The Effect of the Universal Primary Education Program on Labor Market Outcomes : Evidence from Tanzania

E. Delesalle

Discussion Paper 2019-10

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

UCLouvain

The Effect of the Universal Primary Education Program on Labor Market Outcomes: Evidence from Tanzania*

Esther DELESALLE †

April 24, 2019

Abstract

The purpose of this article is to study the effect of education on labor market participation and household consumption in a rural environment. The Tanzanian Universal Primary Education (UPE) program, which provides variations in education across locations and over time, is used as a natural experiment. Exploiting these two exogenous variations to instrument education, I find that education increases household consumption, especially in agriculture and in nonfarm self-employment activities. I also provide evidence that education increases the probability of working in agriculture. These results, initially surprising, suggest that returns to education in agriculture are positive, provided that the skills taught at school are suitable for agriculture.

Keywords: Human capital investment, returns to education, labor market organization, schooling reforms, Tanzania.

JEL Codes: I2, J24,015, 022.

*I wish to thank Pierre André, Paul Glewwe, Flore Gubert, Arnaud Lefranc, William Parienté, Christopher Udry, Gonzague Vannoorenberghe, Philippe De Vreyer for their useful comments, as well as the participants at the following conferences and seminars: NBER Transforming Rural Africa, CSAE, SMYE, CFDS, and IRES.

†IRES, Université catholique de Louvain, and DIAL. Email: esther.delesalle@uclouvain.be. Tel: +336.74.93.26.31

1 Introduction

Education is a cornerstone for economic growth; it eradicates poverty and counters the transmission of inequalities between generations. Given its importance, policy-makers put education at the top of their agenda. More specifically, several governments of developing countries have implemented policies to universalize primary education. These programs turned out to be useful in increasing the access to education, but not necessarily in improving the quality of education. A remaining challenge is to define policies that increase the level of education and insure high returns to education in a rural environment.

However, estimating the returns to education in a rural environment is not straightforward for two reasons. First, education is likely to be endogenous, which prevents obtaining unbiased estimates. Card (2001) reviews papers that aim to identify the causal impact of education on earnings. Among the eleven papers included in the survey, only two authors focus on developing countries: Duflo (2001), who instruments education by a school construction program in Indonesia, and Maluccio (1998), who instruments education by the distance to school in the rural Philippines.

However, both authors restrict their analysis to wage earners, which might affect the representativeness of the samples since wage-earning individuals are likely to be self-selected and to have specific characteristics. Maluccio (1998) does not deal with this sample selection issue, but Duflo (2001) adopts an imputation technique to compute a wage for individuals from the self-employment sector. While this method is suitable for countries with a developed formal sector, it is less adapted to countries that are mainly agriculture-based and where few individuals are wage earners.

Another strand of the literature estimates the returns to education among agricultural households by considering the agricultural production. Lockheed *et al.* (1980) review papers estimating the impact of education on agricultural production and find very mixed results depending on the country and the specification of education. However, these papers do not consider the endogeneity of education of the household head.

Using the Universal Primary Education (UPE) program implemented in Tanzania from 1974 to 1978 as a natural experiment, this paper contributes to this existing literature by investigating the benefits of education not only for wage-workers but also for the entire population.

Since developing countries are often characterized by the large size of both the nona-

gricultural self-employed sector and the agricultural sector, I construct a consumption aggregate to extend the analysis to all sample households. To account for the potential endogeneity of education, I instrument education of the household head by the exposure to the massive UPE program. In 1974, before the implementation of the UPE program, educational levels were low at the national level, with wide variation across regions. To reduce disparities in access to education, the Tanzanian socialist government gave priority to deprived areas, which led the latter to experience higher schooling expansion. The results of the strict enforcement of the UPE program were substantial: 3.3 million children aged 7 to 13 were enrolled in 1980, compared to 1.2 million in 1974 (Bonini, 2003). Thus, the UPE program provides an exogenous variation in education that I use to instrument education and to determine the effect of education on consumption. As returns to education might vary over sectors, I also distinguish the effect for subgroups: the agricultural sector, the nonfarm self-employed sector, and wage-work activities.

The second contribution of this paper is to address the effect of education on the labor market participation, more precisely, on the choice of the sector of activity. Indeed, education might not only increase earnings but also provide access to better paid activities and ease mobility between sectors. To address the endogeneity of education, I adopt the same identification strategy, and I instrument education by exploiting the exposure to the UPE program.

The main findings of this paper suggest that the UPE program reduced inequalities of access to education and that returns to education are positive in every sector. Counterintuitively, they are especially high in agriculture. I justify this finding by the design of the program, which was directed toward agriculture by providing a specific curriculum with agricultural classes. In this specific environment, this paper also demonstrates that education decreases the probability of working in nonfarm self-employed activities in favor of farm activities.

The remainder of the paper is organized as follows: section 2 provides a broad picture of the evolution of education in Tanzania and describes the data and the main variables of the analysis. Section 3.1 introduces the identification strategy; section 3.2 presents the effect of the UPE program on education; section 4.1 and section 4.2, respectively, focus on the effect of education on consumption and on labor market participation. Finally, section 6 concludes.

2 The program and the data

2.1 Data sets

This study uses three data sources: a census data set, a household panel survey, and administrative data. First, the census data are a 10 percent IPUMS sample from the 2002 Population and Housing Census in Tanzania. These exhaustive data, collected by the National Bureau of Statistics (NBS), are representative at the district level and contain basic information on dwelling characteristics, individual demographics and socioeconomics for 500,519 households. To complete the analysis and provide an accurate measure of households' consumption, I combine the LSMS-ISA (LSMS-Integrated Surveys on Agriculture) data and a household panel survey collected by the World Bank in 2008-2009, 2010-2011 and 2012-2013.[1] The LSMS-ISA data include 3265 households in 2008, 3924 households in 2010 and 5015 households in 2012.[2] This dataset gives detailed information on labor activities, on household consumption, and on other individual characteristics. Despite a district reorganization between the dates of the two datasets, both datasets cover the 26 Tanzanian regions. Finally, I use administrative data collected by the Ministry of Economic Affairs and Development Planning that are recorded in Jensen *et al.* (1968). These data comprise information on the distribution of primary schools and on GDP[3] by regions for mainland Tanzania in 1967, just before the introduction of the UPE program. These data are particularly interesting for investigating the effect of the UPE program because they constitute, to the best of my knowledge, the only source of information on primary school provision in Tanzania at this time.[4]

2.2 Historical background and the UPE program

When colonization ended in 1961, access to education in Tanzania was very unequal between regions (Court and Kinyanjui, 1980).[5] At this time, the purpose of primary

[1] From October 2008 to December 2009 for the first wave, from October 2010 to December 2011 for the second wave, and from October 2013 to December 2013 for the third wave.

[2] The number of households is increasing over the three waves due to the high number of split-off households and to the low attrition rate that does not exceed 5 % over the three rounds.

[3] GDP records are divided into subactivities, such as crops, livestock, mining, manufacturing, construction, public utilities, transport, rent, and other services.

[4] The National Bureau of Statistics gives access to the number of schools by region only from 2002.

[5] These spatial disparities were based on ecological endowments and were exacerbated by colonial activities and transport networks. The most privileged zones were the Arusha-Kilimanjaro-Tanga and Mwanza-Shinyanga corridors, as well as the Coast Morogoro-Kigoma (Maro and Mlay, 1979).

education was to prepare for secondary education and to encourage a small number of rural students to find white-collar jobs in urban areas (Kinunda, 1975). To mark a radical change with this elitist system, Prime Minister Nyerere, who came to power in 1964, fully redesigned the education system. With the Education for Self-Reliance (ESR) policy, approved in 1967, education became the mainstay of the Tanzanian socialist economy that would ensure economic growth. The aim of this policy was threefold: i) to improve the equity of access to education, ii) to teach agricultural skills that would be relevant in a rural society, and iii) to offer a political and civic education (Nyerere, 1967). This policy was supposed to lead to radical changes, but in practice, enforcement was slow. It was only in 1974 that the government committed itself to reach Universal Primary Education (UPE) with a forced march by 1978.

To achieve these goals, the Tanzanian government invested massively in primary education. Local resources and existing infrastructures were mobilized for classrooms, and a large number of schools were built. To provide access to schools in remote rural areas, the government proceeded to a villagization process, which consisted of gathering people in community villages commonly called ujamaa. This villagization started in 1968 on a voluntary basis, but from 1974, households living in remote areas were forced to move. As a result, more than 10 million people were moved, and 2,650 ujamaa were built from 1974 to 1977 (see Table 1). Although the distance to their prior dwelling was often less than five kilometers, villagization greatly reduced distances to schools.

To attain agriculture self-sufficiency defined as one of the main priorities: «kilimo cha kufa na kupona», *Agriculture for Life and Death*(Nyerere, 1967), agricultural classes were introduced in the curriculum. These classes were not necessarily taught by traditional teachers but also by farmers, who shared their experiences and their technical skills (Gillette, 1975). By the end of the reform, almost every school had access to a field in which children worked and experimented new harvest methods.

In addition, the starting age was postponed from 5 to 7 years old, and the examination in the middle of the primary cycle was removed. Consequently, pupils leaving the primary schools would be old enough and would have acquired the abilities to work in the fields. To accompany these changes and encourage people to start working after primary school, access to the secondary cycle was drastically limited by regional quotas (Martin, 1988).[6]

[6]Despite this policy, no significant drop in the secondary enrollment rate is observed.

Table 1: Villages in Tanzania

Year	Number of villages	Number of residents
1968	180	58 000
1969	650	300 000
1970	1200	50 000
1971	4484	1 595 240
1972	5556	1 980 862
1973	5631	2 028 164
1974	5008	2 560 474
1975	6944	9 140 229
1976	7658	13 067 220
1978	7768	13 847 000
1979	8200	13 905 000

Source: Shao (1982)

The government also made additional adjustments to improve schools' attractiveness. Tuition fees were eliminated, primary education became mandatory, and Swahili, most pupils' mother tongue, was designated as the language of instruction.

In terms of the education attainment, the results of this UPE program were considerable: from 1974 to 1978, the percentage of enrolled children aged 7 to 13 increased from 43.1 to 90.4 percent, and disparities among regions were drastically reduced (Bonini, 2003).

2.3 Measuring intensity of the UPE program

The UPE program was applied during a limited time frame and targeted areas with poor access to education. Hence, exposure to the program is expected to vary across locations and over time.

2.3.1 Over time

Since the official exit age to primary education was 13, individuals older than 13 years old at the beginning of the program (in 1974) should not have been affected by the program. However, several pilot programs were implemented in some regions from 1968. Thus, I define a pretreatment group T_0 to be household heads not affected by the UPE reform, consisting of individuals who were 13 or were older than 13 in 1967 (born between 1945 and 1954), and I distinguish T_{pt}, the group that consists of household heads who were

likely to be partially treated by the UPE program (born between 1945 and 1960). Then, I define T_{tot}, the treatment group, to include all children who should have been affected by the program. Restricting the data to the pretreatment cohort T_0 (individuals born between 1945 and 1954) and to the treatment cohort T_{tot} (individuals born between 1961 and 1978), I obtain from the census data two samples composed of 111, 818 and 388,701 individuals.

Table 2: Age Cohorts

Age cohorts	Year of birth	Age in 1974	Potential education level during the UPE plan	Obs. IPUMS	Obs. LSMS
T_0	1945-1954	20- 29	postsecondary and over	111,818	1,706
T_{pt}	1946-1960	14-19	secondary and postsec.	83,937	1324
T_{tot}	1961-1978	not born-13	no education-secondary	388,701	5,119

2.3.2 Across location

Since the program aimed at improving equity of access to education, the intensity should be a decreasing function of the schooling supply captured by $N_{j,67}$, the number of primary schools per square kilometer by region of birth j. Zanzibar West, which experienced no increase in years of schooling between 1967 and the end of the program in 1978, is considered as untreated, and I define the intensity index as:

$$I_{j,67} = (N_{Zanzibar\ West,67} - N_{j,67})$$

When $N_{j,67}$ is close to the schooling supply in Zanzibar West, the intensity of the treatment is expected to be small. Inversely, when $N_{j,67}$ is small, the intensity of the treatment is expected to be high.

As a robustness check, I provide another intensity index constructed from the education attainment by district of residence $I'_{d,67} = (S_{Zanzibar\ West,67} - S_{d,67})$. Although this index is likely to be less exogenous,[7] it is available at a lower geographical level.[8]

Figure 1a and 1b depict the education level before and after the introduction of the UPE program and confirm two predictions. The education attainment has been increased,

[7] It reflects the demand for education, and it is computed at the place of residence, which can be the result from endogenous migration decisions.

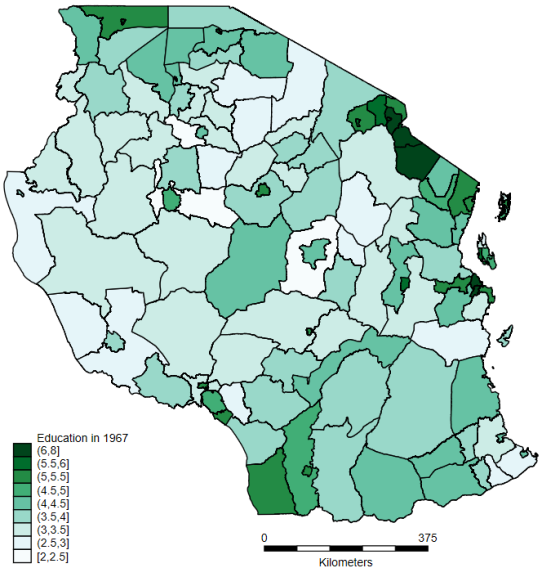
[8] All the results based on this intensity index can be found in the appendix.

and the education gap across location has been narrowed between these two periods.[9] Figure 1c, which illustrates the school distribution, shows that the schooling supply was also very unequal between regions in 1967 and that the supply of schools was correlated with the education attainment at this time.

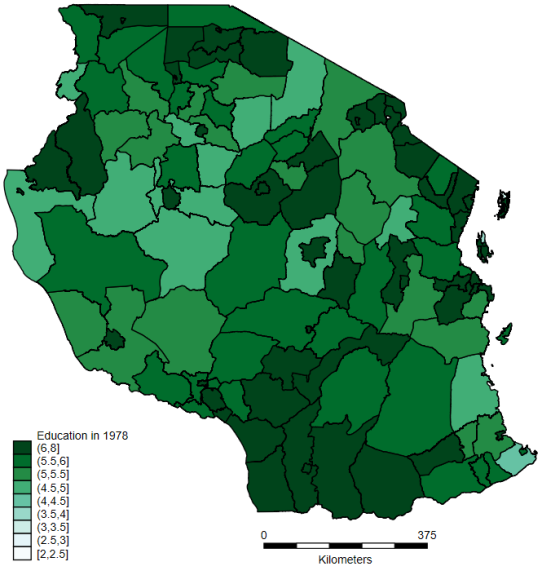
[9]In 1967, Zanzibar West and Kilimanjaro had already reached the maximum years of primary education, while the average education level in other regions had not exceeded two years of education.

Figure 1: Access to education in Tanzania

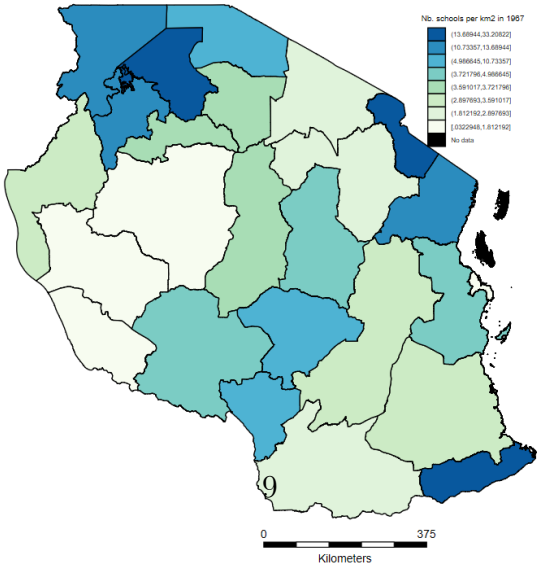
(a) Education level by district in 1967.



(b) Education level by district in 1978.



(c) Distribution of primary schools by region in 1967



2.3.3 Measuring household consumption

As the purpose of this article is to estimate the returns to education, it is necessary to first describe how the households' wealth is measured.

Usually, living standards are measured either by income or by consumption. In developing countries where agriculture is widespread, incomes are very sensitive to current shocks and may not be representative of household well-being (Meyer and Sullivan, 2003), while consumption can be smooth through formal or informal mechanisms. In this respect, consumption has the advantage of being more representative of long-run well-being. The second interest of using consumption stems from the fact that income is not similarly measured between activities,[10] which calls into question the reliability of comparison between sectors. Last, but not least, consumption is available for all households, which allows avoiding selection and imputation issues. Thus, these features advocate the use of consumption rather than income data in developing countries.

Deaton and Zaidi (2002) propose guidelines to construct a consumption variable from rich household survey data.[11] The Living Standard Measurement Study (LSMS) data are particularly well suited for constructing the consumption index since they collect exhaustive information on consumption expenditures. However, the serious limitation of the Deaton and Zaidi (2002) method is that such accurate data are costly to collect and are often not included in large datasets, such as the 2002 Tanzanian census.

Thus, to take advantage of the large sample size of the census data and obtain a monetary value of consumption that eases comparison with the literature (Duflo (2001), Maluccio (1998)), I follow the method developed by Elbers *et al.* (2003) and Tarozzi and Deaton (2009) from census data and household survey matching to construct a consumption proxy (see appendix A.1 for more details).

2.3.4 Descriptive statistics

Table 3 reports descriptive statistics from the 2002 census for the whole sample and for the subgroups of interest: the pretreatment T_0 (old cohort), the treatment group

[10] Self-employment income is rarely a wage, and agricultural income is measured through production.

[11] They consist of defining a weighted per capita consumption variable composed of four components: food items, nonfood items, housing consumption and consumption from consumer durables. To adjust household consumption for variation in household composition, the consumption variable is divided by an equivalence scale made from the household's size: every adult represents one unit of consumption, and each child represents 0.3 units.

T_{tot} (young cohort), "regions +" and "regions -" that gather regions where the education level in 1967 was above and below the national average, respectively. The data indicate

Table 3: Descriptive statistics from the 2002 census

	(1) All	(2) T_0	(3) T_{Tot}	(4) T_0-T_{Tot}	(5) Region$^-$	(6) Region$^+$	(7) Region$^-$-Region$^+$
Age	41,699	51,948	32,113	20,64***	41,806	41,498	0,000271
	(16,252)	(2,800)	(5,051)	(0,0124)	(16,395)	(15,977)	(0,0205)
Men	0,663	0,647	0,692	0,00385**	0,660	0,667	-0,000619
	(0,473)	(0,478)	(0,462)	(0,00130)	(0,474)	(0,471)	(0,000548)
Urban areas	0,425	0,385	0,461	-0,0676***	0,409	0,455	-0,0373***
	(0,494)	(0,487)	(0,498)	(0,00129)	(0,492)	(0,498)	(0,000536)
Years of primary edu,	4,484	3,403	5,612	-2,424***	4,215	4,991	-0,616***
	(3,071)	(2,987)	(2,602)	(0,00732)	(3,147)	(2,854)	(0,00366)
Ended primary edu	0,542	0,321	0,742	-0,439***	0,506	0,610	-0,0729***
	(0,498)	(0,467)	(0,437)	(0,00118)	(0,500)	(0,488)	(0,000582)
Man doesn't work	0,094	0,067	0,062	0,0156***	0,102	0,080	0,0282***
	(0,292)	(0,250)	(0,241)	(0,000802)	(0,303)	(0,271)	(0,000371)
Man works in agri	0,583	0,633	0,560	0,0799***	0,619	0,515	0,0842***
	(0,493)	(0,482)	(0,496)	(0,00148)	(0,486)	(0,500)	(0,000630)
Man is self-emp,	0,177	0,119	0,219	-0,0817***	0,156	0,216	-0,0473***
	(0,382)	(0,323)	(0,414)	(0,00115)	(0,363)	(0,412)	(0,000474)
Man is a wage worker	0,146	0,182	0,159	-0,0138***	0,123	0,190	-0,0652***
	(0,353)	(0,386)	(0,366)	(0,00110)	(0,328)	(0,392)	(0,000453)
Woman doesn't work	0,258	0,202	0,257	-0,0514***	0,255	0,264	0,00241***
	(0,438)	(0,402)	(0,437)	(0,00120)	(0,436)	(0,441)	(0,000503)
Woman works in agri	0,595	0,665	0,567	0,0874***	0,617	0,553	0,0428***
	(0,491)	(0,472)	(0,495)	(0,00135)	(0,486)	(0,497)	(0,000564)
Woman is self-emp,	0,096	0,076	0,117	-0,0319***	0,088	0,110	-0,0175***
	(0,294)	(0,265)	(0,321)	(0,000814)	(0,283)	(0,313)	(0,000331)
Woman is a wage worker	0,051	0,057	0,058	-0,00412***	0,040	0,073	-0,0277***
	(0,221)	(0,231)	(0,234)	(0,000620)	(0,196)	(0,261)	(0,000250)
log ($\widehat{consumption}$)	14,050	14,135	14,083	-0,0530***	13,965	14,211	-0,206***
	(0,720)	(0,771)	(0,680)	(0,00201)	(0,690)	(0,747)	(0,000852)
$\widehat{consumption}$	1702123,000	1961921,000	1688983,000	-45370,1***	1535861,000	2014851,000	-496063,2***
	(1872784)	(2456112)	(1624089)	(8069,4)	(1699485)	(2126491)	(3284,7)
GDP in 1967					231,015	318,822	-91,44***
					(98,673)	(243,845)	(0,199)
Observations	3676116	59326	221677	1026701	529033	280666	3676116

Sources: The 2002 census (IPUMS data). ***,**,* means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. Standard deviations are reported in parentheses columns (1) to (3), and (5) to (6). Standard error for average difference are reported in parentheses columns (4) and (7).

that the majority of households are head by men and live in rural areas, although the prevalence of rural households decreases overtime. A significant share of household heads has validated less than five years of education, meaning that some of them did not finish the primary education cycle. However, the comparison of the young and the old cohorts informs that the education level has been rising: the number of years has increased by 65%, and the percentage of households' heads who ended the primary education cycle has

more than doubled. Further, I examine whether "regions +" and "regions -" significantly differ and have characteristics that could explain different trends. The data highlight that regions with a low initial education level are more rural, more agricultural intensive and are poorer. These results are not surprising since the density of schools in rural areas is expected to be lower.

Table A2 assimilates similar descriptive statistics, but from the LSMS panel data. Overall, the statistics confirm what was described above, and inform that a nonnegligible share of households diversify their sources of income by cumulating different activities.

Although education has been increasing between age cohorts, it does not imply that the UPE program had a causal effect on the education expansion. The following section presents the identification strategy that I adopt to estimate the effect of the UPE program and to measure the returns to education.

3 Empirical strategy

3.1 Identification strategy

The instrumental variable approach that I use exploits the exposure to the UPE program. This assumes that being exposed to the program increases the probability of being enrolled in school but is orthogonal to unobserved household characteristics that determine labor market outcomes. To capture exposure to the UPE program, I adopt a difference-in-differences strategy based on variations in time and in space. It consists of comparing pretreatment cohorts ($T = 0$) with treated cohorts ($T = 1$) for whom the intensity of the program varied across areas and instrumenting education by the interaction term $T * I_{j,67}$, which captures the UPE program's exposure.

This variable is a valid instrument (IV) if two conditions are satisfied: i) the IV is correlated with education, and ii) the IV explains the outcome of interest only through education. In such cases, IV estimates correspond to the local average treatment effect (LATE). Otherwise, IV estimates give inconsistent and biased results. Since the interpretation of IV estimates relies on the quality of the instrument, I now discuss whether the interaction term $T * I_{j,67}$ is a valid instrument.

The IV variable is a relevant candidate if it is highly correlated with the endogenous variable. To check whether the instrument explains the education expansion, I plot in

figure A2 the education increase between the pretreatment and the treatment cohorts as a function of $N_{j,67}$, the number of schools per square kilometers in 1967.[12] The correlation between $N_{j,67}$ and the education increase shows that indeed, the UPE program was more intense in regions with a poor schooling supply. Likewise, the same conclusion can be drawn from figure A3, in which I depict the education increase as a function of the educational attainment by district of residence $S_{d,67}$. However, this relationship is not necessarily causal. The main concern is that the education expansion is not exclusively due to the UPE program but to other factors correlated with the instrument and the outcome of interest. Despite the fact that the exclusion restriction could not be tested, I try to identify all potential sources that could discredit this condition, and I provide evidence that the instrument is exogenous.

First, I check whether the education expansion is due to the introduction of the UPE program and not to a convergence phenomenon. In case of convergence, less educated regions could have had a higher education increase in order to catch up with the more educated regions. If this were to be true, this phenomenon would be observed before and after the introduction of the UPE program. Subsection 3.2 addresses this question and confirms that during the pretreatment program, the education progression was not statistically different between educated and noneducated regions. In contrast, the education expansion was statistically higher for deprived regions during the UPE program period.

Second, the exclusion restriction can still be invalidated if other region's characteristics generate the same trend reversal or are correlated with the outcomes of interest. To insure the exogeneity of the instrument, I add a set of controls. Among these control variables, I add the number of children aged 7 to 13 to account for the possibility that the education expansion may depend on the size of the cohort and the level of wealth, which might influence the development of the schooling supply.[13]

Furthermore, the sectoral specialization may be another source of bias if the regions develop different sensitivities to shocks. One way to ensure the validity of the instrument is to control for heterogeneity in order to capture variations in shocks between regions. In this respect, I add regional GDP by sector of activity interacted with a time trend. Among these sectors, I distinguish between the following economic activities: crops, livestock, mining, manufacturing, construction, and activities from the tertiary sector, including

[12]Each dot depicts a region of birth.

[13]Wealthy regions can have higher needs in skilled labor and invest more in education.

public utilities, transport, rent, and other public services.

In addition, De Chaisemartin and d Haultfoeuille (2015) highlight that IV estimates can be far from returns to education in any location when the homogeneity assumption does not hold. However, the authors show that difference-in-difference (DID) methods with fuzzy treated groups[14] should provide unbiased estimates without relying on any homogeneity assumption, as long as 1) the common trend assumption is valid and 2) there is a control group for which the treatment does not change over time. According to the above results, these two assumptions appear to be satisfied. In this study, 2) implies there is at least one region where education has not evolved between the pretreatment period and the treatment period, which is precisely the case of Zanzibar West.[15] This lack of education increase may be explained by the fact that education had already reached the maximum years of primary education in 1967[16] and that access to secondary education was cut at the time of the UPE program.

Last but not least, IV estimates are biased if the program has influenced outcomes other than education that explain the level of consumption. Regarding the forced villagization process, this assumption is very likely to be unsatisfied. Among the possible channels, the program could have changed the access to other social services and the living conditions. Nonetheless, I argue that this should not call into question the validity of the instrument because both the pretreatment and the treatment cohorts were similarly affected by these changes. Indeed, the specificity of the villagization program is that entire families were asked to move. In contrast, the education reform that was part of the program was beneficial only for the treatment cohorts and had no reason to affect outcomes other than education. This key argument that supports the validity of the instrument will be further discussed in the discussion section.

[14]This refers to DID when the intensity of the treatment varies between treated groups.
[15]The education level decreases by 0.1 year between 1967 and 1978, which is negligible.
[16]Zanzibar, independent in 1964, benefited from a better access to education.

3.2 The effect of the UPE program

3.2.1 On education expansion

Since education may be endogenous, I adopt a two-stage procedure, the first stage of which is:

$$S_{ijt} = \alpha + \beta_j + \beta_t + \gamma T * I_{j,1967} + \delta t * X_{j,1967} + \mu_{ijt} \quad (1)$$

β_j and β_t are region-of-birth fixed effects and birth-cohort fixed effects to account for permanent differences across regions and over time, and $X_{j,1967}$ is a set of region characteristics, including the log of population aged 7 to 13 and regional GDP by sectors of activity in 1967. Each of these controls is interacted with a time variable t. T is a dummy taking the value 0 for people belonging to the pretreatment group and 1 for people belonging to T_{tot}. The coefficient of interest, γ, represents the effect of the UPE program on education (years of schooling). The higher the intensity of the UPE program, $I_{j,1967}$, the larger should be the education expansion between the pretreatment and the treatment groups.

Table 4: Effect of the UPE program on education: γ coefficient of equation (1)

IV	(1)	(2)	(3)	(4)
	Years of education		Education completion	
$T_{tot} * I_{j,1967}$	0.052***	0.045***	0.006***	0.006***
	(0.006)	(0.010)	(0.001)	(0.001)
R-squared	0.271	0.272	0.238	0.239
F-test	69.76	21.26	31.33	19.34
Cohort FE	yes	yes	yes	yes
Location FE	yes	yes	yes	yes
GDP Controls		yes		yes
Observations	433,606	433,606	435,332	435,332

Source: the 2002 census. Notes: Standard errors are clustered at the location level and are reported in parentheses. ***,**,* mean respectively that the coefficient are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household size and the principal sector of activity of the household head.

Table 4 reports the results of equation (1). To consider the possible serial correlation in errors, I cluster standard errors at the regional level (Bertrand *et al.*, 2004). I find that when the predicted intensity $I_{j,1967}$ is raised by one additional school per square kilometer, education increases by 0.05 between T_0 and T_{tot} and that the introduction

of GDP controls marginally lower the coefficients. This result is consistent with the idea that the UPE program targeted regions with low initial education attainment and contributed to the equalization of access to education among regions. Columns (3) and (4) indicate whether the UPE plan had fully reached its goal by convincing people not only to enroll in school but also to complete primary education. Although the UPE program was strictly implemented for a shorter period than the duration of the primary cycle (four and seven years, respectively), the UPE program has significantly increased the education completion.

I also estimate a more flexible regression that allows the effect of the UPE program to vary with the time exposure to the program:

$$S_{ijt} = \alpha + \beta_j + \beta_t + \sum_{t=1945}^{1954} \gamma_t I_{j1967} + \sum_{t=1961}^{1978} \gamma_t I_{j1967} + \delta_t X_{j1967} + \mu_{ijt} \qquad (2)$$

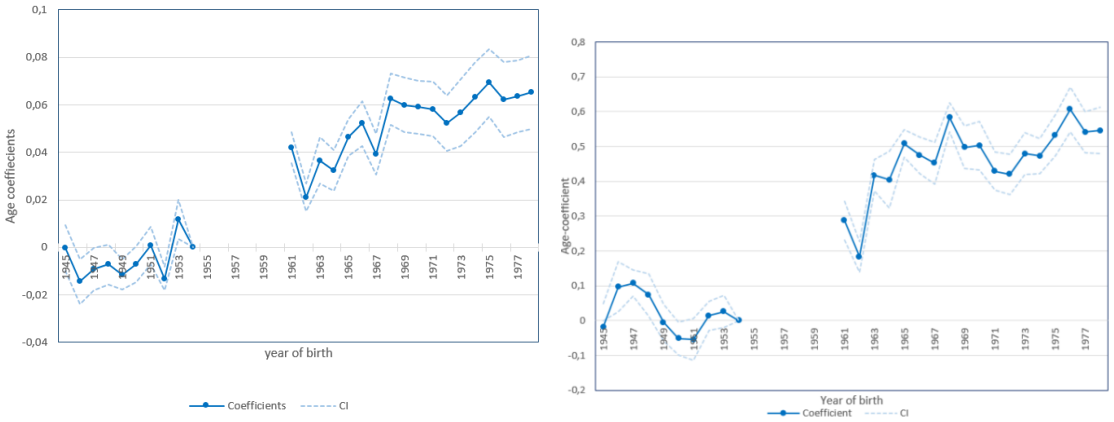
In this equation, γ_t, captures the effect of the UPE program on education by the age cohort, and the difference between γ_t and γ_{t+1} represents the education expansion between t and t+1 generated by the education supply in 1967.

For the pretreatment group, $I_{j,67}$ should have no impact on education expansion and γ_t values should be close to 0. In contrast, one expects the education expansion to be an increasing function of the time exposure to the UPE program for the treated cohorts. This is precisely what is shown in figure 2. Each dot depicts the γ_t coefficients of equation (2), from I_{j1967} at the left-hand side of the panel and from I'_{d1967} at the right-hand side of the panel.[17] For both intensity indexes, almost all the coefficients in the pretreatment group were not statistically different from 0, while γ_t coefficients steadily increased for the treated age cohorts. Cohorts born after 1968 were still exposed, but the slope declines afterwards. This graph confirms that the identification strategy is reasonable: the trend was not present before the program, and the UPE program had a significant impact on education for the treated cohorts (all coefficients are significant at the 1% level).

Thereafter, I instrument education by relying on equation (2), but by imposing each γ_t to be 0 for the pretreatment cohorts.

[17]The reference year is the year before the introduction of the UPE program in 1967, which corresponds to the age cohort born in 1954.

Figure 2: γ_t coefficients of equation 2



(a) From I_{j1967} (b) From I_{d1967}

Source: 2002 census.

$$S_{ijt} = \alpha + \beta_j + \beta_t + \sum_{t=1961}^{1978} \gamma_t I_{j1967} + \delta_t X_j + \mu_{ijt} \tag{3}$$

In this equation, γ_t identifies the effect of the UPE program by age cohort in comparison with the preprogram period T_0. If no regional time-varying characteristics correlated with the program's intensity are omitted, these fuzzy difference-in-differences correctly estimate the impact of the UPE program (results are presented in table A5).

Although I cannot identify the compliers of the UPE program, I can still compare the characteristics of individuals who completed and who did not complete primary education among the treatment age cohorts. Table A3 shows that individuals who completed primary education have a higher consumption level, are more likely to be a wage worker, to live in urban areas, and to be a man.

3.3 On consumption

Instead of looking at the effect of education, I exploit equation (2) to estimate the reduced form between the UPE program and the logarithm of the consumption proxy. Graphic A4 shows that the age coefficients for the pretreatment cohorts are almost never statistically different from 0. By contrast, all the coefficients for the treated age cohorts are positive and significant at the 1% level. Although the size of the coefficient differs, I deduce that the effect of the UPE program on education and on consumption follows the same trend and becomes positive and significant for treated age cohorts.

4 The results

This section presents the main results. The first subsection is devoted to the returns to education for the entire population, and by sectors of activity. Education may also have the benefit of increasing the probability to work in sectors that are better paid. Then, subsection 4.2 investigates whether education changes the labor distribution between the sectors of activity.

4.1 The returns to education

4.1.1 For the entire population

I measure the returns to education by looking at the effect of education S_{ijt} of household head i born in region j at year t on current consumption C_{ijt}. The main equation is:

$$Log(C_{ijt}) = \alpha + \beta_j + \beta_t + \theta S_{ijt} + \delta_t X_j + \varepsilon_{ijt} \tag{4}$$

where β_j and β_t are, respectively, region-of-birth and year-of-birth fixed effects. Regional controls X_j are also included and interacted with a time trend.

I first ignore the potential endogeneity of education and run OLS regressions by using the household consumption computed in the LSMS data with the Deaton and Zaidi (2002) method. Columns (1) and (2) in Table 5 indicate that returns to education are approximately 7%. Comparing columns (2) and (4) shows that the use of the consumption proxy $\widehat{LogC_{ijt}}$ instead of the consumption significantly lowers the returns to education, probably due to the potential downward bias described in appendix A.2. Notwithstanding, the estimates from $\widehat{LogC_{ijt}}$ are very similar between the LSMS data and the census data and are approximately 4%.

To obtain consistent estimates of θ, I instrument education by exploiting the census data.[18]

Table 6 reports the 2-SLS estimates of the effect of education on the consumption proxy. When I add controls for GDP by sectors of activity, I find that one additional year of education of the household head increases the log of household consumption between 7.1 and 9.2%. F-statistics are high, which suggests that the instruments have strong

[18]These first-stage equations are also estimated with the LSMS data, but sample sizes of subsamples are too small and prevent capturing any significant effect (see Table A4).

Table 5: OLS estimates of the returns to education

	$log(C)$		$\widehat{log(C)}$		$\widehat{log(C)}$	
	(1)	(2)	(3)	(4)	(5)	(6)
	0.071***	0.071***	0.044***	0.044***	0.043***	0.042***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.000)	(0.000)
R-squared	0.424	0.425	0.450	0.451	0.579	0.580
Data set	LSMS		LSMS		Census	
Cohort FE	yes	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes
GDP Control		yes		yes		yes
Observations	4,983	4,983	4,983	4,983	430,490	430,490

Notes: Standard errors are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household size and thesector of activity of the household head.

Table 6: IV estimates of the returns to education

	$T * N_{j,1967}$		$\sum_{t=1961}^{1978} \gamma_t * N_{j,1967}$	
	(1)	(2)	(3)	(4)
	0.078***	0.092***	0.071***	0.078***
	(0.019)	(0.024)	(0.019)	(0.021)
R-squared	0.231	0.195	0.244	0.231
F-test	69.42	21.14	93.54	63.38
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP Controls		yes		yes
Observations	430,490	430,490	430,490	430,490

Source: the 2002 census. Notes: Standard errors are clustered at the location level and are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

predictive power, and the results are robust to specifications.

In comparison with OLS estimates, coefficients are larger. With regard to the ability bias, this result is counterintuitive. If educated individuals have higher abilities, θ captures both the education and the ability effect and OLS estimates should be overestimated. However, the opposite effect can be observed when education is measured with error (Griliches, 1977) and when returns to education are heterogeneous[19] (Card, 2001). In this framework, the most plausible explanation is that instrumenting education removes the downward bias introduced with the use of the consumption proxy (see appendix A.2).

4.1.2 Returns to education by sector of activity

Thus far, returns to education have been estimated for the whole population. However, they can vary from one sector of activity to another. In this subsection, I investigate this question and estimate the consumption equation for each sector:

$$Log(C_{iajt}) = \alpha_a + \beta_{aj} + \beta_{at} + \theta_a S_{ijt} + \delta_{ta} X_j + \epsilon_{iajt} \tag{5}$$

the subscript "a" depicts the main activity of the household head and indicates whether the individual: 1) does not work or is unpaid, 2) works in agriculture, 3) works in nonfarm self-employed activities, or 4) is a wage-worker.

The first panel of Table 7 presents the OLS results. It shows that returns to education are much lower in agriculture than in the nonfarm self-employment activities and in wage-work activities. However, 2SLS estimates presented in the middle panel show that returns to education are higher in agriculture and in nonfarm self-employed activities than in the wage-activities. By comparing IV estimates with OLS estimates, one notices that the size of the bias varies between sectors of activity.[20]

This result might be explained by the magnitude of the ability bias between sectors of activity[21], but more plausibly, this discrepancy is explained by the nature of the

[19] When the instrument affects the education choices of less-educated subgroups, which have high marginal returns to education, IV estimates are upward biased. Regarding the UPE program that focused on individuals with restricted access to primary schools, IV estimates may overestimate the average marginal returns to education of the entire population.

[20] IV estimates are two times larger, 1.6 times larger and 20% smaller than OLS estimates in agriculture, nonfarm self-employed activities and wage-workers activities, respectively.

[21] If working in the formal sector requires higher abilities, the ability bias will be larger for wage workers.

Table 7: IV estimates of the returns to education by sector of activity

	(1)	(2)	(3)	(4)
Activity	Unpaid	Agriculture	Self-employed	Wage-work
OLS estimates				
	0.033***	0.033***	0.064***	0.054***
	(0.006)	(0.002)	(0.003)	(0.002)
IV estimates				
	0.027	0.066***	0.105***	0.043**
	(0.018)	(0.017)	(0.018)	(0.017)
R-squared	0.408	0.260	0.384	0.344
F-test	70.58	33.19	151	69.33
IV estimates with sample selection correction				
	0.025	0.067***	0.105***	0.045***
	(0.018)	(0.018)	(0.018)	(0.016)
Mills no work	-0.011			
	(0.011)			
Mills agri.		-0.012***		
		(0.003)		
Mills self.			0.004*	
			(0.002)	
Mills wage				0.021*
				(0.012)
R-squared	0.408	0.257	0.384	0.345
F-test	43.55	32.02	126.1	73.01
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP control	yes	yes	yes	yes
Observations	3,914	278,112	84,365	63,757

Source: the 2002 census. Notes: Standard errors are clustered at the region of birth level and are reported in parentheses. In IV estimations, standard errors are bootstraped. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

treatment. Since the aim of the UPE program was to boost rural productivity through agricultural classes (Kinunda, 1975), it is not surprising that individuals who benefited from this education policy have higher returns in agriculture than in wage activities.

4.1.3 Sample selection bias

The household consumption has the advantage of being available for the entire population, but estimating the returns to education for nonrandom subsamples, such as sectors of activity, might lead to sample selection issues. To address this possibility, I adopt the strategy described by Wooldridge (2010), which is also used by Duflo (2001). This two-stage model allows addressing both the endogeneity of education and the selection of samples (see appendix E for more details).

The results with sample selection correction are reported at the bottom panel of Table 7. The introduction of sample selection corrections does not change the IV estimates. The returns are still much higher in the agricultural sector and in the self-employment sector, while they are lower in the formal sector. Furthermore, coefficients of the Mills ratio are small but statistically significant, suggesting that subsamples are selected.

4.2 Effect of education on the choice of the sector of activity

Education can also ease the access to sectors that require skilled labor. To investigate this question, I estimate a multinomial logit model where A_{ijt} is the sector of activity, taking the value of 1 if the individual does not work or is unpaid, 2 if the individual works in the agricultural sector, 3 if the individual is self-employed in nonfarm activities and 4 if the individual has a wage-work employment. The activity equation has the following functional form:

$$A_{ijt} = \alpha + \beta_j + \beta_t + \theta S_{ijt} + \delta_t X_j + \epsilon_{ijt} \tag{6}$$

To account for endogeneity issues, I instrument education with the exposure to the UPE program and I follow a two-step control function approach (Wooldridge, 2014). After obtaining the predicted residual from the first stage equation, I plug it into equation (6). This predicted residual is also used to test the endogeneity of education.

In contrast to the returns to education estimated at the household level[22], the effect

[22]The statistical power is too low to instrument the education of several members of the households at the same time.

Table 8: Average marginal effect of education on the probability of working in each sector of activity (mult. logit)

Activity	Don't paid Don't work	agri	self employed	formal
	(1)	(2)	(3)	(4)
		Men		
OLS	0.000	-0.006***	-0.003***	0.009***
	(0.000)	(0.001)	(0.001)	(0.000)
IV: $\sum_{t=1961}^{1978} \gamma_t * N_{j,1967}$	-0.013***	0.023***	-0.003	-0.007
	(0.004)	(0.007)	(0.006)	(0.005)
$\hat{\mu_{ijt}}$		-0.444***	-0.281**	-0.117
		(0.125)	(0.110)	(0.103)
F-test		108.32		
Observations		415917		
		Women		
OLS	-0.003***	-0.006***	-0.001	0.010***
	(0.000)	(0.001)	(0.001)	(0.000)
IV: $I'_{j67} * T_{tot}$	-0.009*	0.014***	-0.014***	0.009***
	(0.005)	(0.005)	(0.004)	(0.003)
$\hat{\mu_{ijt}}$		-0.128***	0.112**	0.021
		(0.044)	(0.057)	(0.074)
F-test		162.27		
Observations		452544		
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP control	yes	yes	yes	yes

Sources: 2002 census. Notes: Standard errors, reported in parentheses, are bootstraped and clustered at the birth region level. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. CF-IV: IV estimates with control function method. Additional controls are the population aged 7 to 13 in 1967 and the household's size.

of education on the choice of the sector of activity is estimated at the individual level, which allows distinguishing the effect by gender.

The results are reported in Table 8. From OLS estimates, I observe that education increases the probability to work in the formal sector against agricultural and nonagricultural self-employed activities for men, and against agriculture and unpaid activities for women. However, IV estimates show a completely different picture.

Education raises the probability of working in agriculture and reduces the probability of having an unpaid job for both men and women. This switch towards agricultural activities is probably explained by the curriculum of the UPE program, which is composed of agricultural classes. In addition, the predicted residuals are statistically different from 0 in most specifications, which confirms the importance of dealing with the endogeneity of education.

4.3 Decomposition of the monetary effect of education

The monetary benefits of education occur because education increases the consumption level, conditional on the choice of the sector of activity, and changes the choice of the sector of activity, in which the consumption level varies. Thus, based on the above results and on the expected consumption: $E(C) = \sum_{a=1}^{n} P_a * C_a$, where P_a denotes the probability of working in the sector of activity a and C_a denotes the consumption level of individuals working in activity a, I decompose the returns to education into two components:

$$\frac{\delta E(C)}{\delta S} = \overbrace{\sum_{a=1}^{n} \frac{\delta P_a}{\delta S} * C_a}^{distribution\ effect} + \overbrace{\sum_{a=1}^{n} P_a * \frac{\delta C_a}{\delta S}}^{intra\ sector\ effect} \tag{7}$$

$\frac{\delta P_a}{\delta S}$ depicts the effect of education on the probability of working in activity a, and $\frac{\delta C_a}{\delta S}$ are the returns to education by activity. C_a and P_a are approximated by the predicted values of $\hat{C}_a$ and $\hat{P}_a$ from equation 5 and equation 6, respectively.

The left-hand side term represents the monetary benefit of education due to the change in the choice of the sector of activity, while the second term corresponds to the returns to education within sectors.

Table 9 provides the results from equation (7). OLS estimates show that both the distribution and the intrasector effects are positive and significant. In contrast, IV esti-

Table 9: The cumulative effect of education

	(1)	(2)	(3)	(4)
Model	OLS		IV: $I_{j,1967}$	
Mob effect	0.002***	0.002***	-0.003***	-0.0146***
	(0.001)	(0.001)	(0.009)	(0.002)
Within effect	0.040***	0.049***	0.087***	0.097***
	(0.012)	(0.008)	(0.003)	(0.001)
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP Control	no	yes	no	yes

Source: the 2002 census. Notes: Standard errors are clustered at the region of birth level and are reported in parentheses. Since results are produced from a multi-stage procedure, standard errors are bootstraped. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

mates suggest that the monetary benefits are mostly explained by the "intrasector effect", while the distribution effect is much smaller and negative. Since education increases the probability of working in the agricultural sector (see section 4.2) and because the average consumption is lower in this sector, this effect is not surprising and illustrates the specificity of the UPE program.

5 Discussion and robustness checks

To test whether 2SLS estimates are unbiased, I implement a series of robustness checks.

Regarding the identification strategy, one of the main concerns is that the exclusion restriction is not satisfied. The villagization process, which consisted of gathering people in community villages, probably had an impact over other concerns than education that could influence consumption. However, this should not put into question the validity of the instrument because the UPE program affected both pretreated and treated cohorts, with the exception of the education component that benefited the treated cohorts only. However, the identification strategy is invalidated if the age at which individuals were affected by the villagization has a direct effect on the consumption. To empirically check this issue, I reproduce the results by minimizing the age difference between the pretreatment and the treatment cohorts (see table A6). This entails excluding the youngest individuals of the treatment group (column 2) and the oldest individuals of the

pretreatment group (column 3). I find that the point estimates are slightly higher than the former results (column 1), but I do not reject the equality of the coefficients. I also test whether the introduction of individuals who were likely to be partially affected by the UPE program (individuals who were 13 years old between 1968 and 1974) change the results, but I do not find any significant difference (column 4 of table A6).

Heretofore, the instrument was constructed from the schooling supply by region of birth, and standard errors were clustered at the same level. However, a small number of clusters lead to overrejection of standard asymptotic tests (Cameron *et al.*, 2008). To check whether I underestimate the standard errors, I instrument education with the intensity $I'_{d,67}$, constructed from the educational attainment by district of residence in 1967.[23] Table A7 shows that the 2SLS estimates are slightly lower but are still significant at the 1% level. This entails that that overrejecting issue is negligible but that $I'_{d,67}$ is not purely exogenous.

One of the main limitations of this study is that the returns to education are estimated for the household head only, while everyone in the household can contribute to the consumption.[24] Although the household head is likely to take most of the decisions that influence the household decision, the education of the most educated individual also matters if the knowledge is shared among individuals. Thus, I also provide estimates of the returns to education for the highest educated individual in the household, and I find very similar estimates (see table A8).[25]

6 Conclusion

This paper studies the benefits of education in Tanzania and considers two particular dimensions: household consumption and the choice of the sectors of activity. To address endogeneity issues, I instrument education of household heads by exploiting variation in time and in space of the exposition to the Universal Education Program.

I find that this massive primary education program contributed to a reduction in inequalities among regions. After this program ended, its effects persisted for the next

[23] In 2012, there were 31 regions against 169 districts in Tanzania.

[24] The statistical power is too low to instrument the education of different members of the households at the same time.

[25] The sample excludes approximately 10% of households, for which the highest education level is reached by two or more individuals.

age cohorts. Despite the controversial means of villagization, the Tanzanian government fulfilled its goals by improving access to basic education, even in remote areas. Unfortunately, several changes were implemented at the same time, which prevents one from identifying the effectiveness of each policy.

By using a household survey, census data, and records on the number of schools, I find that education increases household consumption between 7.3 and 9.3 percent, depending on the specification and the instrument. The main contribution of this analysis is to focus on the entire population, instead of wage workers who are in the minority in most developing countries and are very likely to be self-selected. I also compare the returns to education between sectors of activity. I find that the returns to education are higher in agriculture and in nonfarm self-employed activities than in wage-work activities. This conclusion, initially surprising, is consistent with the Tanzanian governmental policy that aimed to put education at the service of agriculture by teaching agricultural skills. Compared to the studies on the benefits of primary education in agriculture in African countries that find low returns (Appleton *et al.*, 1996; Jolliffe, 2004), I argue that returns to education in agriculture are positive, provided that the curriculum at school is consistent with agriculture. This conclusion is closer to Foster and Rosenzweig (1996)'s results, suggesting that returns to education are positive only during specific contexts, such as during technological changes, when education helps farmers to adopt new technologies.

These findings suggest that the introduction of agricultural classes could help households to escape poverty by increasing the farmers' productivity. In terms of public recommendations, this result is all the more relevant in most African countries, where the large majority of the population works in agriculture and where the agricultural productivity remains low.

References

Appleton, S., Balihuta, A. *et al.*, 1996, *Education and agricultural productivity: Evidence from Uganda*, University of Oxford, Centre for the Study of African Economies.

Barham, B. and Boucher, S., 1998, 'Migration, remittances, and inequality: estimating the net effects of migration on income distribution', *Journal of Development Economics*, 55(2), 307–331.

Bertrand, M., Duflo, E. and Mullainathan, S., 2004, 'How much should we trust differences-in-differences estimates?', *The Quarterly Journal of Economics*, 119(1), 249–275.

Bonini, N., 2003, 'Un siècle d'éducation scolaire en Tanzanie', *Cahiers d'études africaines*, (1), 40–62.

Cameron, A.C., Gelbach, J.B. and Miller, D.L., 2008, 'Bootstrap-based improvements for inference with clustered errors', *The Review of Economics and Statistics*, 90(3), 414–427.

Card, D., 2001, 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica*, 69(5), 1127–1160.

Court, D. and Kinyanjui, K., 1980, 'Development policy and educational opportunity: the experience of Kenya and Tanzania', Occasional Paper 33, Nairobi: Institute for Development Studies, University of Nairobi.

De Chaisemartin, C. and d Haultfoeuille, X., 2015, 'Fuzzy differences-in-differences', Cemmap working paper No. CWP69/15, Centre for Microdata Methods and Practice.

Deaton, A. and Zaidi, S., 2002, *Guidelines for constructing consumption aggregates for welfare analysis*, volume 135, World Bank Publications.

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.

Elbers, C., Lanjouw, J.O. and Lanjouw, P., 2003, 'Micro–level estimation of poverty and inequality', *Econometrica*, 71(1), 355–364.

Foster, A.D. and Rosenzweig, M.R., 1996, 'Technical change and human-capital returns and investments: evidence from the green revolution', *The American economic review*, 931–953.

Gillette, A., 1975, 'L'éducation en Tanzanie: une réforme de plus ou une révolution éducationnelle?', *Revue Tiers Monde*, 735–756.

Griliches, Z., 1977, 'Estimating the returns to schooling: Some econometric problems', *Econometrica*, 1–22.

Gubert, F., Lassourd, T. and Mesplé-Somps, S., 2010, 'Transferts de fonds des migrants, pauvreté et inégalités au Mali', *Revue économique*, 61(6), 1023–1050.

Jensen, S., Mkama, J. and ya Uchumi na Mipango ya Maendeleo, T.W., 1968, *District data, Tanzania, 1967*, Ministry of Economic Affairs and Development Planning.

Jolliffe, D., 2004, 'The impact of education in rural Ghana: Examining household labor allocation and returns on and off the farm', *Journal of Development Economics*, 73(1), 287–314.

Kinunda, M.J., 1975, *Experience in Tanzania in identifying and satisfying local needs in Education: A Contribution to the IIEP Seminar on "The Planning of learning arrangements of all kinds for local Communities", 9-17 December 1974*, volume 14, International Institute for Educational Planning.

Lockheed, M.E., Jamison, T. and Lau, L.J., 1980, 'Farmer education and farm efficiency: A survey', *Economic Development and Cultural Change*, 29(1), 37–76.

Maluccio, J., 1998, 'Endogeneity of schooling in the wage function: Evidence from the rural Philippines', *Food Consumption and Nutrition Division Discussion Paper*, 54.

Maro, P.S. and Mlay, W.F., 1979, 'Decentralization and the organization of space in Tanzania', *Africa*, 49(03), 291–301.

Martin, D., 1988, *Tanzanie: l'invention d'une culture politique*, KARTHALA Editions.

Meyer, B.D. and Sullivan, J.X., 2003, 'Measuring the well-being of the poor using income and consumption', National Bureau of Economic Research Working Pape No. 9760.

Nyerere, J.K., 1967, 'Education for self-reliance', *The Ecumenical Review*, 19(4), 382–403.

Shao, I.F., 1982, 'A neo-colony and its problems during the process of attempting to bring about socialist rural transformation: the case of Tanzania', In: *Taamuli: a Political Science Forum*, volume 12, 29–46.

Tarozzi, A. and Deaton, A., 2009, 'Using census and survey data to estimate poverty and inequality for small areas', *The Review of Economics and Statistics*, 91(4), 773–792.

Wooldridge, J.M., 2010, *Econometric Analysis of Cross Section and Panel Data, second edition*, MIT press.

Wooldridge, J.M., 2014, 'Quasi-maximum likelihood estimation and testing for non-linear models with endogenous explanatory variables', *Journal of Econometrics*, 182(1), 226–234.

A The proxy for consumption

A.1 Construction

By following a two step-procedure, I predict household consumption from a set of predictors P that are common to both household and census data.[26] The idea behind this method is first to estimate the joint distribution of the consumption, $logC$, and of P from the household survey:

$$LogC_{ijt} = bP_{ijt} + \delta_j + \nu_{ijt} \tag{8}$$

where ν_{ijt} is the error term of household i. Then, I use the estimated distribution $\hat{b}P_{ijt}$ to predict $\widehat{logC_{ijt}}$ in the census data. This method is valid if the predictors P are similarly measured in both data sets, and if questions have the same wording for the two questionnaires.[27] Table A1 shows coefficients of equation 8 in the LSMS data. As predicted, all dwelling characteristics have a positive and significant impact on consumption. The R-squared coefficient is high, meaning that the predictors have good explanatory power. In table A2, the comparison between $LogC_{ijt}$ and $\widehat{logC_{ijt}}$ shows that the share of the consumption explained by the dwelling characteristics reaches approximately two-thirds, and confirms the fact that dwelling characteristics have a high explanatory power. Graph A1 plots the relationship between $\widehat{logC}$ and $logC$ in the LSMS data.[28] I find a clear positive linear relationship between these two variables. Although $\widehat{logC}$ may not capture all the variation in consumption, especially at the tails of the distribution, this does not appear to be empirically the case. The dispersion for extreme values of $\widehat{logC}$ is larger, but this effect stays negligible.

To account for the artificially low variance of the consumption proxy, I adopt the method proposed by Barham and Boucher (1998) and Gubert *et al.* (2010). This approach recommends adding to an error term drawn from a normal distribution with the same variance $\widehat{\nu_{ijt}}$ that is observed in the survey data. To make sure that the results are independent from the random draw, this procedure is replicated a large number of times. Due to this method, the standard errors can be normally interpreted.

[26] The number of rooms in the dwelling, whether the household has drinking water, electricity, a phone, a flush toilet, a high-quality roof, high-quality walls, etc.

[27] To avoid anachronism issues, I do not include in the list of predictors "having a phone", which may have a different meaning across time and across the data.

[28] For each value of $\widehat{logC}$, I compute the average value of $logC$ depicted by a dot.

Table A1: Effect of dwelling characteristics on consumption

VARIABLES	$log(C_{ijt})$
Solid wall	0.148***
	(0.015)
Housing water	0.124***
	(0.019)
Flush toilet	0.040**
	(0.016)
Electricity	0.388***
	(0.019)
Permanent floor	0.379***
	(0.017)
Solid roof	0.478***
	(0.055)
Nb. of bedrooms	0.093***
	(0.005)
Age of household head	-0.002***
	(0.000)
Gender of household head	-0.107***
	(0.014)
Number children aged 5-15	0.092***
	(0.004)
Number adults aged 16-65	0.157***
	(0.004)
Constant	12.566***
	(0.041)
R-squared	0.532
Observations	12,178

Sources: The three pooled waves of the LSMS data. Notes: additional controls: Regions dummies, survey year dummies. Standard errors are reported in parentheses. ***,**,* mean respectively that the coefficient are significantly different from 0 at the level of 1%, 5% and 10%.

Figure A1: Relationship between the expected consumption $\widehat{lnC}$ and lnC.

Source: LSMS data (2008, 2010, 2012)

A.2 Consequences of the use of a consumption proxy

To obtain unbiased estimates of the returns to education by using the consumption proxy, one additional validity assumption should be satisfied: education and $\widehat{\nu_{ijt}}$ obtained from equation (8) should be uncorrelated. Heretofore, $\widehat{\nu_{ijt}}$ is assumed to be exogenous and has been drawn from a normal distribution. However, $\widehat{\nu_{ijt}}$, which represents the consumption part unexplained by households' dwelling characteristics, may result from households' preferences and may be correlated with the education of the household's head S_{ijt}. For instance, educated household heads may be more willing to spend money for the education or health of their children. If so, there is a remaining endogenous part of the residual denoted ν''_{ijt} ($\nu''_{ijt} = \nu_{ijt} - \widehat{\nu}_{ijt}$) that is not captured by the drawn residual $\widehat{\nu_{ijt}}$. Thus, by combining equations (8) and (4):

$$\left.\begin{aligned} LogC_{ijt} = bP_{ijt} + \nu_{ijt} = \hat{b}p_{ijt} + \widehat{\nu_{ijt}} + \nu''_{ijt} = \alpha + \beta_j + \beta_t + \theta S_{ijt} + \delta_t X_j + \varepsilon_{ijt} - \nu''_{ijt} \\ \widehat{LogC_{ijt}} = \hat{b}X_{ijt} + \widehat{\nu_{ijt}} \end{aligned}\right\} \quad (9)$$

I deduce that:

$$LogC_{ijt} = \hat{b}X_{jt} + \widehat{\nu_{ijt}} = \alpha + \beta_j + \beta_t + \theta S_{ijt} + \delta_t X_j + \varepsilon_{ijt} - \nu''_{ijt} \quad (10)$$

and $\widehat{\theta} = \theta + \frac{cov(\varepsilon_{ijt}, S_{ijt})}{V(S_{ijt})} - \frac{cov(\nu''_{ijt}, S_{ijt})}{V(S_{ijt})}$.

The positive correlation between education and ε_{ijt} leads to the traditional upward bias, while the positive correlation between education and ν''_{ijt} causes downward bias in the coefficient of interest. Thus, if ν_{ijt} is not purely exogenous, using the proxy for consumption adds an additional source of bias.

B Sample and statistic descriptives

Figure A2: Evolution of education attainment by region from T0 to T1 according to the number of schools in 1967.

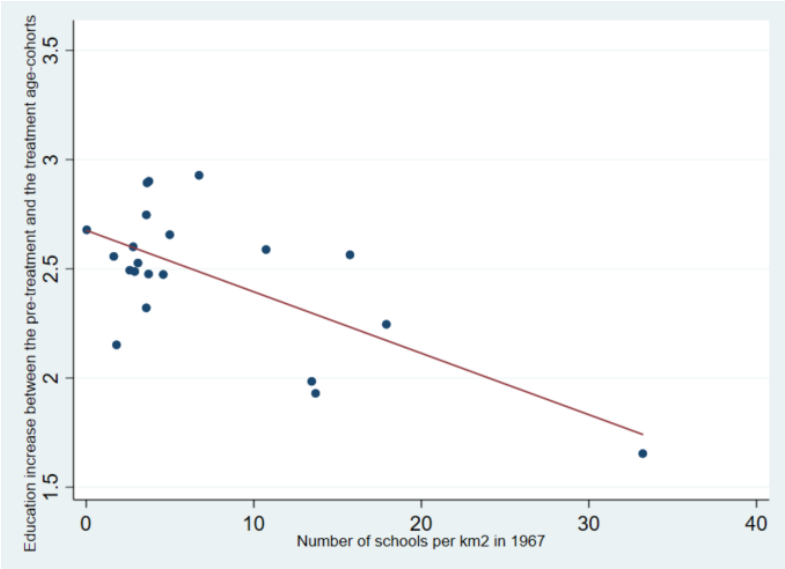
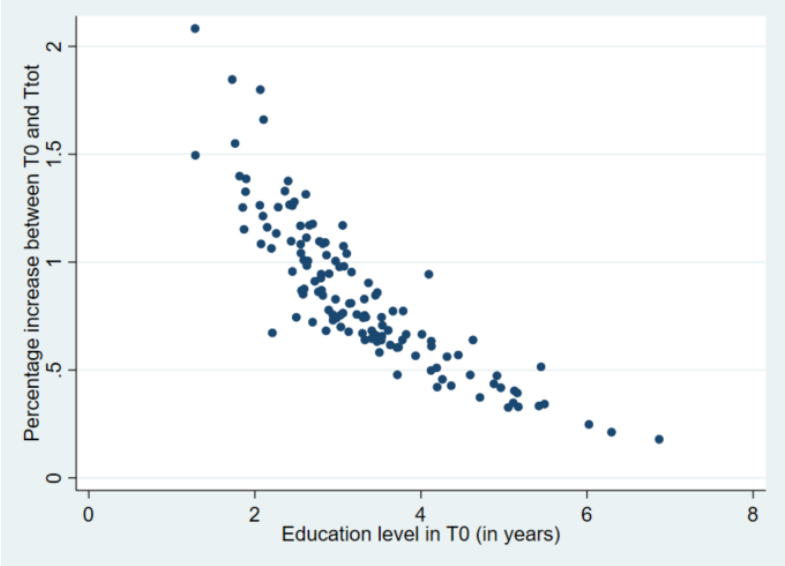


Figure A3: Evolution of education attainment by region from T0 to T1 according to the education level in T0.



Sources: The 2002 census.

Table A2: Descriptive statistics from the LSMS panel data

	(1) All	(2) T_0	(3) T_{Tot}	(4) T_0-T_{Tot}	(5) Region$^-$	(6) Region$^+$	(7) Region$^-$-Region$^+$
Age	46,992	60,283	40,396	19,89***	47,129	46,778	0,351
	(15,875)	(3,128)	(5,470)	(0,139)	(16,257)	(15,314)	(0,292)
Men	0,248	0,307	0,211	0,0952***	0,255	0,238	0,0163*
	(0,432)	(0,461)	(0,408)	(0,0118)	(0,436)	(0,426)	(0,00793)
Urban areas	0,352	0,309	0,364	-0,0541***	0,306	0,416	-0,110***
	(0,477)	(0,462)	(0,481)	(0,0133)	(0,461)	(0,493)	(0,00871)
Years of primary edu,	4,921	3,856	5,763	-1,906***	4,577	5,403	-0,826***
	(2,898)	(3,011)	(2,428)	(0,0728)	(3,026)	(2,637)	(0,0531)
Ended primary edu	0,598	0,387	0,749	-0,362***	0,547	0,669	-0,122***
	(0,490)	(0,487)	(0,434)	(0,0125)	(0,498)	(0,471)	(0,00894)
Man's activity							
Wage worker	0,242	0,211	0,250	-0,0388**	0,202	0,296	-0,0940***
	(0,428)	(0,408)	(0,433)	(0,0149)	(0,402)	(0,457)	(0,00943)
Self-employed	0,160	0,107	0,178	-0,0704***	0,139	0,188	-0,0497***
	(0,366)	(0,310)	(0,382)	(0,0129)	(0,346)	(0,391)	(0,00809)
Works in agriculture	0,317	0,435	0,268	0,167***	0,344	0,279	0,0658***
	(0,465)	(0,496)	(0,443)	(0,0159)	(0,475)	(0,448)	(0,0103)
Wage-worker and self-employed	0,026	0,023	0,027	-0,00428	0,027	0,024	0,00331
	(0,158)	(0,150)	(0,163)	(0,00560)	(0,162)	(0,152)	(0,00350)
Wage-worker and agriculture	0,123	0,108	0,133	-0,0245*	0,132	0,110	0,0226**
	(0,328)	(0,311)	(0,340)	(0,0117)	(0,339)	(0,313)	(0,00726)
Self-employed and agriculture	0,133	0,115	0,144	-0,0287*	0,155	0,103	0,0520***
	(0,340)	(0,319)	(0,351)	(0,0120)	(0,362)	(0,304)	(0,00750)
Woman's activity							
Wage worker	0,159	0,114	0,156	-0,0415*	0,153	0,168	-0,0144
	(0,366)	(0,318)	(0,363)	(0,0202)	(0,360)	(0,374)	(0,0150)
Self-employed in non-agri	0,227	0,196	0,272	-0,0762**	0,210	0,256	-0,0457**
	(0,419)	(0,397)	(0,445)	(0,0249)	(0,407)	(0,437)	(0,0172)
Works in agriculture	0,359	0,464	0,259	0,205***	0,369	0,343	0,0264
	(0,480)	(0,499)	(0,438)	(0,0265)	(0,483)	(0,475)	(0,0197)
Wage-worker and self-employed	0,027	0,007	0,036	-0,0294**	0,023	0,034	-0,0110
	(0,161)	(0,083)	(0,187)	(0,00943)	(0,149)	(0,180)	(0,00664)
Wage-worker and works in agri	0,093	0,079	0,097	-0,0178	0,101	0,082	0,0188
	(0,291)	(0,270)	(0,296)	(0,0167)	(0,301)	(0,274)	(0,0120)
Self-employed and works in agri	0,134	0,140	0,180	-0,0401	0,144	0,118	0,0259
	(0,341)	(0,347)	(0,384)	(0,0216)	(0,352)	(0,323)	(0,0140)
log ($\widehat{consumption}$)	13,790	13,983	13,932	0,0515	13,675	13,948	-0,273***
	(1,046)	(1,107)	(0,944)	(0,0276)	(1,033)	(1,044)	(0,0191)
$\widehat{consumption}$	1924560	2901587	2032611	868975,8**	1862755	2010673	-147918,3
	(8956441)	(18700000)	(7569948)	(319670,2)	(11100000)	(4636514)	(164533,7)
consumption	3424040	3654149	3782781	-128631,4	3130704	3835226	-704522,7***
	(8707449)	(4886988)	(11700000)	(290802,0)	(9612158)	(7248920)	(159844,2)
Observations	12195	1706	5119	6825	7096	5092	12188

Sources: The 2002 census (IPUMS data). ***,**,* means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. Standard deviations are reported in parentheses columns (1) to (3), and (5) to (6). Standard error for average difference are reported in parentheses columns (4) and (7).

Table A3: Descriptive statistics of the treated-cohort T_{tot}, depending on the education status.

	Did not complete primary education	Completed primary education	T-test
Age	32.463	31.991	-4.404***
	(5.337)	(4.943)	(0.0209)
Men	0.591	0.727	-0.0454***
	(0.492)	(0.445)	(0.000585)
Urban status	0.297	0.518	-0.182***
	(0.457)	(0.500)	(0.000565)
HH head doesn't work	0.127	0.094	0.0239***
	(0.333)	(0.292)	(0.000407)
HH head works in agri	0.676	0.511	0.139***
	(0.468)	(0.500)	(0.000574)
HH head is self-emp.	0.151	0.223	-0.063***
	(0.358)	(0.416)	(0.001)
HH head is a wage worker	0.045	0.173	-0.099***
	(0.208)	(0.378)	(0.001)
$log(\widehat{consumption})$	13.751	14.198	-0.313***
	(0.543)	(0.685)	(0.001)
$\widehat{consumption}$	1122746	1884895	-775023.4***
	(999477.600)	(1747951.0)	(3601.3)
Observations	100121	288580	3069955

Sources: The 2002 census (IPUMS data). ***.**.* means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. Standard deviations are reported in parentheses columns (1) to (3), and (5) to (6). Standard errors for average difference are reported in parentheses columns (4) and (7).

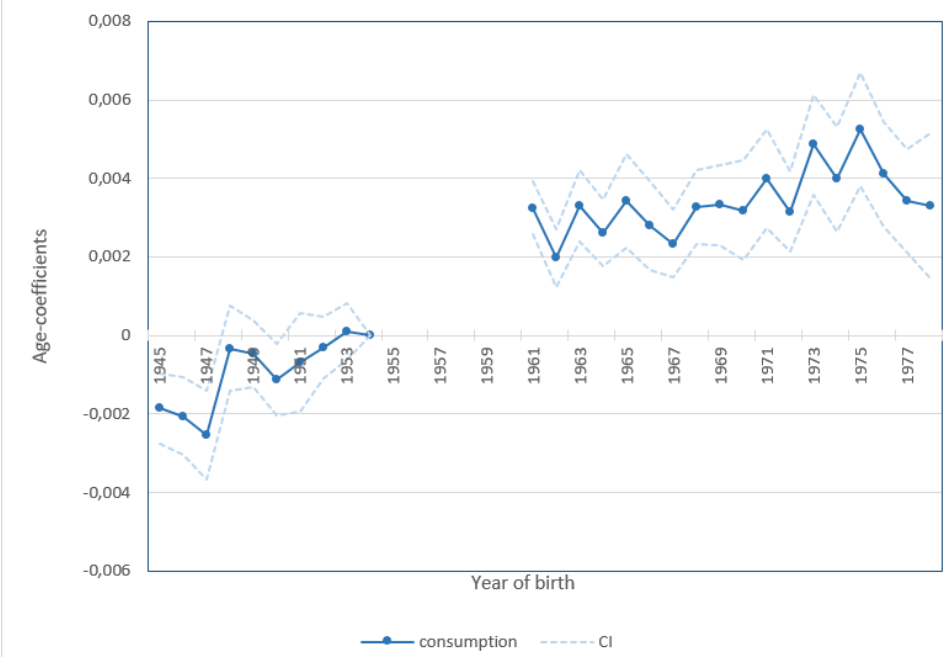
C First stages

Table A4: Effect of the UPE program on education (γ coefficients of 1) with LSMS data.

	(1)	(2)	(3)	(4)
Instrument	$I_{j1967} * T_{tot}$		$I'_{d1967} * T_{tot}$	
	0.007	-0.003	0.186	0.368
	(0.014)	(0.018)	(0.124)	(0.239)
R-squared	0.196	0.198	0.197	0.198
F-test	0.227	0.0307	2.244	2.383
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP Control	no	yes	no	yes
Observations	4983	4983	4983	4983

Source: the pooled LSMS survey (2008, 2010, 2012). Notes: Standard errors are clustered at the region of birth level and are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

Figure A4: Effect of the UPE program on the logarithm of the consumption proxy.



Sources: The 2002 census.

Table A5: Effect of the program on the education level: γ_t coefficients of equation (3)

	(1)	(2)	(3)	(4)
	$I'_{d,1967}$		$I_{j,1967}$	
1961	0.266***	0.268***	0.046***	0.048***
	(0.053)	(0.054)	(0.008)	(0.010)
1962	0.158***	0.164***	0.025***	0.027**
	(0.037)	(0.038)	(0.009)	(0.010)
1963	0.390***	0.398***	0.041***	0.042***
	(0.042)	(0.044)	(0.009)	(0.010)
1964	0.374***	0.385***	0.036***	0.038***
	(0.073)	(0.073)	(0.008)	(0.009)
1965	0.480***	0.489***	0.050***	0.052***
	(0.037)	(0.039)	(0.006)	(0.009)
1966	0.445***	0.456***	0.056***	0.058***
	(0.048)	(0.050)	(0.007)	(0.010)
1967	0.424***	0.433***	0.043***	0.045***
	(0.052)	(0.054)	(0.008)	(0.010)
1968	0.555***	0.564***	0.066***	0.068***
	(0.041)	(0.043)	(0.006)	(0.012)
1969	0.466***	0.479***	0.063***	0.065***
	(0.051)	(0.053)	(0.007)	(0.013)
1970	0.472***	0.483***	0.062***	0.064***
	(0.055)	(0.057)	(0.006)	(0.013)
1971	0.397***	0.410***	0.061***	0.063***
	(0.046)	(0.048)	(0.007)	(0.013)
1972	0.392***	0.400***	0.055***	0.057***
	(0.063)	(0.064)	(0.007)	(0.014)
1973	0.447***	0.460***	0.059***	0.062***
	(0.050)	(0.053)	(0.008)	(0.016)
1974	0.441***	0.453***	0.065***	0.068***
	(0.050)	(0.052)	(0.008)	(0.017)
1975	0.496***	0.513***	0.071***	0.074***
	(0.053)	(0.056)	(0.007)	(0.016)
1976	0.574***	0.588***	0.064***	0.067***
	(0.069)	(0.074)	(0.008)	(0.017)
1977	0.506***	0.522***	0.065***	0.069***
	(0.053)	(0.054)	(0.008)	(0.016)
1978	0.510***	0.527***	0.067***	0.070***
	(0.068)	(0.071)	(0.007)	(0.017)
R-squared	0.285	0.287	0.272	0.272
F-test	32.99	31.47	62.30	51.46
Observations	433,606	433,606	433,606	433,606

Source: the 2002 census. Notes: Standard errors are clustered at the birth region level and are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

D Robustness Checks:

Table A6: 2SLS estimates of the returns to education with different age-cohorts

	(1)	(2)	(3)	(4)
Control age-cohorts	1945-1954	1945-1954	1949-1954	1945-1954
Treatment age-cohorts	1961-1978	1961-1965	1961-1965	1955-1978
	0.078***	0.082***	0.089***	0.074***
	(0.021)	(0.018)	(0.017)	(0.018)
R-squared	0.231	0.341	0.301	0.292
F-test	63.38	13.77	10.08	10.21
Cohort FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
GDP Controls	yes	yes	yes	yes
Observations	430,490	176,359	145,126	503,156

Source: the 2002 census. Notes: Standard errors are clustered at the region of birth level and are reported in parentheses. ***,**,* mean respectively that the coefficient are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household size and the principal sector of activity of the household head.

Table A7: IV estimates of the returns to education

	(1)	(2)	(3)	(4)
IV:	$T * S_{d,1967}$		$\sum_{t=1961}^{1978} \gamma_t * S_{d,1967}$	
	0.059***	0.061***	0.054***	0.056***
	(0.010)	(0.009)	(0.009)	(0.009)
R-squared	0.276	0.289	0.282	0.295
F-test	101.5	97.12	33.43	31.74
Cohort FE	yes	yes	yes	yes
District FE	yes	yes	yes	yes
GDP Controls		yes		yes
Observations	430,490	430,490	430,490	430,490

Source: the 2002 census. Notes: Standard errors are clustered at the location level and are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

Table A8: 2SLS estimates of the returns to education of different members

	(1) The household head	(2) The most educated individual
	0.068***	0.062***
	(0.023)	(0.020)
R-squared	0.234	0.229
F-test	43.60	41.94
Cohort FE	yes	yes
Region FE	yes	yes
GDP Controls	yes	yes
Observations	361,923	312,256

Source: the 2002 census. Notes: Standard errors are clustered at the location level and are reported in parentheses. ***,**,* mean respectively that the coefficients are significantly different from 0 at the level of 1%, 5% and 10%. Additional controls are the population aged 7 to 13 in 1967, the household's size and the sector of activity.

E Measuring the effect of education by sector of activity with the Heckman selection model

To overcome endogeneity and selection issues, I follow the method proposed by Wooldridge (2010) based on the three following equations: the consumption equation, the selection equation where A_{ijt} represents the sector of activity of the household's head, and the endogenous education equation.

$$\left.\begin{aligned} Log(C_{iajt}) &= \alpha_{1a} + \beta_{1aj} + \beta_{1at} + \theta_{1a}S_{ijt} + \delta_{1ta}X_j + \epsilon_{1isjt} \\ A_{ijt} &= \alpha_{2a} + \beta_{2aj} + \beta_{2at} + \theta_{2a}I_{j,67} * T + \gamma_{2a}N_{ijt} + \delta_{2ta}X_j + \epsilon_{2iajt} \\ S_{ijt} &= \alpha_{3a} + \beta_{3aj} + \beta_{3at} + \theta_{3a}I_{j,67} * T + +\delta_{3ta}X_j + \epsilon_{3iajt} \end{aligned}\right\}$$

To obtain unbiased estimates of the returns to education, I first regress the choice of the sector of activity on the instrument in order to deduce the predicted probabilities of working in the different sectors of activity. Then, to control for sample selection, I compute the inverse Mills ratios $\hat{\lambda_{ia}}$ that I introduce into the consumption equation:

$$Log(C_{iajt}) = \alpha_{1a} + \beta_{1aj} + \beta_{1at} + \theta_{1a}S_{ijt} + \delta_{1ta}X_j + \gamma_{1a}\hat{\lambda_{ia}} + \epsilon_{1iajt} \tag{11}$$

Standard errors are bootstrapped to account for it as a two-step procedure.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2019/3082/10