014)	(0.021)	(0.019)
Age parent	−0.013	0.031	−0.016	0.275**	−0.019	0.420**
	(0.010)	(0.063)	(0.011)	(0.114)	(0.017)	(0.166)
Age spouse	0.005	−0.007	0.006	−0.011	0.006	−0.008
	(0.004)	(0.025)	(0.004)	(0.027)	(0.006)	(0.036)
Age squared child	−0.000	−0.000	−0.000	−0.000	−0.001	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age squared parent	0.000	−0.000	0.000	−0.004***	0.000	−0.006***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.002)
Age squared spouse	−0.000	0.000	−0.000	0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rural area (when child)	−0.076***	−0.001	−0.083***	−0.021	−0.115***	−0.077**
	(0.011)	(0.023)	(0.012)	(0.023)	(0.018)	(0.032)
Observations	3337	3337	3115	3115	1948	1948
R^2	0.093	0.228	0.087	0.253	0.038	0.335
F-test (first stage)	/	3.89	/	4.24	/	9.10

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

5 Conclusion

This thesis applies two identification strategies to examine which family environmental factors affect children's education in Indonesia. The first strategy aims to disentangle nature from nurture factors in the intergenerational schooling correlation. I compare OLS estimates of the schooling persistence between biological children and adopted children. If the underlying assumptions that (1) adoptive parents are similar to biological parents and that (2) adoptees' and adoptive parents' background characteristics are uncorrelated hold, then the difference between the estimate for biological and adopted children can be attributed to innate characteristics. That is, the OLS estimate for adopted children represents the intergenerational schooling persistence net of genetic effects (nature). When estimating this relationship with respect to years of schooling I find that an additional year of either parent's schooling increases child schooling by about 0.23 years and that the estimates for adoptees are lower than those for biological children by about 0.05. However this difference is insignificant. When estimating this relationship with respect to proceeding to secondary school I find estimates for adoptees that are larger than those for biological offsprings. This is difficult to reason, but the difference between the two estimates is insignificant.

The comparison of biological children and adoptees is prone to violations of identifying assumptions. These are (1) adoptive parents differ from biological parents, (2) non-random allocation of adoptees to their adoptive parents, (3) adoption at a too old age and (4) the sample of adoptees is too small to identify differences from biological children. These shortcomings potentially lead to bias in the OLS estimates for adopted children where the net direction of the bias remains unclear.

The second strategy identifies a causal effect of parental education on child education. The causal effect can be thought of as the extent to which parental education contributes to the nature component. I use an IV strategy that creates variation in parents education by instrumenting it with a large scale school construction program. For years of schooling, I find an IV estimate of 0.285 which is lower than the OLS estimate of 0.294. For proceeding to secondary school, I find an IV estimate of 0.217 which is much higher than the OLS estimate of 0.100. The comparison between OLS and IV estimates potentially underlies some deficits: (1) the OLS estimate is prone to downwards directed error-in-variables bias, (2) the IV estimate may reflect a local average treatment effect, (3) the instruments are weak and (4) the exogeneity of the instruments is contestable. Therefore, at least for years of education, the OLS estimate is likely to be downward biased and the IV estimate is likely to be upward biased. For proceeding to secondary school the net direction of bias is more difficult to predict.

In sum, I find no evidence that the transmission of inheritable characteristics matters to explain the intergenerational schooling persistence. Although the estimates are likely to suffer from biases discussed above, the pattern is consistent throughout estimation

techniques and measures of educational attainment. This is surprising in the light of studies from Scandinavia and the United States that find the larger part of the intergenerational schooling persistence to be due to nature instead of nurture. In some cases they even find a causal relationship close to zero. My results also contradict evidence from Rwanda that suggests that at most parts of the intergenerational schooling persistence can be explained by nurture.

Additionally, I find that parental education contributes a major part to the nurture factor. This is particularly encouriging for policy makers as it underlines the effectiveness of schooling to improve parents nurturing skills which in turn have a positive effect on the future generation's schooling outcomes.

I hypothesize that the effect of an additional year of parental education has larger effects on nurture in Indonesia where education levels and child rearing standards are low compared to more advanced countries. If educational attainment in the parents' generation can be improved not only will the current generation benefit in terms of private returns to schooling but also will the schooling of the next generation be increased. That is social returns to education exist. However, I am unable to draw any definite conclusions from this study regarding the importance of nature and nurture. Therefore, further evidence from other developing countries is essential and inevitable in order to understand the nurture production functions of developing countries.

References

Behrman, J. R., Gaviria, A. & Székely, M. (2001), 'Intergenerational mobility in latin america', *Economia* **2**(1), 1–31.

Behrman, J. R. & Rosenzweig, M. R. (2002), 'Does increasing women's schooling raise the schooling of the next generation?', *American Economic Review* pp. 323–334.

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2005), 'Why the apple doesn't fall far: Understanding intergenerational transmission of human capital', *American Economic Review* **95**(1), 437–449.

Chevalier, A. (2004), Parental education and child? s education: A natural experiment, Technical report, IZA Discussion paper series.

De Walque, D. (2005), 'Parental education and children's schooling outcomes: Is the effect nature, nurture, or both? evidence from recomposed families in rwanda', *Evidence from Recomposed Families in Rwanda (January 2005). World Bank Policy Research Working Paper* (3483).

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment', *American Economic Review* **91**(4), 795–813.

Emran, M. S. & Sun, Y. (2011), Magical transition? intergenerational educational and occupational mobility in rural china: 1988-2002, *in* 'American Economic Association meeting, Denver, and Northeast Universities Development Consortium conference, Yale University'.

Holmlund, H., Lindahl, M. & Plug, E. (2011), 'The causal effect of parents' schooling on children's schooling: a comparison of estimation methods', *Journal of Economic Literature* **49**(3), 615–651.

Klein, R. & Vella, F. (2009), 'A semiparametric model for binary response and continuous outcomes under index heteroscedasticity', *Journal of Applied Econometrics* **24**(5), 735–762.

Oreopoulos, P., Page, M. E. & Stevens, A. H. (2006), 'The intergenerational effects of compulsory schooling', *Journal of Labor Economics* **24**(4), 729–760.

Plug, E. (2004), 'Estimating the effect of mother's schooling on children's schooling using a sample of adoptees', *American Economic Review* pp. 358–368.

Sacerdote, B. (2007), 'How large are the effects from changes in family environment? a study of korean american adoptees', *The Quarterly Journal of Economics* **122**(1), 119–157.