Education and Human Capital Externalities: Evidence from Colonial Benin[*]

Leonard Wantchekon[†]
Natalija Novta[‡]
Marko Klašnja[§]

October 6, 2012

Abstract

We use a unique dataset on students from the first regional schools in colonial Benin, to investigate the effect of education on income, occupation and political participation. Because of the near random selection of the school location and the first student cohorts, we can estimate the effect of education by comparing the treated to the untreated living in the same village, as well as those living in villages where no schools were set up. We find a significant positive treatment effect of education on a number of outcomes. For example, the treated have better living standards, are less likely to be farmers and more likely to be politically active. Second, we look at the outcomes of the descendants. Similar to the first-generation effects, parents' education have a large positive effect on their children's educational attainment, living standards, and social networks. Third, there are large positive externalities of education in the second generation – descendants of the untreated in villages with a school have substantially better outcomes than descendants in villages without a school. We find evidence that these externalities run through the parents' enhanced social networks. Fourth, the strength of extended families is documented as nephews and nieces directly benefit from education of their uncles – they are almost as equally educated as the students' children, and are more educated than descendants without any educated members in a family. We demonstrate that these within-family externalities represent a "family-tax," as educated uncles transfer resources to the extended family.

[†]*Corresponding author*: Department of Politics, Princeton University. E-mail: lwantche@princeton.edu
[‡]Department of Economics, New York University
[§]Wilf Family Department of Politics, New York University

"An educated child is like a lantern in your house at night." Eloi Gainsi, farmer and religion teacher, Zagnanado (Benin)

1   Introduction

In a seminal paper, Acemoglu et al. (2001) establish a strong correlation between earlier colonial settlement patterns and long-term development. The mechanism responsible for this relationship, they argue, is institutional – more precisely, they suggest that higher settler mortality led colonizers to choose more extractive institutions, while lower rates led them to establish growth-promoting institutions. In contrast to this view, Gennaioli et al. (2012) argue that much of cross-country and within-country variations in regional incomes are explained by human capital. This may be the case because human capital plays a special role in promoting development at the initial stages of colonization. This view appears to be consistent with Huillery (2009), which suggests that regions within Francophone Africa where schools were established early in the colonial period have seen better development outcomes today.

There are thus two competing views of how settlement patterns influence long-run growth: one emphasizes human capital while the other argues for institutional quality as the fundamental cause of development. In the absence of variation in micro-level data, it is quite difficult to disentangle these two competing explanations for the relationship between colonial settlement and long-term economic growth. For instance, if settlers simultaneously invested fewer resources in education and established weaker institutions where mortality was higher, then both human capital and institutions could be influencing long-term development outcomes. One way to solve the problem of joint determination is by examining sufficiently rich micro-level evidence on how European settlements affected a range of social outcomes at both the aggregate and individual level, and empirically discern the relationship between average levels of education, the quality of institutions and development outcomes at the local level. This paper is a first step in this direction. Using a unique dataset, we provide micro-evidence on the immediate impact of the first colonial schools in Benin on measures of income, professional achievement and occupational choice. We consider only schools established in regions with no prior exposure to European influence at the time. In other

words, we use data collected in regions where colonial institutions were established after, not before, education opportunities were made available to the local population. The data originates from Wantchekon (2012) and uses information provided by school and church archives and face-to-face interviews with "informants." The data was collected from four sites (Kandi, Natitingou, Save and Zagnanado), spans several generations and includes measures of income, professional achievement, and educational attainment.

The subjects are divided into 3 groups. The first group is comprised of those who were members of the first two cohorts from the newly built schools.[1] We call this group Treatment 1. The second group or Treatment 2 is comprised of children from the village where the school was located (or from its surrounding hamlets), and who were not selected to go to school but were part of the social networks of those who did.[2] The third group or control group is comprised of children from villages that were geographically separated or otherwise sufficiently distant from villages that had schools.[3] These children were neither selected to attend a school nor in frequent contact with those who were.

Wantchekon (2012) compiled a list of directly treated subjects through church and school records. Since no such records exist for those in Treatment 2 or for those in the control group, he relied on fieldwork and face-to-face interviews to generate samples of members of these groups. Their second or third generation descendants were interviewed after being randomly sampled from the current population census. The information they provided was verified and supplemented by church and colonial records such as ID cards. The teams of enumerators also visited local cemeteries to verify birth year and death year of several generations of villagers included in the samples. The data also includes information on the subjects' siblings, children, nephews and nieces.[4] The process was repeated for the third and fourth generation of all subjects in the three distinct groups enumer-

---

[1] As shown below, the selection process that determined who would attend the new schools was quasi-random-school officials selected children by convenience from the area in which the school was built, who were then enrolled.

[2] It is assumed that members of this group (treatment 2) interact with those selected to go to school (treatment 1) several times a week.

[3] The historical evidence suggests that these villages were very similar to the "treated village" in terms of income, social economic structure. For example, more than 95% of the population were subsistence farmers (D'Almeida Topor 1995).

[4] Note that, if a subject had more than ten siblings, ten were randomly selected. The same procedure was used to select the sample children (up to ten), and nephews and nieces (up to thirty).

ated above. The final dataset thus includes social and demographic characteristics on the initial subjects, their extended families, and an additional two generations, allowing for the measurement of the effect of education over time and across space.[5]

The evidence provided by the data suggests a significant positive effect of education on a number of outcomes. For example, the first generation of treated individuals and their children experience better living standards, are less likely to be farmers and have more friends among white settlers. We also find large positive externalities of education in the second generation: descendants of the untreated in villages with a school have substantially better outcomes than descendants in villages without a school. In addition, the results suggest that nephews and nieces greatly benefit from education of their uncles: they are almost equally as educated as the students' children, and are far more educated than descendants without any educated members in a family. The evidence suggests that these externalities manifest through parents' enhanced social networks and fostering aspirations.

2 Literature

Our paper contributes to several strands of literature in economic history and development economics. It provides empirical evidence on the specific way human capital influenced the process of development in West Africa. We build on recent literature on the colonial legacy in the area of education (e.g. Nunn 2009, Nunn 2010, Woodberry and Shah 2004), by tracking down the first students and their descendants.

The paper also contributes to studies on extended families and human capital externalities. While the importance of extended families have been questioned in the developed world (see Altonji et al. 1992), they play significant role in developing countries (see Angelucci et al. 2010, Cox and Fafchamps 2007, La Ferrara 2003, Shavit and Pierce 1991). We provide rigorous empirical support for these findings in the context of colonial Benin. The evidence also suggests that we having an educated parent or uncle is nearly as beneficial as growing up in a village with a newly

---

[5]The dataset includes information about gender, number of siblings, number of children, occupation, years of schooling, housing quality, clothing type or style, social network (e.g., number of white friends, travel frequency), political participation, party affiliation, and membership in civil society organizations.

established school. The result is consistent with Gould et al. (2011), which found that, in Israel, the infrastructure quality of the community in which Yemenite immigrants were (randomly) placed in 1949 still affects their, and their offspring's, education nearly half-a-century later.

Our results are also consistent with recent findings on peer effects. For example, Lalive and Cattaneo (2009) and Bobonis and Finan (2009) find that ineligible students have benefited from the PROGRESA program in Mexico, due to neighborhood peer effects. In the United States, Borjas (1992) and Borjas (1995) have shown that the ethnic community in which children grow up to a large extent determines their later labor market outcomes, and Topa (2001) shows local spillovers are particularly strong in areas with less educated workers.

Why is the social environment in which children grow up important for later social and financial outcomes? In line with Chiapa et al. (2010), we argue that aspirations may play a role. They find that exposure to typically female health professionals through the PROGRESA program increases educational aspirations, especially for education of girls. Mookherjee et al. (2010) and Ray (2006) provide a theoretical model and discussion of how aspirations may be important for education and escape from poverty.

3   Context and Experiment Design

Benin, formally Dahomey, was colonized in 1894 when French troops, led by General Alfred Dodds, defeated the army of the kingdom of Dahomey. The country was officially colonized when Behanzin, the King of Abomey, surrendered to Dodds after three years of war. Prior to colonial administration and in the shadow of the slave trade, Catholic missions were established in the coastal town of Agoue (1874), Porto Novo (1864) and the interior towns of Zagnanado (1895). There were two types of missions: those established in regions with prior European presence in the form of commercial trading posts and military settlements, such as Porto Novo and Agoue, and those, with no prior European influence such as Zagnanado.

[Figures 1 and 2 about here.]

Following the conquest of the Kingdom of Dahomey, formal colonial administration was first established in Southern and Central Benin in areas with prior exposure to European influence and formal education. This was not the case in the Northern part of country (Borgou and Atakora) where the formal colonial institutions were established after creation of regional schools – examples include Save (1913), Kandi (1911) and Natitingou (1921).

In light of the chronological order of events, it is reasonable to assume that the locations of schools were quasi-randomly selected, since villages where they were set up were under no prior European influence. To see this more clearly, consider for example the process by which a team of missionaries of the Societes des Missions Africaines (SMA) set up the Catholic school in Zagnanado in 1895 (Dupuis 1961). They took a boat in Porto Novo with the ambition to create a church in the interior territories of Benin. After several days of travel, they were suddenly and unexpectedly invaded by "armies" of mosquitoes around Sagon, a village located about 20 miles from Zagnanado, at the other side of Oueme River. The following day, they made the decision to stop the mission and to settle in Zagnanado. Had they had more perseverance or had mosquitoes not attacked them, they could have continued the mission and settled in a village located further North from Zagnanado, at the other side of the Oueme River. Therefore, it is reasonable to view Zagnanado catholic school as a natural experiment.

Similar happenstance lead to the founding of Natitingou primary school. According to several historical sources, the colonial administration was unsure whether to set up the regional school in Kouande, Boukombe and Djougou, the three main cities in the Atakora region. In the end, they decided to settle in Natitingou, a small town, which was equidistant from the three cities. The decision was in part driven by lack of human resources and bad road quality (Cornevin 1981, Garcias 1970, Mercier 1968).

The selection of the first cohort to attend the newly built schools was also quasi-random. Typically, school officials – often a priest – selected children from a communal area and without prior knowledge of the child or the child's family. There was no positive selection effect that could have resulted in wealthy or predisposed children to be members of the first cohorts. The local chiefs, in fact, were often reluctant to send their own children to the new schools, fearing that

doing so would compromise their ability to remain independent of the colonial administration and the missionaries (Cornevin 1981, Dupuis 1961). Thus, the children that became members of the first cohort were, if anything, more likely to be poor and poorly disposed toward achieving positive educational outcomes.

Given the quasi-random selection of both the treated villages and the treated children within those villages, we are able to identify the effect of education by comparing outcomes across several treatments.

4 Results

4.1 Summary statistics

Table 1 summarizes the most important variables for the first-generation inhabitants of the villages in our sample. In the first row, we list by treatment status the mean values for the only pre-treatment variable we were able to collect – the number of siblings. The first-generation students did not differ from the rest of the sample in the number of siblings, nor were those inhabitants who resided in villages with school but did not attend school different from inhabitants of villages without a school.[6] Ideally, we would examine other pre-treatment covariates, but we were unable to collect reliable data.

Looking down Table 1, we see that setting up schools appears to have had a profound and apparently long-lasting effect on the children that were chosen to attend schools and their descendants. Table 1 presents the summary statistics that compare the first generation of students and their contemporaries. Among the children chosen to go to school, almost all (96%) completed primary education and 10% of them further competed secondary education. In the first generation no one went on to university, which is hardly a surprise given that these children were born at the turn of the 20th century and that no universities were available in Western Africa at the time.

In terms of income, those randomly chosen to attend school clearly have superior outcomes to

---

[6]Several tests show that this pre-treatment covariate is balanced across the three groups. Any pairwise difference in means fails to be significant at conventional levels. Moreover, comparing treatment 1 to controls gives similar results, irrespective of whether we use the difference-in-means test, the rank-sum test, or the Kolmogorov-Smirnov test. Finally, the Kruskal-Wallis test for multiple treatments also returns a negative result.

Table 1: Summary Statistics for the First Generation

	Treated parents	Untreated parents in village w/ school	Untreated parents in village w/o school
Pre-treatment			
Number of siblings	3.370	3.059	2.964
	(2.366)	(2.326)	(2.114)
Education			
Primary or more	0.963	0.008	0.008
	(0.189)	(0.092)	(0.091)
Secondary or more	0.098	0.000	0.000
	(0.299)	(0.000)	(0.000)
University	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Income			
Farmer	0.143	0.784	0.842
	(0.352)	(0.414)	(0.367)
Water	0.258	0.146	0.092
	(0.440)	(0.355)	(0.290)
Electricity	0.101	0.024	0.007
	(0.303)	(0.155)	(0.081)
Vehicle	0.476	0.182	0.195
	(0.502)	(0.387)	(0.397)
Income scale	0.677	-0.195	-0.188
	(1.159)	(0.887)	(0.835)
Networks			
French language	0.955	0.085	0.013
	(0.208)	(0.280)	(0.114)
White friends	0.457	0.084	0.035
	(0.502)	(0.278)	(0.186)
Social networks scale	1.661	-0.350	-0.451
	(0.864)	(0.539)	(0.425)
Observations	89	164	152

Note: Standard deviations in parentheses.

either the uneducated from the same village, or those from untreated villages. For example, only 14% of the educated students become farmers, while farming is clearly the dominant occupation among the uneducated (about 80%). We also observe that the educated are more likely to have running water in their homes (26%), electricity (10%) and to own a vehicle (48%). The uneducated in villages with and without a school have worse income outcomes, and do not seem to be different from each other, as we will formally show in the next section (Table 3).

We also include a measure of income based on factor analysis using several indicators of living standards such as those listed in the Table 1. Other variables include house wall material, house/land/shop ownership, household equipment, means of transportation, travel patterns and type of attire. We see that also in terms of this composite measure of income, the educated have clearly higher scores than the uneducated.

The presence of a school in a village, however, does seem to have some indirect effect on the uneducated in addition to the direct effect on education. We expect to observe that the educated are more likely to speak French, have friends among whites and score higher on a social networks scale. The interesting observation is that the uneducated in villages with school seem to also score higher than those in village without a school. The social networks scale was again coded based on factor analysis using information about membership in organizations (religious, business, sports), languages spoken (national, foreign), friends among whites and other local ethnic groups, and participation in local politics.

4.2 First-Generation Effects: Income, Social Networks, and Political Participation

We first evaluate the effects of being individually treated and treated at the village level among the first generation students and their contemporaries. Because we argue that children were randomly chosen to attend the schools, and that villages in which schools were set up were chosen in a quasi-random manner, any effects that we find in the first generation can be interpreted as causal effects.

The simple reduced-form OLS regressions we estimate are of the following form:

9

$$\text{Outcome}_{ij} = \alpha + \beta_1 I_{ij} + \beta_2 V_j + \epsilon_{ij}. \tag{1}$$

Our outcome variables are education, income and social ties. The variables I and V are binary, and they indicate whether the individual was chosen to attend school and whether he lived in a village where a school was set up. If an individual child i was chosen to go to school, then both $I_i = 1$ and $V_i = 1$. If a child grew up in a village with school, but was not chosen to attend the school, then $I_i = 0$ and $V_i = 1$, and if a child grew up in a village with no school, then $I_i = 0$ and $V_i = 0$. The key coefficients are β_1 and β_2 which estimate the causal effect of individual and village-level treatment, respectively.

Table 2 presents the coefficients on individual and village-level treatment with education as the outcome variable. The results thus represent a manipulation check. As expected, the coefficient on individual-level treatment is positive and highly statistically significant. In the first column in Table 2, education is measured on a scale from 0 to 3, where 0 indicates no education, 1 indicates primary school only, 2 indicates secondary school only and 3 indicates university education. From Table 1 we know that most of the treated children have completed primary school and only about 10% have completed secondary school. Accordingly, the individual-level coefficient in column 2 of Table 2 is very close to 1, while the coefficient in column 3 is about 0.1.

Table 2: First-Generation Education Effects

	(1) Education	(2) Primary or more	(3) Secondary or more
Individual-level treatment	1.053***	0.955***	0.098**
	(0.060)	(0.033)	(0.032)
Village-level treatment	0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)
Observations	324	324	324

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Note: Dependent variables are indicated in the column header. Standard errors are clustered by common.

Looking at the effect of individual and village-level treatment on income we see that in the first generation *only* the individual-level treatment contributed to higher income, as shown in Table 3.

This result is very strong and intuitive – we can deduce that the students put their knowledge of the French language, their literacy and math skill and understanding of the colonial state and culture to good use. They were able to get better jobs, and secure better living standards for their families. For example, students were as much as 65 percent less likely to be farmers compared to those who were not chosen to go to school, or those who lived in a village without a school.[7] In contrast, the coefficients on the village-level treatment variable are all very close to zero and statistically insignificant. This indicates that for those living in villages with school but who did not get education, their income level was no different from the level of income of those living in villages with no school.

Table 3: First-Generation Income Effects

	(1) Farmer	(2) Water	(3) Electricity	(4) Vehicle	(5) Income scale
Individual-level treatment	-0.641***	0.112**	0.077**	0.294***	0.872***
	(0.109)	(0.040)	(0.020)	(0.026)	(0.188)
Village-level treatment	-0.060	0.055	0.018	-0.012	-0.004
	(0.107)	(0.052)	(0.011)	(0.020)	(0.180)
Observations	291	406	406	388	379

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Income scale is a factor score comprising a number of variables. Standard errors are clustered by common.

What is particularly interesting is that the uneducated who grew up in treated villages did learn some French and in general had better social ties than those in untreated villages. These results are shown in Table 4, and are evidence of within-village externalities from the introduction of a school. These positive externalities, however, were not generated through interaction between the uneducated villagers and the colonialists. Rather, Column 2 of Table 4 suggests that the uneducated did not have more friends among the white people than those from untreated villages. Hence, it appears that these uneducated villagers learned French through neighbors and friends. This is

---

[7]Since most first-generation students finished only elementary school, the marginal effect of an additional year of education is quite large. Primary school consisted typically of four years of education, and the effect of having finished primary school on the probability of being a farmer is -0.61, or a decrease in 61%. Assuming a linear effect of additional schooling, each year of education decreased the probability of being a farmer by 15%, or around one fifth of the likelihood of being a farmer in treatment 2.

corroborated by the result in column 4 of Table 4. Using the coordinates of all the settlements within our four sites with school, we calculated the distance between each individual's home (to the extent we could identify and verify its location during the relevant time after treatment) and the the location of the school.[8] We find that those closer to school had higher social networks, as measured by our factor scale suggesting that the effect indeed may run through the neighbors. The difference in social networks score between the untreated in villages with and without school (column 3) is statistically significant at the 5% level, suggesting a development of greater social activity and organization in the villages that had a school.

Table 4: First-Generation Social Networks Effects

	(1) French language	(2) White friends	(3) Network scale	(4) Network scale
Individual-level treatment	0.870***	0.373***	2.010***	1.999***
	(0.035)	(0.017)	(0.255)	(0.265)
Village-level treatment	0.072*	0.049	0.100**	
	(0.027)	(0.041)	(0.029)	
Distance from school				-1.102**
				(0.251)
Observations	406	355	252	238

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Network scale is a factor score comprising a number of variables. The last column includes only individuals under treatment 1 and treatment 2 condition. Standard errors are clustered by common.

These differences in social networks among the uneducated in villages with and without schools are already suggestive evidence that the introduction of schools may have long-lasting effects that go beyond those individuals who receive education. These positive externalities that education might generate are likely particularly important in a state of utter underdevelopment, as was the case in turn of the 20th century Dahomey.

---

[8]What we refer to as a "village" is in fact a group of interconnected smaller settlements – groups of homes. For example, in Zagnanado, treatment 1 and treatment 2 include 16 settlements: Agnangon, Assiadji, Assiangbome, Ayogo, Azehounholi, Dezonde, Doga, Dovi Dove, Gbenonkpo, Hougbodji, Kinbahoue, Kotyngon, Legbado, N'Dokpo, Sowe, and Zomon. We assign a location for each individual to a settlement, and calculate the distance from the location of the school. For Zagnanado, the school was closest to the settlement of Gbenonkpo and farthest from the settlement of Ayogo.

Part of the social network effect of education may run through higher political participation. Table 5 shows that students were significantly more likely to campaign for political parties, or even become full-fledged members. While very few people stood for election to political office in the period we cover in the first generation (only 12 people in our sample, or 3.22%), they are by and large concentrated among the treated individuals, allowing for quite precise estimate of the treatment effect, despite the low power.[9] These findings show a clear effect of education on political participation. To the best of our knowledge, the results are the first (quasi) experimental evidence in the support of the positive effect of education on political participation in developing countries.[10]

Table 5: First-Generation Political Participation Effects

	(1) Campaign for party	(2) Member of party	(3) Candidate in election
Individual-level treatment	0.339***	0.317***	0.117**
	(0.052)	(0.061)	(0.035)
Village-level treatment	0.045	0.057	-0.021**
	(0.053)	(0.069)	(0.007)
Observations	365	362	373

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by common.

The statistically significant results in the first generation of students are hardly a surprise, but they are important to document as a social phenomenon. Education has brought important change to the lives of the first students.[11] The bigger question is whether there were long-lasting effects of education on the descendants of the first students, and whether the differences between the descendants of the educated and the uneducated grow or diminish through generations. We turn to these questions next.

---

[9]The negative and statistically significant effect for the village-level effect is due to the fact that no individuals in treatment 2 ran for election, whereas two individuals in the control group did.

[10]See Berinsky and Lenz (2011), Campante and Chor (2011), Dee (2004), Glaeser et al. (2007), and Kam and Palmer (2008), among others.

[11]Note that in Tables 2, 3 and 4 we have no additional controls and the standard errors are clustered at the common-level. If decade and common fixed effects are included the estimated coefficients are very similar, but sample sized drop by about 25% due to missing information about the year of birth of some students. Results are also robust to controlling for the number of siblings.

4.3 Direct and Indirect Second-Generation Effects: Education, Income, and Social Networks

The second-generation effects of education are of paramount importance for human development and social mobility in Benin. If the introduction of education only affects the educated and their descendants, the country's development path may be quite different than if education also indirectly affects everyone who lives in a village with school. In this section, we will show in several ways that descendants of uneducated people in villages with school catch up with the descendants of the educated. In particular, they catch up both in terms of primary education outcomes as well as in terms of income and measures of social networks.

We begin our analysis of the second-generation effects by estimating regressions of the following type:

$$\text{Outcome}_{ij} = \alpha + \beta_1 I_{ij} + \beta_2 V_j + \beta_3 \mathbf{X_{ij}} + \mu_j + \tau + \epsilon_{ij}. \tag{2}$$

As before, our outcome variables are education, income and social ties, the binary variables I and V indicate individual-level and village-level treatment of the first-generation individuals, and β_1 and β_2 are our coefficients of primary interest. Now we also add a matrix of controls, $\mathbf{X}$, which includes gender and the number of siblings, we also include common fixed effects, μ, and decade fixed effects, τ.

Note that in the second generation, the binary variable I is equal to 1 for both children as well as nieces and nephews of former students. This coding was chosen because extended families were and still are a crucial social unit in African countries. Of course, there may be differences in the opportunities available to children and nieces and nephews of the original students as they grow up. However, for the moment we disregard these differences, and we return to them in depth in the next section.

Table 6 presents the second-generation results for education. The most striking finding from Table 6 is that the coefficient on village-level treatment, unlike in the first generation, is large and statistically significant. This indicates that descendants of the uneducated from villages with

Table 6: Second-Generation Education Effects

	(1) Education	(2) Primary or more	(3) Secondary or more	(4) University
Indivudal-level treatment	0.387***	0.135***	0.164***	0.088***
	(0.084)	(0.039)	(0.036)	(0.022)
Village-level treatment	0.566***	0.334***	0.173***	0.058***
	(0.062)	(0.033)	(0.028)	(0.013)
Observations	2396	2396	2396	2396

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, indicator for child or nephew/niece, number of siblings, and common and decade fixed-effects.

school have significantly more education than descendants of the uneducated from villages without school. This difference in education outcomes is substantively large, statistically significant at the 1% level, and it appears at all education levels – primary education, secondary education and university eduction.

Perhaps even more striking is the finding from columns 1 and 2, that the coefficient on village level treatment is greater than the coefficient on the individual-level treatment indicator. This means that simply growing up in village with school has a big positive effect on descendants education, while the additional positive effect of having an educated parent or uncle is smaller. Looking at the individual and village-level coefficients for secondary and university education (columns 3 and 4), both are still highly statistically significant, but now they are of comparable magnitude. This suggests that at the higher levels of education, the descendants of educated fathers or uncles are twice as likely to go to secondary school or university as descendants of uneducated parents from villages with school. For example, in the case of secondary education a descendant of uneducated parents from a village with school, *ceteris paribus*, has about a 17% chance of attending secondary school, while the chance that a descendant of an educated parent or uncle attends secondary school is 16 percentage points higher. These are sizable effects.

A similar pattern emerges for income among the second generation descendants, as shown in Table 7. We see that simply having been raised in a village with school has important positive

Table 7: Second-Generation Income Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Farmer	Water	Electricity	Television	Telephone	Car	Income scale
Indivudal-level treatment	-0.085*** (0.029)	0.111*** (0.038)	0.124** (0.049)	0.206*** (0.044)	0.218*** (0.046)	0.124*** (0.034)	0.424*** (0.091)
Village-level treatment	-0.300*** (0.033)	0.069** (0.030)	0.383*** (0.045)	0.277*** (0.034)	0.201*** (0.034)	0.031 (0.031)	0.550*** (0.064)
Observations	2263	2426	2426	2426	2425	2387	2087

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, indicator for child or nephew/niece, number of siblings, and common and decade fixed-effects.

effects on measures of income and living standards. For example, results from column 1 of Table 7 suggests that being born in a village with school reduces the descendants probability of being a farmer by about 30%, and having an educated father or uncle reduces the likelihood of being a farmer only by an additional 8.5%. Hence, while being a descendant of an educated person clearly matters, descendants of the uneducated in villages with school have caught up over the course of only one generation.

For most other measures of living standards, such as having running water in the house (column 2), having a television or a telephone[12] (columns 3 and 4) the individual and village level effects are of comparable magnitude. Also, when we look at the composite measure of income, generated by factor analysis, the individual and village-level effects are comparable. In the case of car ownership, however, descendants of educated parents or uncles have a greater additional likelihood of ownership.

The effect of village-level treatment on social networks is also very large, statistically significant and consistent across measures. In particular, when looking at knowledge of French we again see that just growing up in a village with school increases the likelihood that the descendent speaks French by about 33%, and the additional effect of being a descendent of an educated person is a further 16 percentage points. In the case of knowledge of English and having white friends, however,

---

[12]Note that in the first generation we did not report results for ownership of telephone or television set because in early 20th century neither the educated nor the uneducated had this equipment.

Table 8: Second-Generation Social Networks Effects

	(1) Speaks French	(2) Speaks English	(3) White friends	(4) Social networks scale
Indivudal-level treatment	0.159***	0.058***	0.057***	0.442***
	(0.039)	(0.016)	(0.022)	(0.084)
Village-level treatment	0.329***	0.014*	0.036***	0.461***
	(0.034)	(0.008)	(0.013)	(0.071)
Observations	2427	2426	1849	2331

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, indicator for child or nephew/niece, number of siblings, and common and decade fixed-effects.

the additional effect of being a descendant of an educated person is large which is reasonable since it requires interaction with people outside the traditional social milieu.

Overall, there is one very big difference in results across the first and second generation. In the first generation, only those who were randomly picked to attend schools reaped the benefits of education. In other words, in the first generation, only the individual level treatment variable produces positive and statistically significant effects on our two main outcomes of interest – education and income. The only discernible positive effect on the contemporaries of students who did not go to school is that they learned a bit more French and began to develop better social networks than those in village where no school was set up.

In contrast, in the second generation across all outcomes we see that just having grow up in a village with school positively affects education and income. That is, the village-level treatment effect is now consistently positive and statistically significant, in addition to the individual level treatment effect. We take this as an indication that the descendants of the uneducated are catching up, and catching up fast, especially in terms of income and social networks.

4.4 Family Tax: Do Nieces and Nephews Perform as Well as Children?

So far we have shown that in the first generation the educated have better outcomes than the uneducated, and that in the second generation the descendants of the educated have better out-

comes. Under "descendants" we included both the direct descendants (i.e. children of the original students) as well as the indirect descendants (i.e. nieces and nephews of the students). The natural question arises – do the children accrue higher benefits from their parent's education than nieces and nephews?

The answer to this question is given in Table 9. In this table, we compare the average outcomes of children of the students to the average outcomes of nieces and nephews of the students. Because this table uses only descendants of the original students the number of observations is 738. The regression we estimate is very similar to equation 2, except that the individual-level treatment is defined a bit differently – now $I_i = 1$ for children of the students, and $I_i = 0$ for nieces and nephews. Also we exclude the village-level indicator since there would be no variation as all the descendants would have $V_i = 1$.

Table 9: Outcomes for Children and Extended Family Descendants of the Students

	(1) Education	(2) Primary or more	(3) Secondary or more	(4) University
Catholic student child	0.155	0.071	0.087*	-0.003
	(0.098)	(0.044)	(0.050)	(0.040)
Observations	738	738	738	738

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, number of siblings, and common and decade fixed-effects.

Some readers may find it surprising that children of the students do not seem to be performing any better than their nieces and nephews. We find that this demonstrates the strength of extended family networks in Western Africa and the pressure on successful individuals to support their kin. It is true that the children of the former students tend to have more primary education than nieces and nephews, as the coefficient in column 2 of Table 9 is positive, but this difference is not statistically significant. Only for secondary school, do we see that children have better outcomes, and this result is only significant at the 10% level. At the university level, however the children are indeed no different from nieces and nephews, as indicated by the zero coefficient estimated in column 4 of table 9.

If we acknowledge the strength of extended family networks, we would expect that nieces and nephews of the former students, even though they were born to uneducated parents, do significantly better than descendants of uneducated parents who do not have any educated members in the extended family. This is confirmed in Table 10.[13]

Table 10: First-Generation Student Extended-Family Externalities

| | (1) | (2) | (3) | (4) |
	Education	Primary or more	Secondary or more	University
Catholic student niece/nephew	0.503***	0.205***	0.185***	0.114***
	(0.087)	(0.040)	(0.039)	(0.029)
Observations	2066	2066	2066	2066

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, number of siblings, and common and decade fixed-effects.

In Table 10 we see that across all education outcomes having just one educated person in the extended family makes a large difference to the outcomes of the nieces and nephews. These descendants have better education at all education levels than descendants (either children on nieces and nephews) in families where no one was educated. These effects are statistically significant and substantial – they are 20% more likely to have primary school education, 19% more likely to have secondary school education and 11% more likely to go to university.
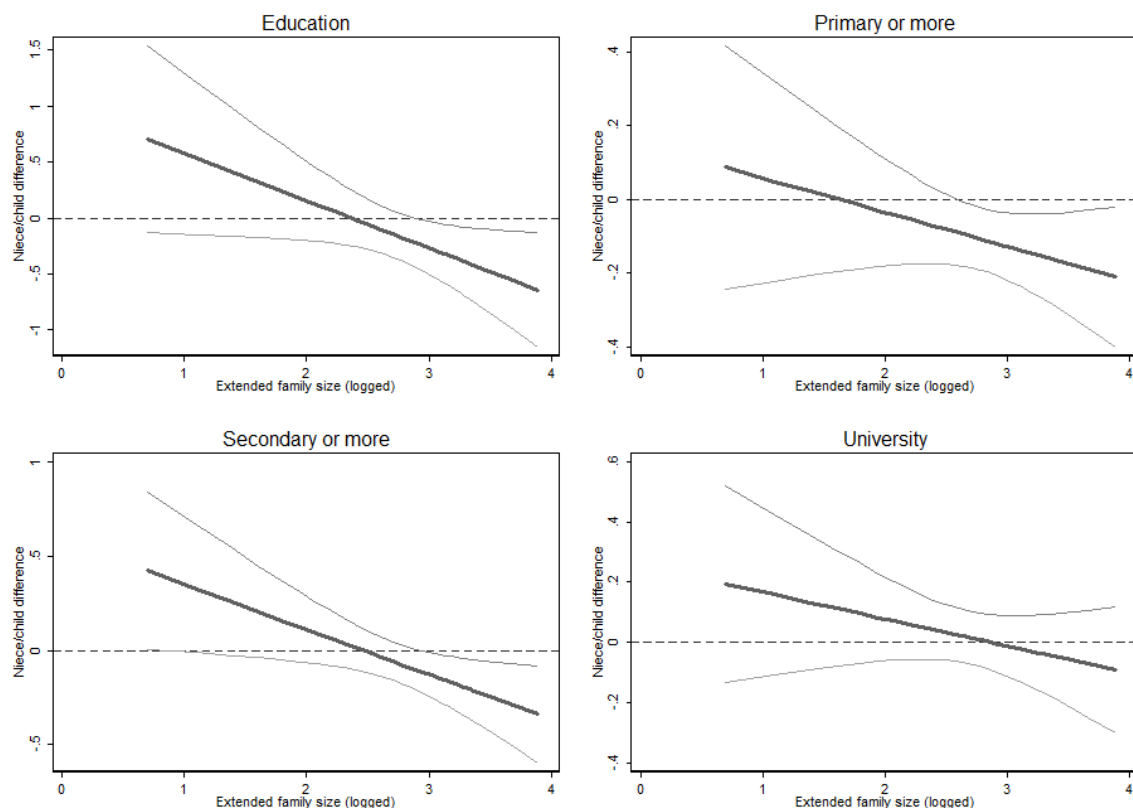
What may be happening is that educated uncles tend to support their nieces and nephews almost as much as their own children – we call this the *extended family tax* on education. One way to test this mechanism is to compare educational attainment of children and nieces and nephews in small and large extended families. If the family tax mechanism exists, we could imagine that as the extended family increases, the ability of the educated uncle to support all the nieces and nephews may be stretched too thin. In order words, the difference between nieces and nephews and children may be increasing as the size of the extended family size increases.[14] Results presented in Figure

---

[13]We confirm that children and nieces and nephews in treatment 2 and control do not have different outcomes – as they should not, given that none of the parents in their extended family had formal education. Results are available upon request, but they are indirectly presented in Table 10, since children and nieces/nephews from treatment 2 and control serve as the baseline in these results.

[14]An alternative plausible explanation might be that extended family externality runs through *aspirations*. The educated uncle may serve as a role model to both nieces and nephews and their parents. Similarly, nieces and nephews

1 seem to support this mechanism.

Figure 1: Education and Family Tax in Extended Families

Note: All models control for gender, number of siblings, parents' wealth, and common and decade fixed-effects. Marginal effects are calculated by keeping all remaining regressors at their means or medians. Gray lines represent the 95 percent confidence interval based on the standard errors clustered by extended family.

In Figure 1 we see that the difference in education outcomes between nieces and nephews and children becomes negative and statistically significant if the logged extended family size exceeds about 3 (i.e. the true extended family size exceeds about 20). Given such a large extended family, the educated uncle must prioritize between educating his own children and educating the extended family and the data suggest that at around this threshold level, education of own children becomes more important than education of the nieces and nephews.

may increase their educational attainment through emulation and learning from the children of the educated uncle. It is possible that as the extended family grows, ties to the educated uncles of any one niece and nephew become weaker, thus weakening the strength of aspirations and emulation. However, based on our knowledge of extended family networks in Benin, this is unlikely.

Note that our finding of an extended family tax is in discord with findings in the developed world that extended families are *not* altruistically linked (Altonji et al. 1992).[15] How does the existence of an extended family tax affect the human development of West Africa, and Benin in particular? Clearly, in the aggregate, there is a positive side of the family tax as it allows more promising children to get high levels of education, especially university education. However, there is also a negative side. As shown in Table 12 uneducated siblings of initial students choose to have more children than their uneducated counterparts in the same villages with school. Hence, these parents choose to have more children then they could raise independently. Educated parents, knowing that their siblings will expect support, may decide that exerting high effort to earn more may not be optimal given that they will have to give up an increasing amount of what they earn to their increasing extended family. With our analysis here, we only acknowledge the apparent existence of family tax. Currently, we cannot discern the magnitude of the positive and negative effects of family tax and we leave these challenges for future work.

4.5 Aspirations: A Determinant of Village-Level Externalities

In this section we aim to explain how the children of the uneducated in villages with school began to catch up with the children of the educated. We saw in section 4.3 that village-level treatment is associated with higher education, income and social network outcomes in the second generation. But, of course, there is variation in outcomes. In this section, we identify which uneducated families in villages with school are more likely to produce educated children. The channel that we focus on is that of higher aspirations developed to greater social ties with the educated locals and the white colonialists.

A problem in estimating the causal effect of parent's social network on his children's education is possible endogeneity. Parents with extensive social networks are more likely to be more ambitious and possess superior abilities which both lead to greater social networks and to greater aspirations

---

[15]Our findings are also related to the literature on sibling rivalry in developing countries. In Burkina Faso, Akresh et al. (2010) have found that if one child has higher IQ than his or her sibling, they receive a disproportionately large share of the families investment in education. In other words, a child is picked as the "hope of the family" and supported at the expense of less-abled siblings. Other papers that have found evidence of sibling rivalry in developing countries include Morduch (2000), Garg and Morduch (1998), Parish and Willis (1993) and Binder (1998), often in the context of allocation of resources across male and female children.

for their children's education, income, etc. One way to solve this problem is to instrument for parent's social networks with their distance from the location of the school where they could interact with the better educated locals and colonialists. This is what we do in Table 11.[16]

Table 11: Parents' Social Networks as Determinants of Descendants' Education

	(1) Education	(2) Primary or more	(3) Secondary or more	(4) University
Social networks scale	0.808**	0.406**	0.293**	0.109*
	(0.349)	(0.207)	(0.139)	(0.065)
First stage				
Distance from school	-3.623***	-3.623***	-3.623***	-3.623***
	(1.228)	(1.228)	(1.228)	(1.228)
Observations	499	499	499	499
F-statistic	8.698	8.698	8.698	8.698
p-value	0.004	0.004	0.004	0.004

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Social network scale is a factor score comprising a number of variables. Social network scale is that of the parents. It is instrumented by the distance of the parents' household from the school in the nearest village. Standard errors are clustered by extended family. All regressions control for gender, number of siblings, and common and decade fixed-effects.

The first stage in Table 11 is the same in each regression because we are using distance to instrument for parents' social networks. Next we use the variation in parents' social network that can be explained by distance and find that it is associated with higher education of their children, and this positive relationship is statistically significant at the 5% level.

What do these findings suggest? Our interpretation is that greater interaction with the educated locals and the colonialists, simply because they set up a school in the vicinity, increases the aspirations that parents have for their children. This then induces parents to invest more in their children's education and leads to better outcomes at the village level where schools were set up.

---

[16]In Table 11 we can only use information on parents and their own children, hence the number of observations is relatively low (499). We cannot use the full set of nephews and nieces because we do not have information about the social networks of the parents of nephews and nieces (we only have information about their one uncle).

5   Robustness Checks

5.1   Addressing Possible Bias due to Different Birthing Patterns

Despite finding significant differences between the descendants of the educated and the uneducated from villages with and without school in section 4.3, we must be careful to treat these results as suggestive, not causal, relationships. For causal interpretation, we need the individual-level and village-level assignment to be random. This condition is satisfied in the first generation, but not in the second. It is not satisfied in the second generation because parents **choose** how many children to have. In particular, the parent's treatment assignment may determine their chosen number of children – in Benin, more educated parents tend to be richer and to have more children. Moreover, siblings of initial student may also choose to have more children, because they count on extended family support to raise and educate these children. These patterns are documented in Table 12.

Table 12: Treatment Assignment and Family Size

	Treated parents	Untreated parents in village w/ school	Untreated parents in village w/o school
Children			
Average number	5.49	4.98	3.17
Difference from treated		-0.51	-2.32
p-value		0.20	0.00
Descendants			
Average number	5.88	4.02	2.88
Difference from treated		-1.86	-3.00
p-value		0.00	0.00

Note: Extended family descendants include all reported nieces, nephews and foster children. For very large families, our sampling design includes only a random subsample of all extended family descendants; this design should not affect the accuracy of the test reported. p-values are based on the Mann-Whitney-Wilcoxen difference of means test. Results are qualitatively equivalent if using the traditional two-groups difference of means t-test or the Kolmogorov-Smirnov test of the equality of distributions.

   How may this bias our results? When the treatment assignment affects the number of children and nieces and nephews born to the educated, we are faced with a selection problem. In Table 12 we see that the educated, on average, have half a child more than the uneducated in villages with school have, and, on average, 2.3 children more than the uneducated in villages without school.

Moreover, the educated also have more nieces and nephews – on average they have almost 2 and 3 more nieces and nephews than the uneducated in villages with and without school, respectively. These differences in the number of descendants are statistically significant and substantial in size given that the educated parent on average has 5.5 children and 5.9 nieces and nephews.

A good way to think about this problem is in terms of "principal strata" (Frangakis and Rubin 2002). Let the decision to have a child for a parent under certain treatment status be denoted C_{SB}, where $S \in \{c, t\}$ denotes individual-level treatment status (c means untreated, and t means treated), and $B \in \{u, b\}$ denotes the parenting decision (u stands for deciding not to have a child (unborn), and b stands for deciding to have a child (born)). Ignore for the moment the village-level treatment. Then, children can be divided into the following four strata:

1. *Always-born*: children who would be born irrespective of whether the parent is treated or not;

2. *Compliers*: children who would be born only if the parent is treated;

3. *Never-born*: children who would not be born irrespective of treatment assignment;

4. *Defiers*: children who would be born only if the parent is not treated.

Figure 3: Principal Strata

	Treatment (Educated)	Control (Uneducated)
Born	$C_{tb} = \{$Always born, Compliers$\}$	$C_{cb} = \{$Always born, Defiers$\}$
Not born	$C_{tu} = \{$Never born, Defiers$\}$	$C_{cu} = \{$Never born, Compliers$\}$

Figure 3 places the principal strata in a matrix to indicate what mix of populations we can observe among the descendants of the educated and the uneducated. The observed children of the treated parents, C_{tb}, will be a combination of the always-born and the compliers. The children of the treated that we do *not* observe (because they are not born, but who are theoretically defined), C_{tu}, are a combination of the never-born and the defiers. The observed children of the untreated

parents, C_{cb}, will be a combination of the always born and the defiers. Finally, the unobserved children of the untreated, C_{cu}, are a combination of the compliers and the never-born. Since we can only measure the outcomes of the born children, in our estimation, we focus on C_{tb} and C_{cb}.

The estimator in equation 2 makes a "naive" comparison of these two groups (C_{tb} and C_{cb}), i.e. it assumes that the underlying populations are the same. However, this is problematic as we are comparing outcomes in the set of always-born and compliers (C_{tb}) to outcomes in the set of always-born and defiers (C_{cb}) – these two groups do not represent the same populations and are therefore not directly comparable. We infer that the two groups do not represent the same population because of the evidence shown above – that treated parents have more kids than parents in control.

The problem in naive comparison arises if the compliers in the treatment group and the defiers in the control group have different potential outcomes than the always-born in the two groups. For example, if we assume that there are no defiers, and that the compliers (children only born when the parent is treated, and not born otherwise) have lower education on average than the always-born, then our naive estimator from equation 2 will *underestimate* the treatment effect. This is because the presence of compliers in the treatment group lowers the mean education of the treated group compared to the untreated group. Conversely, if either the compliers are expected to have better educational outcomes than the always-born, or if defiers (if present) are expected to have lower education than the always-born in the untreated group, or both, our estimator will *overestimate* the true treatment effect.

The literature that has dealt with this issue in other contexts recommends focusing only on the *always-takers* (Horowitz and Manski 2000, Lee 2009, Zhang and Rubin 2003, Zhang et al. 2009), as only for this subpopulation are the outcomes defined in both treatment conditions (other subpopulations, compliers and defiers, have undefined outcomes for at least one of the treatment conditions). In our case, that means trying to estimate the treatment effect only for the always-born, by comparing the always-born in the treatment group to the always-born in the control group. Of course, who exactly is an always-born, and who a complier or a defier is unobservable. These approaches therefore resort to calculating *bounds* on the treatment effect using various assumptions and possible worst-case and best-case scenarios.

One useful approach, that we apply here, has been developed by Lee (2009). The key assumption that we need for this approach is that monotonicity holds, i.e. that there are no defiers in the control group. This assumption allows us to assume that the only subpopulation in the control group are the always-born, i.e. those descendants who would be born irrespective of the education level of their parents or uncles. The treatment group is composed of always-takers – that we want to isolate – and the compliers (those who are born only because they have an educated uncle or father).

Is the monotonicity assumption reasonable in our case? We believe that it is. If this assumption were violated, then there exist people who have fewer kids if they are educated then if they had been uneducated. In the aftermath of the slave trade that decimated the local population over four centuries, people in 20th century Dahomey had as many children as they could afford. (see Manning [1995]). Hence, the educated would almost never have fewer children then the uneducated. In order to calculate the lower and upper bounds for our treatment effect, we need to focus on the compliers in the treatment group. To determine the share of compliers in the treatment group, we should take the difference between those who were born in the treatment group and those who were born in the control group (i.e. the difference between the always-born and compliers in the treatment group and the always-born in the control group), and express that as a share of the born individuals in the treatment group. Next, since we cannot identify exactly who these compliers are, just how many of them there are, we construct the best- and the worst-case scenarios, as in Lee (2009):

1. **Best case**: all compliers have the *lowest* education level among the treated who were born. The procedure is then to "trim" the low end of the distribution of education among the treated by the share of the compliers, and recalculate the mean education among treated and calculate the treatment effect with this mean (by subtracting the mean education of the control group). Since the low end of the distribution is trimmed, the new mean of the treated will be higher, and the new treatment effect will be higher. This is the *upper bound*.

2. **Worst case**: all compliers have the *highest* education level among the treated who were employed. The procedure is then to trim the high end of the distribution of education among the treated by the share of the compliers, and recalculate the treated mean and the treatment

effect. Obviously, now the treatment mean and the treatment effect will be lower, which gives the *lower bound.*

Table 13: Bounds on Treatment Effect for Children with Selective Birth

	Treatment	ATE	"Worst Case" Bound	"Best Case" Bound
Education				
	Individual-level	0.483 (0.108)***	0.400 (0.102)***	0.539 (0.105)***
	Village-level	0.427 (0.075)***	-0.115 (0.048)**	0.929 (0.076)***
Income scale				
	Individual-level	0.502 (0.116)***	0.416 (0.112)***	0.568 (0.118)***
	Village-level	0.569 (0.077)***	-0.069 (0.054)	1.093 (0.082)***
Networks scale				
	Individual-level	0.444 (0.121)***	0.293 (0.103)***	0.529 (0.118)***
	Village-level	0.231 (0.090)**	-0.260 (0.075)***	0.777 (0.087)***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: The main entries in cells are estimates from the regression of each dependent variable, indicated in the first column, on individual-level and village-level treatment. The entries in the parentheses are the standard errors, clustered by extended family. Bounds are obtained using the method of Lee (2009). No other controls are used, but the results (available upon request) are qualitatively similar when controls from the models reported in previous tables are included. The share of unborn children in any group (the "never-born"), needed for the trimming procedure, is not observed. It is assumed that the largest family within the wealthiest 50 percent in the treatment group had attained an ideal family size. The unborn are obtained for every other family by subtracting their number of children from the family with the largest number of children. Results are qualitatively similar when the share of the unborn children in each control group is alternatively calculated by taking the ratio of the average number of children in that group and the average number of children in the treatment group.

The calculated best and worst case bounds are presented in Table 13. Results in Table 13 are calculated only for children, i.e. nieces and nephews are excluded.[17] We see that individual-level effects are positive and even the lower and upper bound of the ATE are positive and statistically significant. This is true for all outcomes.

The worst and best-case bounds for the village-level effect are much wider, because the difference in the number of descendants in villages with and without school is quite large (see Table 12). The estimated lower bound for the village-level effect is typically just below zero, suggesting that in the worst case scenario, we cannot claim the existence of a village level effect. Yet, the worst case

---

[17]If nieces and nephews were included, then the ATE shown in this table would be the same as the ATE in column 1 of Table 6, column 7 of Table 7 and column 4 of Table 8. We exclude nieces and nephews because we do not know which nuclear family they belong to (i.e. how many brothers and sisters they have) which is necessary for the computations.

scenario – that compliers have higher potential outcomes than the always born – is pretty extreme and most likely a positive effect remains.

5.2   Addressing Possible Bias due to Non-Random Missingness

A natural concern is that our dataset fails to capture the less successful and prosperous individuals in the first generation, as well as their descendants. Since we have shown that success is correlated with education, this may imply that we are less likely to observe individuals in control groups than in treatment. Therefore, our comparisons may overestimate the returns to education. There are two ways in which this bias may arise. First, we may fail to observe any data on less successful individuals due to biased sampling. However...[HERE, DISCUSS SAMPLE SELECTION IN MORE DETAILS].

Second, conditional on sampling, we may fail to observe less successful individuals if they are more likely to have missing values for outcomes of interest. This may be a consequence of recall bias – our respondents may be more likely to remember the outcomes of the more successful relatives. There is some evidence of this in our data. For example, the rate of missingness on education is significantly lower among the treated first-generation individuals (8%) than those in control 1 and control 2 (27% and 20%, respectively). Since our estimates in the previous sections discard missing values, our estimates may be biased. We therefore perform several checks of the robustness of our findings to the potentially non-random patterns of missingness.

First, we perform a worst-case scenario exercise similar in logic to that in the previous section. We assume that a missing value on some variable of interest is due to the value of that particular variable, i.e. that missingness is non-ignorable (Rubin 1987). As we focus on the outcomes examined in the previous sections, we are assuming that missingness is caused by treatment status. We further assume that all missing values in treatment contain the lowest outcome, and all missing values in control groups contain the highest outcome. This is the worst-case scenario for our estimates: assigning the lowest (highest) outcome in treatment (control) will down-weight the effects of education shown above in proportion to the share of missing values in each treatment group.

Table 14 shows the first-generation results of this exercise.[18] While the results are somewhat weaker, even under the worst-case scenario the sign and the significance of most of our earlier results are entirely preserved. This is the case even for the outcomes for which missingness is relatively substantial, such as the farmer indicator, where almost 30% of observations are missing.
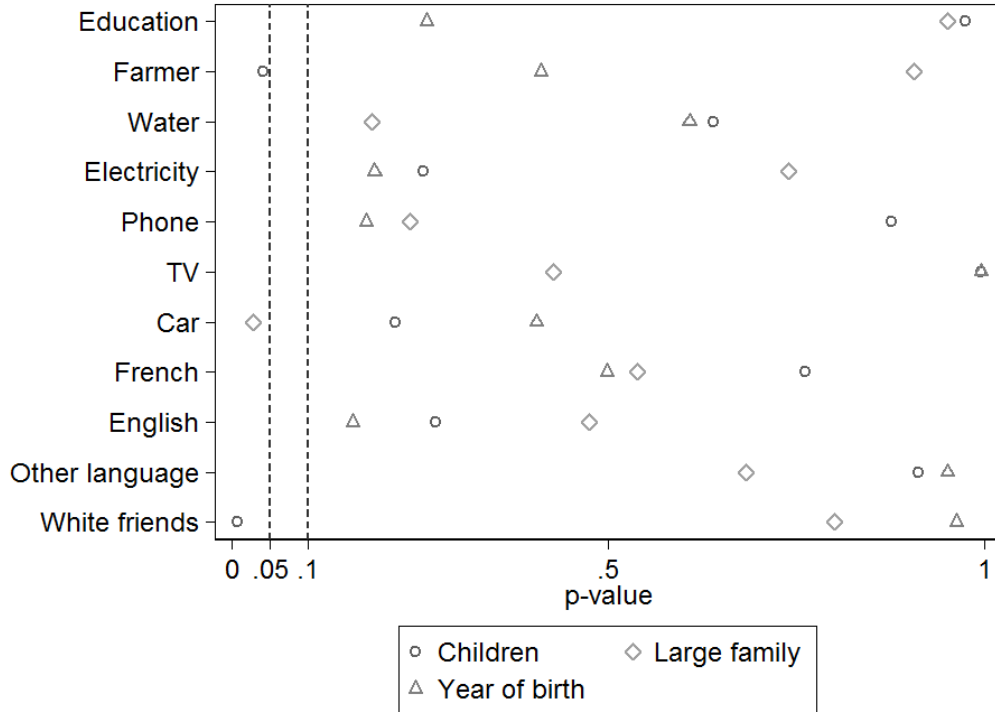
[Tables 14 and 15 about here.]

Table 15 shows the results of the same exercise for the descendants. Again, our results are mostly identical. Note for example that the worst-case scenario assumes that all missing data on education in the control group contain the achievement of university education, whereas all missing data in treatment indicate individuals without education. Nevertheless, the worst-case scenario estimates still point to significant positive effects of parents' education on descendants' outcomes.

Recall bias may not be entirely non-ignorable, i.e. missingness on our outcome data may be due to some other factors observable in the data. For example, our respondents may be more likely to recall outcomes for children than for nephews and nieces, for smaller families, or for individuals who were born later. Figure 4 examines the evidence for such possibilities. It plots the p-value from separate regressions of the missing value indicator for our outcome variables on: the dummy variable for children (circles), (b) the log of the number of siblings (diamonds), and the year of birth (triangles). The figure shows that there is no systematic evidence of recall bias based on any of the three plausible sources. In most regressions, the coefficient on all three variables of interest is not significant at the conventional levels, indicated by the vertical dashed lines.

We perform one more check of the evidence for recall bias. We have shown in section 4.4 that children do not have higher educational attainment than extended family descendants. Even though Figure 4 does not suggest that missingness is less likely among children, it may be that our respondents are less likely to recall outcomes for less successful nieces and nephews than less successful children. This would bias our results towards zero when comparing children and nieces/nepwhews. One way to examine the robustness to such recall bias is to compare only sons and nephews, as

---

[18]We exclude the factor scales for income and networks because we show the results of many of their individual components. Moreover, these scales are continuous, and it is less clear what value to assign. For example, assigning the minimum in treatment and the maximum in control represents an extremely conservative test.

Figure 4: Evidence for Recall Bias

male descendants were likely more successful on average than females.[19] In this subsample, the recall bias towards zero – if it exists – should be lower. Table 16 reruns models from Table 9 on the subsample of men. Our results are unchanged, further suggesting that recall bias is not an issue.

Table 16: Outcomes for Children and Extended Family Descendants of the Students – Males only

	(1)	(2)	(3)	(4)
	Education	Primary or more	Secondary or more	University
Catholic student child	0.034	0.033	0.031	-0.030
	(0.113)	(0.048)	(0.059)	(0.055)
Observations	452	449	449	449

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. Standard errors are clustered by extended family. All regressions control for gender, number of siblings, and common and decade fixed-effects.

---

[19]One reason to believe this is that we are more likely to observe nephews than nieces relative to the ratio of sons to daughters.

6   Conclusion

We use data from first elementary schools established in areas of Colonial Benin with no prior European influence to estimate economic and social effects of education. We find a large positive impact of education on measures of income, professional achievements and political participation. We also find significant peer effects and intergenerational income effects. The results provide rigorous estimates of human capital externalities and illustrates their impact on development.

One possible extension of the current paper includes a micro level investigation of the interaction between institutions and human capital by exploiting the timing of creation of schools and the setting up of colonial administration.

References

Acemoglu, D., S. Johnson, and J.A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *The American Economic Review*, 2001, *91* (5), 1369–1401.

Akresh, Richard, Emilie Bagby, Damien de Walque, and Harounan Kazianga, "Child Ability and Household Human Capital Investment Decisions in Burkina Faso," 2010. working paper.

Altonji, Joseph, Fumio Hayashi, and Laurence Kotlikoff, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," *American Economic Review*, 1992, *82(5)*, 1177–1198.

Angelucci, M., G. De Giorgi, M.A. Rangel, and I. Rasul, "Family networks and school enrolment: Evidence from a randomized social experiment," *Journal of Public Economics*, 2010, *94* (3-4), 197–221.

Berinsky, Adam and Gabriel Lenz, "Education and Political Participation: Exploring the Causal Link," *Political Behavior*, 2011, *33*, 357–373.

Binder, M., "Family background, gender and schooling in Mexico," *The Journal of Development Studies*, 1998, *35* (2), 54–71.

Bobonis, G.J. and F. Finan, "Neighborhood peer effects in secondary school enrollment decisions," *The Review of Economics and Statistics*, 2009, *91* (4), 695–716.

Borjas, George, "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 1992, *107(1)*, 123–150.

⸺ , "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review*, 1995, *85(3)*, 365–390.

Campante, Filipe R. and Davin Chor, "Schooling, Political Participation, and the Economy," *Review of Economics and Statistics*, 2011, *May 19 (Early Access)*.

Chiapa, Carlos, Jose Luis Garrido, and Silvia Prina, "The Effect of Social Programs and Exposure to Professionals on the Educational Aspirations of the Poor," 2010. working paper.

Cornevin, Rober, *La République populaire du Bénin. Des origines dahoméennes à nos jours*, Paris: Maisonneuve & Larose, 1981.

Cox, D. and M. Fafchamps, "Extended family and kinship networks: economic insights and evolutionary directions," *Handbook of Development Economics*, 2007, *4*, 3711–3784.

Dee, Thomas S., "Are There Civic Returns to Education?," *Journal of Public Economics*, 2004, *88*, 1697–1720.

Dupuis, Paul Henry, *Histoire de l'Eglise du Benin: Aube Nouvelle (Tome 2)*, Lyon, France: Societes des Missions Africaines, 1961.

Ferrara, E. La, "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana," *American economic review*, 2003, *93* (5), 1730–1751.

Frangakis, Constantine and Donald Rubin, "Principal Stratification in Causal Inference," *Biometrics*, 2002, *58*, 21–29.

Garcias, L., "Les mouvements de résistance au Dahomey (1914-1917)," *Cahier d'Etudes Africaines*, 1970, *37*, 144–178.

Garg, A. and J. Morduch, "Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana," *Journal of Population Economics*, 1998, *11* (4), 471–493.

Gennaioli, Nicolla, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human Capital and Regional Development," *Quarterly Journal of Economics*, 2012, *Forthcoming*.

Glaeser, Edward, Gicamo Ponzetto, and Andrei Shleifer, "Why Does Democracy Need Education?," *Journal of Economic Growth*, 2007, *12*, 77–99.

Gould, Eric, Victor Lavy, and Paserman Daniele, "Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes," *Reveiw of Economic Studies*, 2011, *78*, 938–973.

Horowitz, Joel L. and Charles F. Manski, "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," *Journal of the American Stastical Association*, 2000, *95*, 77–84.

Huillery, Elise, "History Matters: the Long Term Impact of Colonial Public Investments In French West Africa," *American Economic Journal - Applied Economics*, 2009, *1*, 176–215.

Kam, Cindy D. and Carl L. Palmer, "Reconsidering the Effects of Education on Political Participation," *Journal of Politics*, 2008, *70*, 612–631.

Lalive, Rafael and Alejandra Cattaneo, "Social Interactions and Schooling Decisions," *Reveiw of Economics and Statistics*, 2009, *91(3)*, 457–477.

Lee, David S., "Training, Wages and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Reveiw of Economic Studies*, 2009, *76*, 1071–1102.

Mercier, Paul, *Tradition, changement, histoire. Les Somba du Dahomey septentrional*, Paris: Anthropos, 1968.

Mookherjee, Dilip, Stefan Napel, and Debraj Ray, "Aspirations, Segregation and Occupational Choice," *Journal of the European Economic Association*, 2010, *8(1)*, 139–168.

Morduch, J., "Sibling rivalry in Africa," *The American Economic Review*, 2000, *90* (2), 405–409.

Nunn, Nathan, "Christians in Colonial Africa," 2009. working paper.

——— , "Religious conversion in Colonial Africa," *American Economic Review: Papers & Proceedings*, 2010, *100(May2010)*, 147–152.

Parish, W.L. and R.J. Willis, "Daughters, education, and family budgets: Taiwan experiences," *Journal of Human Resources*, 1993, *28(4)*, 863–898.

Ray, Debraj, *What Have We Learnt About Poverty*, Oxford University Press, 2006.

Rubin, Donald B., *Multiple Imputation for Non-Response in Surveys*, New York, NY: John Wiley and Sons, 1987.

Shavit, Yossi and Jennifer Pierce, "Sibship Size and Educational Attainment in Nuclear and Extended Families: Arabs and Jews in Israel," *American Sociolgical Review*, 1991, *56(3)*, 321–330.

Topa, G., "Social interactions, local spillovers and unemployment," *The Review of Economic Studies*, 2001, *68* (2), 261–295.

Topor, Helene D'Almeida, *Histoire Economique du Dahomey (1890-1920), Tome 1 et Tome 2*, Editions l'Harmattan, 1995.

Wantchekon, Leonard, "Mobilité Sociale des Premiers Elèves du Benin (1864-1922)," 2012. Base de données, IREEP. Cotonou (Benin).

Woodberry, R.D. and T.S. Shah, "The pioneering protestants," *Journal of Democracy*, 2004, *15* (2), 47–61.

Zhang, Junni L. and Donald B. Rubin, "Estimation of Causal Effects via Principal Stratification When Some Outcomes are Truncated by 'Death'," *Journal of Educational and Behavioral Statistics*, 2003, *28*, 353–368.

___ , ___ , **and Fabrizia Mealli**, "Likelihood-Based Analysis of Causal Effects of Job-Training Programs Using Principal Stratification," *Journal of the American Statistical Association*, 2009, *104*, 166–176.
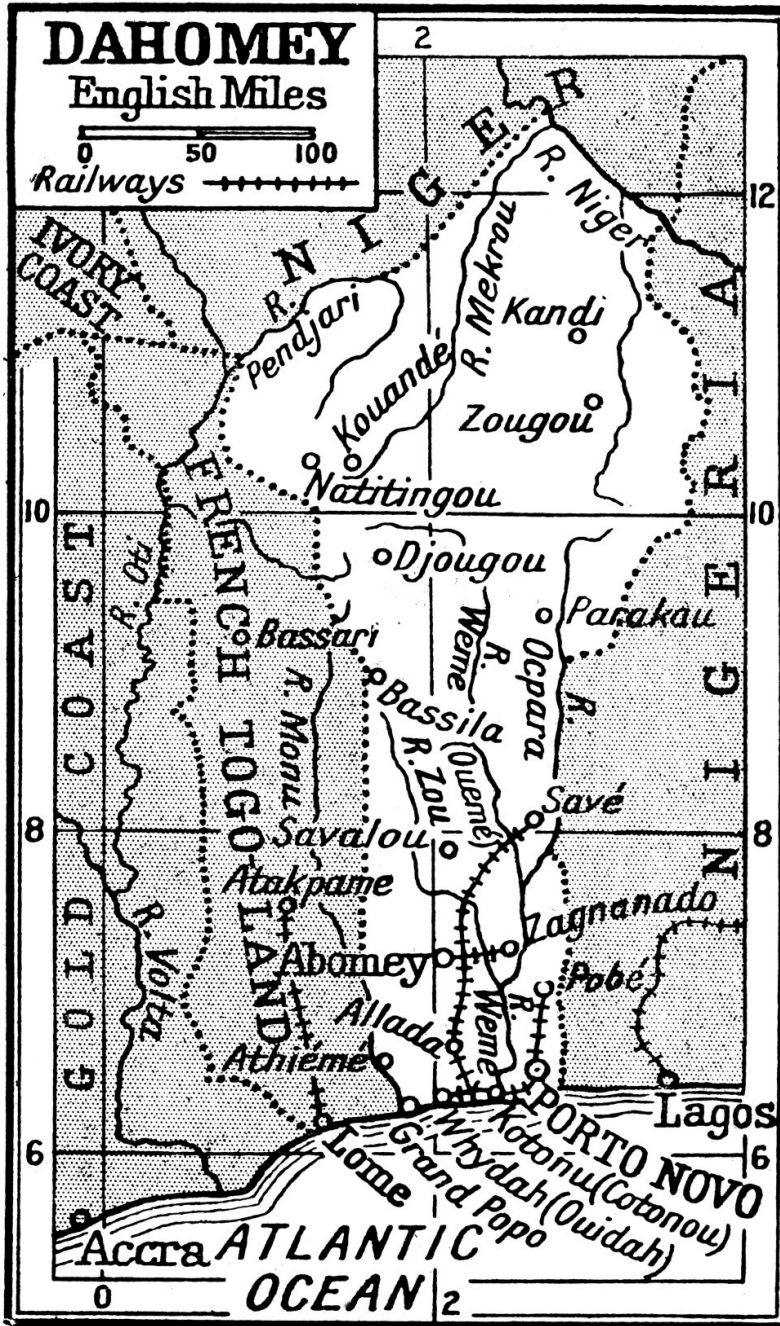
Figure 1: Map of Benin, 1938

Figure 2: List of the first students in the school in Natitingou (fragment)

Table 14: Robustness to Non-Ignorable Missingness – First Generation

	(1) Education	(2) Farmer	(3) Water	(4) Electricity	(5) Vehicle	(6) French language	(7) Other language	(8) White friends
Individual-level treatment	0.697**	-0.339*	0.112**	0.077**	0.218**	0.870***	0.061**	0.158*
	(0.170)	(0.141)	(0.040)	(0.020)	(0.057)	(0.035)	(0.014)	(0.059)
Village-level treatment	0.050	0.004	0.054	0.018	0.021	0.072*	0.006	0.096**
	(0.044)	(0.222)	(0.052)	(0.011)	(0.056)	(0.027)	(0.006)	(0.027)
Observations	405	405	405	405	405	405	405	405
Missing share	0.20	0.28	0.01	0.01	0.04	0.01	0.01	0.13

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. The worst-case scenario is explained in the text. Standard errors are clustered by common.

Table 15: Robustness to Non-Ignorable Missingness – Second Generation

	(1) Education	(2) Farmer	(3) Water	(4) Electricity	(5) TV	(6) Phone	(7) Car	(8) French	(9) English	(10) White friends
Indivudal-level treatment	0.301***	-0.052**	0.073	0.130***	0.193***	0.204***	0.061*	0.146***	0.043***	-0.227***
	(0.101)	(0.024)	(0.061)	(0.048)	(0.043)	(0.043)	(0.035)	(0.049)	(0.014)	(0.030)
Village-level treatment	0.451***	-0.211***	0.107*	0.377***	0.270***	0.198***	0.056*	0.277***	0.012*	0.169***
	(0.079)	(0.034)	(0.059)	(0.039)	(0.028)	(0.028)	(0.030)	(0.042)	(0.007)	(0.037)
Observations	3311	3311	3311	3311	3311	3311	3311	3311	3311	3311
Missing share	0.01	0.07	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.24

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Dependent variables are indicated in the column header. The worst-case scenario is explained in the text. Standard errors are clustered by extended family. All regressions control for gender, number of siblings, and common and decade fixed-effects.