Africa's Education Enigma? The Nigerian story.*

Ruth Uwaifo Oyelere†

October 20, 2005

Abstract

In the last two decades, the social and economic benefits of formal education in sub-Saharan Africa has been debated. Anecdotal evidence points to low and time varying returns to education in Africa. Unfortunately, there has been little econometric evidence to support these claims at the micro level. Here I focus on Nigeria, a country that holds 1/5 of Africa's population, and use instruments based on the exogenous timing of the implementation and withdrawal of free primary education across regions in this country to precisely estimate the returns to education in the late 1990s. In addition, claims of time differences in returns are investigated. The results show that the average returns to education are particularly low in the 90s, in contrast to conventional wisdom for developing countries (3.6% for every extra year of schooling in 1998). In addition, there have been significant changes in returns to education for head of households over short time periods. These results shed new light on both the changes in demand for education in Nigeria and the increased emigration rates from African countries that characterized the 90s.

(**JEL classification:** C13,C30, J24, I21, I29, O12; **Keywords:** Human capital, Instrumental variables, Nigeria, Returns to education, Schooling)

---

*This research has received funding partly from the Center for African Studies University of California, Berkeley. The author would like to especially thank Brian Wright for invaluable advice, patience and support of this work. I am also indebted to Edward Miguel, Jenny Lanjouw, Chris Udry and Alain de janvry for their comments and useful advice. Adebayo Aromolaran and Nosa Nwosu were instrumental in securing the datasets. This work has benefitted from helpful comments from participants in the development workshops and seminar at University of California Berkeley and development workshop at Yale University. Of course, all remaining errors are mine. Any comments are greatly appreciated. Please direct correspondence to uwaifo@are.berkeley.edu

†PhD Candidate, Department of Agriculture and Resource Economics, UC Berkeley.

1 Introduction

Over the last three decades, questions have been raised on why many developing countries are not experiencing significant growth and development especially in sub-Saharan Africa. Explanations have included a combination of poor technology, bad governments, extractive institutions, weak policy choices, health crises and poor education (see Easterly (2001)[19]). In the last ten years several authors have considered these hypotheses regarding lack of growth in several African countries. The education sector has been examined extensively, but one important question, the return to education, is still unresolved.

Though attempts have been made to estimate returns to education in the past, the econometric techniques used in these estimations are prone to bias because of measurement error and unobservables correlated with schooling. With the development of new econometric techniques early in the 90s to deal with these problems, there has been a resurgence of interest in the estimation of returns to education in other parts of the world. However, most of the recent studies on Africa have not made use of these new econometric techniques, for lack of appropriate instruments. Hence, estimates for return to schooling were still derived using ordinary least squares (OLS)[1]. As the endogenous nature of schooling is not addressed with the OLS estimator, the estimated returns to education could be biased. Hence, there is still room for improvement in estimating returns in Africa.

In this paper, returns to education are estimated using the instrumental variable approach. I consider the most populous country in Africa, Nigeria. The Nigerian case is especially interesting because of its importance in Africa in terms of population size (one out every five Africans is Nigerian), diversity (one of the

[1]Relevant papers are highlighted in the literature review

most ethnically diverse with over, 354 languages), and key position in oil and gas production in Africa. As with some other African countries, the role and importance of formal education in Nigeria has been debated since the economic downturn in the early 80s. This controversy was linked primarily to the lack of significant growth in the economy over the 80s and 90s, despite the massive increase in human capital investment via education in the 1960's and 1970's. Also contributing to this controversy was the fall in living standards and real income of many well-educated Nigerians between 1983 and 1998, relative to some of their uneducated counterparts. This situation has raised many unanswered questions about the private and social value of education in Nigeria. Two of these questions will be addressed in this paper.

The main research question I consider is what were the returns to education in Nigeria? The goal here is to precisely estimate the returns to education as revealed in income late in the 90s in Nigeria.[2] The answer should not only provide estimates of the average returns to education in an African country where the economic value of education is the subject of debate, but can also help us evaluate the extent of bias of ordinary least squares estimates of returns to education in the Nigerian case. The second question I would be considering is do time differences in returns to education exist? Here, I would test the null hypothesis that there are no time differences in returns to education in Nigeria.

The returns to education are estimated in this paper using two stage least squares (2SLS). The instrument used in this analysis is based on a free primary education program called Universal Primary Education (UPE), designed to increase educational attainment, exploiting differences in the periods of implementation

[2]In this paper, private return to education is simply referred to as return to education.

of this program across states/regions over time in Nigeria, along the lines of the approach used in Duflo (2001)[18].

The instrument can be constructed in different ways. I construct the instrument as the length of exposure to free education. The argument here is the longer an individual is exposed to free education, the higher the school attainment.

To highlight the importance of including appropriate controls in the estimation, the 2SLS estimation of returns to education was carried out, both with and without additional variables. Furthermore, as a benchmark to compare these estimates, the OLS technique is also used to estimate the returns to education. Using these techniques, I estimate a 3.6% and 3.0% increase in income for every extra year of schooling in Nigeria in 1997/98 and 1998/99 respectively. This estimate of return to education is low and far from what the conventional wisdom expects for a developing country in terms of returns to education. Furthermore, these estimates are much lower than other estimates in other sub-saharan countries. The review of Psacharopoulos and Patrinos (2002)[52] reports average returns to education in Africa of 11.7%[3]. Aromolaran's (2002)[6] estimates of returns to education in Nigeria, which did not correct for potential sources of bias, are also higher than these estimates.

Using these techniques, I also reject the null hypothesis of no time differences in returns to education. In fact, for head of households, the average returns to education was extremely high in the 80s, fell to insignificant levels by 1992, and rose to 5.3% by 1996. Several robustness checks were carried out including correcting for potential sources of selectivity and the above results still hold. Finally, I find

[3] Also, see Schultz (2004)[54] for a review on estimates for selected African countries. It should be mentioned that my estimates are not directly comparable to the studies highlighted in Schultz (2004)[54] which estimate returns at each level of education.

that OLS estimates of returns to education are biased but the direction of the bias differs when considering different groups in the population. Also, omitting important control variables from the wage equation can bias returns to education estimates significantly.

The present study therefore provides the first estimates of returns to education, using a credible instrument, in a West African country. Furthermore, the results draw attention to two important issues with education outcomes in Africa, not highlighted prior to now: low returns to education and time differences in returns to education. Low and time differences in returns are important since low returns can lead to a fall in the demand for education over time and fluctuating returns make investment in education risky and could also have similar consequences. A fall in education investment could be a problem if education investment has large externalities or social returns despite low and fluctuating private returns. Furthermore, this paper draws attention to the importance of including controls in the estimation of returns to education. Finally, several explanations have been sought for the changing demand for education and increased emigration rates in the 90s. The low returns to education in Nigeria suggests one possible explanation for these phenomena.

The remainder of this paper is organized as follows: In the next section I review the relevant literature on schooling. Section 3 gives a review of the general theoretical framework for the analysis. Section 4 presents the data, Section 5 highlights initial data analysis. Section 6 highlights the empirical and identification strategies and section 7 presents the results. Section 8 highlights robustness checks and the last section provides implications, concluding remarks and directions for future research.

2 Literature Review

According to economic theory, earnings are a function of worker productivity. An important policy issue is the extent to which productivity and consequently earnings, are influenced by educational attainment. A school of thought advanced by Spence (1973)[58] and Arrow (1973)[7] in the 1970s points to education as a signal or a screening process of innate ability. This view is linked to the "sheepskin effect hypothesis". On the other hand Bhagwati and Srinivasan(1977)[10] view education as a tool for job competition in a distorted labor market. The third, and most common approach to looking at education came from Becker's (1964)[11] seminal paper in which he views education as an investment in human capital.

From the 1950s, different models have been proposed and tested to evaluate the hypothesis that education affects earnings. Though this relationship has been explored in different ways, recently, schooling and its relationship to wage determination have most often been analyzed in the framework of Mincer's (1974)[41] wage equation. Over the years, several authors have noted various flaws to this human capital approach. These flaws include omitted variables in the estimation equation, and problems of endogeneity of the education coefficients. Hence, non-observed post-schooling on the job training and the absence of suitable comparison groups (as it is almost never possible to observe what particular persons would have earned had they obtained more or less schooling than they did, the closest exception being the identical twin studies) can lead to omitted variable bias and endogeneity.

Adjustments have been suggested to the earnings function in order to deal with the problems stated above. Much of the schooling literature, starting from the late

70s, focuses on disentangling education's independent effect on wages. Examples of papers attempting to do this using different techniques are Griliches (1977)[28], Angrist & Kruger (1991)[3], Ashenfelter et al (1998)[8], Harmon et al (1998)[30], Card (1999)[14] and Duflo 2001[18]. The most commonly used new technique relies on finding instrumental variables (IV) to correct for the endogenous nature of schooling.

Most of the studies using an IV strategy to properly estimate returns to education have focused on developed countries. Studies using the IV approach are less common for developing countries (see Psacharopoulos and Patrinos (2002)[52] and Card (1999)[14]). The best known paper using the IV technique in a developing country is Duflo (2001)[18] on Indonesia. Since this paper, other attempts have been made in developing countries but there has been little progress considering African countries (see Glewwe (2002)[26] for a review of related literature for developing countries).

Up to now, most authors estimating the returns to education in Africa have relied on methods of estimation that do not adequately deal with the endogenous nature of schooling. Hence, estimates of returns to education could be biased. Some simply estimated average returns and returns at each level of education using the OLS framework[4]. Examples of such papers are Mwabu and Schultz(1996)[45] for South Africa, Knight, Sabot and Hovey (1992)[33] for Kenya, Aromolaran (2002)[6] for Nigeria. Other authors maintain the OLS framework but go a step further to account for the endogenous choice of sector of employment, correct for selectivity and control for omitted variables like ability.[5] Also, some of these authors

[4]It is possible OLS might not be biased in some cases as Grilliches (1977)[28] noted, unobservable and measurement biases may actually cancel out leaving the OLS estimates very close to the true return to education.

[5]See for example Kalzianga (2002)[37] for Burkina Faso, Glewwe(1996)[27] for private and

like Glewwe (1996)[27] make use of alternate estimators like maximum likelihood all in an attempt to improve estimates. However, even with this improvements, estimates of returns could still be biased due to reasons highlighted above.

Yet another approach to the returns to education estimation with some examples for African countries involves estimating returns based on surveys of firm based employees rather than households. (See for example Jones(2001)[34] for Ghana, Tekaligne,(1997)[61] for Zimbabwe, and Kahyarara et al (2004)[35] for Kenya and Tanzania .) As noted in Psacharopoulos and Patrinos (2002)[52], this methodology is problematic, as ideally a rate of return to investment in education should be based on a representative sample of the country's population not a minuscule group of workers with formal sector jobs. Firm-based employees are likely to be highly selective.

The only known papers prior to this, using the instrumental variable approach on data from sub-Saharan countries, are Kahyarara et al (2004)[35] for Kenya and Tanzania and Dabalen (1998)[16] for Kenya and South Africa. Both papers make use of instruments such as distance to school and parents education. However, results could still be biased because of common issues with the exogeneity of some of the instruments used and problems with the dataset for Dablen (1998)[16].

As with Dabalen (1998)[16], many papers using the instrumental variable (IV) approach have been critiqued. Staiger and Stock(1997)[59] argued that many studies using IV have weak instruments which led to even more biased estimates of returns to education. Carneiro (2002)[12] argued along similar lines, stating that most of these instruments are correlated with unobservables such as ability,

---

government sector workers in Ghana, Siphambe (2000)[56] for Botswana, Westergard and Nielsen(2001)[46] for Zambia and most recently Lassibille and Tan (2005)[36] for Rwanda.

and hence lead to inconsistent estimates of returns to education[6].

Finally in the recent literature, new general and country specific approaches to estimating returns to schooling have emerged, some general, others country specific. For example, the return to education is estimated when allowing for heterogeneous returns among individuals selecting into schooling based on these differences. Heckman and Li (2003)[31] used this new general approach in the context of China, making use of recently developed semiparametric methods to identify the parameters of interest. Another specific approach described by Hogan and Rigobon(2003)[32] uses unobserved shocks to individual education attainment leading to heteroscedasticity in education attainment across regions, to estimate the return to education for men in the UK using a large panel dataset.

3 General Theoretical Framework for analysis

As mentioned above, the literature on education has been approached from several theoretical perspectives. The most commonly-used framework, which will form the basis for my work, is the human capital approach. At the heart of the human capital model is the notion that education is an investment of current time and money in anticipation of increased earnings.

The human capital model of household or individual decision-making has its roots in Becker's 1964[11] model. However, I will be alluding to the simplified and tractable version of this model presented by Card (1995)[13]. This model is an endogenous schooling model and hence shows some of the biases that would result from OLS estimation of returns to schooling using a simple Mincer earning function. Let $y_i = \Omega(S_i)$ denote the expected level of earnings an individual i

[6] In section 6.2 arguments are presented for the validity of the instruments used in this paper.

would receive if he or she acquires schooling level S_i. Furthermore, I assume that the individual's utility function $U(.,.)$, is a function of level of schooling S_i and average earnings, y_i. I also assume individuals maximize their utility functions by choosing their level of schooling S_i. The utility function takes a simple form

$$U(S_i, y_i) = log(y_i) - \psi(S_i) \tag{1}$$

$\psi(S_i)$ is an increasing weakly convex function representing the disutility or costs from schooling[7]. Earnings y_i in this simple model are solely a function of S_i. I rule out other benefits from education, considering only the private benefits and assume individuals earn nothing while in school and y afterwards [8]. If I also assume individuals discount their stream of future earnings at rate r, then a discounted present value objective function on earnings over years of school for individual i, sets $\psi(S_i) = rS_i$ [9]. Hence, if individual i chooses schooling level S to maximize utility, then an optimal schooling choice would satisfy the first-order condition

$$\psi\prime(S_i) = \Omega\prime(S_i)/\Omega(S_i) \tag{2}$$

in which I am equating marginal benefits of schooling with marginal costs of schooling. I assume the cost/ taste for schooling $\psi(S_i)$ differs across individuals and the economic benefit which I represent as marginal returns $\Omega\prime(S_i)/\Omega(S_i)$ also differs across individuals. Then it follows that there is individual heterogeneity in the optimal schooling choice. Card (1999) gave a simple specification of this heterogeneity.

$$\Omega\prime(S_i)/\Omega(S_i) = b_i - k_1 S_i \qquad (k_1 \geq 0) \tag{3}$$

[7] $\psi(S_i)$ can be strictly convex if the marginal cost of each extra year of schooling rises more than the foregone income for that year.

[8] This assumption implicitly rules out part-time students.

[9] See Card (1999)[14] and Willis (1999)[62] for details on how this was derived.

$$\psi\prime(S_i) = r_i + k_2 S_i \qquad (k_2 \geq 0) \tag{4}$$

Here $\Omega\prime(S_i)/\Omega(S_i)$ is the marginal return to schooling and $\psi\prime(S_i)$ is the marginal cost of schooling and both b_i and r_i are random variables with mean $\bar{b}$ and $\bar{r}$, while k_1 and k_2 are nonnegative constants. In the above specification, optimal schooling choice is linear in the individual-specific heterogeneity terms. Given equation 3 and 4, the optimal years of schooling can be determined

$$S_i = \frac{b_i - r_i}{k} \qquad (k = k_1 + k_2) \tag{5}$$

and integrating equation (3) helps to recover a log earnings function

$$logy_i = \tau_i + b_i S_i - \frac{1}{2} k_1 S_1^2 \tag{6}$$

Here τ is the person-specific constant of integration. The inclusion of this allows for heterogeneity in earnings that arises from factors like ability independent of schooling levels. Equation (5) and (6) are sometimes estimated in schooling studies when estimating returns to education. However, many researchers exclude the non-linearities and heterogeneity terms in these equations and use a schooling earning system as follows:

$$logy_i = \alpha + \Phi C_i + \beta S_i + \epsilon_i \tag{7}$$

$$S_i = \lambda_0 + \lambda_1 Z_i + v_i \tag{8}$$

Here C_i and Z_i are vectors of explanatory variables, ϵ_i and v_i are uncorrelated error terms, α and λ_0 are the intercept terms and β is the return to education/schooling.

The Mincer earning function is compatible with equation (7) as the $C's$ could simply contain variables like experience, $(experience)^2$ and other exogenous factors affecting earnings, standard to the Mincer functional form. I intend on using variants of equation (7) and (8) in my estimation analysis.

4 Description of datasets

In this paper, I made use of two datasets highlighted below:

4.1 National Consumer Expenditure Survey

The National Consumer Expenditure Survey (NCS) is a cross-sectional survey organized by the Federal Office of Statistics (FOS) in Nigeria. The survey years I have are 1985, 1992 and 1996. These surveys cover 9317 households in 1985, 9697 households in 1992 and 14395 households in 1996. These surveys are supplemental modules of the National Integrated Survey of Households (NISH) which is run in line with the United Nations Household Survey Capability Program. This survey sample was drawn randomly from all the states in Nigeria in 1985, 1992 and 1996. The NISH sampling design is a two-stage replicate sample method, which is a common random sampling procedure. Data from these three surveys are comparable as the same sampling procedure was used in the three surveys. The sample size was larger in 1996 because the FOS had less financial constraints and could survey more randomly chosen households especially in the rural areas.

The NCS data set is appropriate for the analysis since it consists of detailed information on households' expenditure, household head income, location and other household characteristics. Also its data covers a 15 year period allowing us to test for time differences in returns. The main drawbacks of this dataset are, first, that all other variables such as gender, level of education, earnings and age, are available only for household heads FOS[21]. Second, the key variable for analysis is reported in education levels (e.g, primary education, secondary education, etc.) and not in years of household head's education[10]. Third, it appears urban areas

[10] The potential problem of overstating amount of schooling when level of education is reported

were oversampled in 1985 and 1992.

Due to the over sampling in the NCS dataset and other limitations of the dataset mentioned earlier, I focus attention on the second dataset I am yet to describe, in precisely estimating the returns to education. However, to test the null hypothesis of no time differences in returns to education, I would make use of the NCS dataset[11].

To ensure that the data are comparable over time and across regions, as is necessary when using income data, monetary variables were deflated to base year prices. Also, regional price differences were corrected for by making one state in the country a base and data from other points in the country were deflated to the price level of the base point[12]. Finally, to improve survey estimates, a standard weighting procedure computed at the World Bank was used. This is well described in FOS (1999) [21].

4.2 General Household Survey (GHS)

The second dataset used is the General Household Survey (GHS). The GHS is one of the major sample surveys carried out under the National Integrated Sample Survey of Households (NISH) program of the FOS in Nigeria and also makes use of a two-stage replicate sample design. It is the only survey in Nigeria that resembles the Living Standards Measurement Survey (LSMS) of the World Bank in terms of variable coverage. The federal office of statistics in Nigeria conducts this survey yearly and data are collected from randomly selected households during the

was checked by using data from the General Household Survey (GHS) of 1996/97 as a comparison for the NCS 1996 data, indicating no such problem.

[11] The NCS dataset covers both the 80s and 90s. Hence, it can be used for time comparisons.

[12] Deflation was done by FOS separately for both urban and rural areas. Lagos state was the base point and separate deflators were computed for food and non-food items.

four quarters of the year.[13] A drawback of the survey is that different households are surveyed in each survey year. The survey periods I use are 1997/1998 and 1998/1999. I have data on 32024 households in 1997/98 and 24889 households in 1998/99.[14] The part of the GHS I am most interested in is the Labor Force Survey (LFS), which is conducted as a part of the GHS. This data set, although only available for 2 consecutive years, unlike the NCS dataset, offers information not on household heads alone, but also on all other members of the household. For example, I have information on the education of each member of the household not only by level, but also by years of schooling. I will explore the range of this data set in answering the main question.

5 Initial Data Analysis

Before highlighting the empirical strategy used to answer each of the questions earlier stated, it is useful to review some descriptive statistics. Table 1 presents summary statistics of some important variables. It is important to recall that the GHS survey contains data on every household member, whereas the NCS for the most part gives only information on the household head so its summary statistics differ substantially for sex, age income and so on. Also, one cannot help noticing the drop in income post 1992. However, mean income in 1992 was high due to the temporary rise in oil prices during the gulf war. Mean income in Nigeria had been falling steadily over the 80s and only rose in 1991/1992 due to this temporary boom. The steady fall in mean income in the 80s was due to the economic recession after the collapse of petroleum prices in 1980. The downward trend in mean income is however not fully consistent with the general trends in GDP per capita over the

[13] Note different households in each enumeration area are interviewed in each quarter.

[14] For the first quarter of 1998/99 the data set was not available.

same time period in Nigeria(see Figure 1) (even though GDP per capita fell in 1980 yet it began to rise in 1986 but mean income as documented by FOS fell steadily until 1991). Furthermore, this drastic fall in income, is consistent with the finding by the World Bank(1996)[63] and Okojie (2002)[49] of an over 300% increase in poverty incidence (from 12% in 1980 to over 50% in 1996).

Table 2 summarizes mean incomes over time by education levels using the NCS and GHS data set. No education implies less than complete primary education and primary education indicates less than complete secondary education, but at least primary education while secondary education indicates less than a higher degree, but at least complete secondary education. There a few things worth noting from this table. First, the higher the education level the more the drop in income over time. This result is compatible with anecdotal evidence pointing to the rise in the educated poor in the early 90s in Nigeria.

Another point worth nothing is that though income rose in the early 90s due to the short oil boom in the early 90s, only the uneducated benefitted from this boom in terms of income rise from 1985 to 1992. However, this kind of result is very compatible with a Nigerian Dutch disease or resource curse story in which an economy totally dependent on a natural resource experiences a boom, and people leave productive work to rent-seek. Here, benefitting from the boom in terms of income increase would have less to do with education than with governmental connections and social networks. However, one can note that immediately after the oil boom ended post 1992, the least educated had the biggest fall in income. These preliminary results are interesting and call for further investigation.

6 Estimation techniques

I will first summarize the methodology used for adequately answering the two questions I posed previously. Subsequently, I describe the instrument used.

Question 1: To answer the question on what are the returns to education in Nigeria in the late 90s, only the 1997/98 and 1998/99 GHS survey data were used. This survey covers the whole labor force and contains more information than the NCS. Also, wage and schooling information is more precisely stated in this dataset and this survey is for the late 90's, which is the period I am estimating returns to education for.

Some simplifying assumptions on the endogenous schooling model were imposed. These are:

1. Log earnings are linear in schooling.
2. There is individual variation in ability and earnings.
3. There is a correlation between the determinants of schooling and the determinants of earnings. This means cov(S_i ,v_i)$\neq 0$ (It is this correlation between the determinants of schooling and earnings that would still make OLS biased even in this simplified case).

As a benchmark, returns to education were first estimated using OLS on a simple Mincer-type earnings function as in equation (9).

$$log(y_i) = \alpha + \lambda S_i + \phi X_i + \kappa X_i^2 + \rho D_i + \epsilon_i \tag{9}$$

Here X_i is experience of individual i and D_i are all other possible exogenous/control variables including dummies for year and regions, for individual i, gender, cohort dummies and so on.

Subsequently, equation (9) and (8) are estimated to derive the returns to schooling using the instrumental variables(IV) approach. This method hinges on finding observable covariates affecting schooling but uncorrelated with the ability factors or other possible omitted variables. These covariates become the instruments that are used in a two stage least squares (2SLS) estimation of returns to schooling. For completeness, yearly estimates of returns, and estimates pooling the data of each survey together are presented. The returns to education are estimated for the whole working population. However, estimations restricting the sample to those above age 22 (average age when college education is completed) do not change the results. Also, issues of potential selectivity are addressed.

Question 2

To test the hypothesis that there are no time differences in returns to education in Nigeria, the NCS datasets was used. First, annual estimates of returns to education are derived using similar methods to those described above. Subsequently, the estimates are compared for significant differences. If estimates are significantly different, the null hypothesis is rejected.[15]

6.1 History and Impact of UPE

As precise identification and estimation of the returns to schooling depends on the instrument, it is important to clearly explain the instruments used to address the endogeneity of schooling. The potential instrument for schooling is length of exposure to free primary education. The idea of using exposure to the UPE as an instrument originated from the paper of Osili and Long (2003)[50] on the impact of education on fertility in Nigeria. Using a difference in difference approach similar

[15] As the NCS datasets and GHS are quite different, the two datasets were not pooled together for any estimation. However, all estimates are presented in stating the results.

to Duflo (2001) [18], Osili and Long(2003) identify a clearly significant impact of the program on primary school attainment over the period of its implementation across regions.

The UPE was a nation-wide program designed to increase educational attainment by providing tuition-free primary education with different periods of implementation across states/regions. This program was first initiated during the colonial period in Nigeria. At this time, Nigeria was divided into 4 regions, the Northern, Western, Eastern and the federal capital, Lagos. The first region to implement free primary education was the former Western region. The regional implementation of this program was not linked to this region's riches or being most favorable toward more education, but determined by a choice of policy by the regions' colonial officer in charge of education. This officer believed strongly that free education was the only way the western region could catch up to the western world. It is also noted historically, that he convinced the regional leader of the west to implement the program. Hence, the policy reflected his own preference and not the preference of the populace of the region as in a democracy (see Fafunwa (1974)[20] and Adesina (1988) [1] for the history of education in Nigeria).

The program started on the 17th January 1955. In January 1957, the Lagos region that used to be the capital region of the federation initiated the program. Subsequently, in February 1957, the regional government of the Eastern region also started the program. Hence at this time, the only region not involved in the program was the North.

However by 1960, the Eastern region decided to restrict the free education program to only the first two years of primary school. In 1963, Nigeria became a republic and in the same year, the Mid-western region was carved out of the western region and was no longer part of the free education policy of the Western region. On the 6th of September 1976 the head of state (Nigeria was under military rule during this period) launched the mandatory program for the whole country, formally naming it UPE.[16]

The Program came to an end in 1981 during the first civilian government when the responsibility of education financing moved from the federal government to the state. However, for the duration of the civilian regime (1979-1983) free education was extended to all levels of education in states won by the United party of Nigeria (UPN) in the 1979 gubernatorial election[17].

Figure 2 is a timeline of program implementation and Figure 3 is a snapshot of the variation in free education across regions over time caused by the program. It is this variation in cohorts exposed to free education, over time and across regions, that I exploit as an instrument for school attainment.

6.2 Why the UPE makes a good instrument

Does the program constitute a good instrument? We know that any good instrument must satisfy three characteristics.

First, a good instrument must be relevant. The relevance/importance of the free primary education program for school attainment and education development in Nigeria has been documented extensively by several authors. For example, Nwanchukwu (1985)[47], Casapo(1983)[15] and Osili and Long (2003)[50] successfully highlight the impact of the UPE program on school attainment. Other descriptive data point to the impact of the program. By 1947, the Eastern region

[16]In this paper the instrument will be called UPE

[17]These states include all the states in the western regions and also Bendel state from the South South region which is presently divided into Edo and Delta

of Nigeria had the highest primary enrollment of 320,000, followed by the West at 240,000 and the North 66,000. Between 1947 and 1957, there was 212% increase in primary enrollment in North, 278% in the East and a 309% increase in the West. The faster growth in enrollment in the West, even though population growth was similar across the regions, has been attributed to this program. More specifically, the rise in primary enrollment from 475,000 in 1954 in the Western region to 800,000 by 1956 one year after the program's implementation, is attributed to introduction of UPE.

In the 70s, the rise in primary enrollment from 4.4 million in 1974 to 14.5 million by early 1982 was attributed to the reintroduction of the program. Specifically there was a 124% rise in primary enrollment from 1975-76 when to program was implemented to 1980-81, in contrast to an increase of only 4.5% from 1980-81 to 1984 when the program ended (see figure 4). This evidence provides further support for the impact of this program, especially as growth of the population of school age children was quite steady over this period (1960-1980).

Another possible argument that the jump in enrollment was caused by the oil boom in the 70s does not hold as the oil boom started in the early 70s and the significant rise in enrollment was in the mid 70s coinciding with the implementation of the program nationwide. Apart from this descriptive evidence, using a difference in difference approach similar to Duflo (2001)[18], Osili and Long(2003) identify a clearly significant impact of the program on primary school attainment over the period of its implementation.

Second, a good instrument must satisfy exclusion restrictions and the UPE program meets this criterion too, as the only means through which the program affects income is exclusively through its effect on schooling. This condition could be violated if the program implementation affected the quality of teachers and their present income. This possibility was investigated, noting no such relationship. Also, the possibility of the temporary fall in quality of education during the phase in period of the program affecting an individual's present income was ruled out upon investigation, using simple tests similar to those in Duflo 2001[18]. For example, I find no systematic correlation between teacher-student ratios and program implementation over time.

Third, a good instrument is strictly exogenous, meaning it is not correlated with any unobservable in the earnings equation. This criterion is the hardest to prove. However, I argue that this instrument is exogenous for many reasons. First, the implementation of the policy was not as a result of a democratic choice, and hence to a large extent does not reflect popular preferences. As the program was implemented in a colonial and military setting, program implementation across region and time reflects various commanders' preferences[18]. Besides, the initial phase in of the program was not in any way related to the western region having a higher value for education than the east or midwest. In fact prior to the program implementation enrollment rates were highest in the eastern region of the country. Also, the program was the idea of an officer in charge of education in a particular region, who had a particular ideology or preference.

A clear example of how an individuals' preference drove policy implementation is the case, of then military ruler Olusegun Obasanjo who made the program nationwide in 1976 when he assumed power. Though the program was scrapped at the end of his regime, he has once again reintroduced the program in 1999, (when oil prices were at its lowest in more than 10 year) over 20 years later,

[18] It is possible to tell a story where commanders try to meet people's preferences but this can be ruled out in the Nigerian case based on historical facts leading to program implementation.

when he was sworn in as the first civilian president of Nigeria after decades of military rule, further extending the program to the first three years of secondary education. Unlike many other past leaders, he is convinced this program is essential to Nigeria's educational progress and shares a similar ideology to the officer who first suggested the idea.

Detailed documentation on the history and administration of the program confirm that timing of implementation was arbitrary and not influenced by resource booms or regional/political factors. This means the choice of location for the initial implementation and length was not linked to non-random regional factors. For example the phasing in of the program in the 50s was not linked to a resource boom in the west neither was the collapse of the program linked to the fall in oil prices but a shift of handling education to the state government. Based on the above arguments and other research into the program implementation, I argue the UPE instrument is exogenous.

Finally, as mentioned earlier, recent studies have critiqued the instrumental variable approach for several reasons such as the instrument being weak with insignificant estimates and estimates being inconsistent as they are correlated with unobservable ability in the wage function. In the case of our instrument, ability does not affect exposure to the free primary education and in general the instrument is not weak.

6.3 Construction of the instrument

As stated in the introduction, the UPE instrument is constructed based on the length of exposure of an individual to free education. The argument here is the longer an individual is exposed to free education, the higher the years of school attainment. The length of exposure to the UPE program makes a good instrument for several reasons. First, for every extra year of free education a parent can get for a child, the lower is the cost of achieving any higher levels of education. Furthermore, if parents, due to lack of knowledge, are apprehensive of western education, as was the case in Nigeria (see Ozigi & Ocho (1981)[48] for the Northern Nigeria case), the longer their children are exposed to education, the higher the probability parents will appreciate its value and be willing to pay for further education. In constructing these instruments, length of exposure to free primary education, or length of exposure to free education, whether primary or higher, can be used. The estimation results using either alternative are not significantly different. However, for completeness, I constructed the instrument as exposure to free education.

It is important to note that Osilli and Long construct their instrument differently (see pp 14-16 Osilli and Long(2003)[50]. They focus only on the formal implementation of the UPE in the 70s. I focus on implementation of free primary education since the idea started in 1955. Furthermore, they limit their sample to women of two cohorts: those born between 1958 and 1963 (age 13 to 18 when the program started) and those born between 1970 and 1975. I consider both men and women truly exposed to the program of free education in its different phases of implementation from 1955 on. I however tried to replicate their estimation of the impact of the UPE using the GHS dataset. Both estimates, though different, are not statistically different. In both cases the estimates show the strong impact of the UPE on schooling.[19]

[19] The estimate of UPE impact (0.65) I tried to replicate was from table four of Osilli and Long 2003[50]). My estimate was 0.54, but one can expect to find slight differences as different datasets are being used. They combine 1990 and 1999 of the Demographic household survey (DHS) while I am using 1997-1999 of the GHS. They also have control variables like religion which are not in the GHS dataset.

The instrument is constructed based on an interaction between year of birth and location. For example, individuals born in the north in 1970, were six years in 1976 when the program started nationwide. Since the program ended in 1981, such individuals would have been exposed to free primary education for six years. The variation in the instrument comes from different cohorts in different areas of the country being exposed to free education for different lengths of time.

The instrument is expected to capture individuals' exposure to free education, but if individuals lived in parts of the region where schools did not exist during the period of program implementation, then such individuals were not actually exposed to free education because it was not an option for them. Several authors have written on changes in the education sector in Nigeria and highlight this problem with the implementation of the UPE. Hass et al(2003)[29] explicitly state that during the UPE implementation there was a recognition that those receiving a primary education tended to be male, urban, well-to-do, and resident in a south-eastern or southwestern states in Nigeria. The reason for this bias was the location of most primary schools in selected urban areas and different ethnic beliefs about sending girls to school.

The lack of schools in towns and villages was common in the early periods of the program implementation, especially in the late 50s to early 70s. Even in the 80s, some rural areas of the north lacked primary schools. Hence, constructing the instrument without taking into account the fact that many people did not have schools in their towns and villages though in a region with program implementation can attenuate the impact of the instrument if the sample is small or contains fewer people truly exposed to free education as in the initial phase in of the program. In the case of the sample size being small, the issue is noise. However, if the effect of the instrument is strong enough not to be attenuated by the noise associated with small sample sizes, this will not be a problem. In terms of the other condition, the issue is wrongly assigning exposure to a large number of observations who were not really exposed to schooling and hence did not try schooling simply because schools did not exist though schooling was free. In this scenario, the instrument would be weak.

The issue raised above is relevant to the analysis using the NCS dataset since the 1985 data years of the NCS naturally contain a higher proportion of observations who were in school before or during the early phases of the program when true exposure was limited. Also, the NCS data set has a relatively smaller sample sizes compared to the GHS and noise could be an issue especially for the 1985 dataset since the sample really exposed to free education was limited. To get around potential problem when using this 1985 dataset, as I do not know exactly which towns in the regions did not have schools, I do two things. First, I focus on specific cohorts, like Osilli and Long [50], to estimate the returns to schooling. I know those in the 1970s phase of the UPE are not in this dataset, so I focus on the cohorts born just before the first phase of the program in 1955 and those who would be above primary school when the program started. Second, I consider only the urban areas, where exposure was more likely. This is a credible way to do this as several authors writing on the spread of education in Nigeria up until 1980 have shown that the main factors that prevented people from going to school once it became free were inaccessibility in many rural areas and also customs, in the case of girls (see Fafunwa (1974)[20], Ozigi & Ocho, (1981)[48] and Mazonde I(1995)[38]). By considering the urban area only in 1985 I can capture more of those who were truly exposed to free education.

In the 1992 and the 1996 dataset I do not need to focus on one specific group or on urban areas only. The reason is that true exposure was higher in the 70s phase of the UPE program. However, the 1996 dataset would have more observations truly exposed to the program than the 1992 dataset because more of the beneficiaries of the 70s phase in of the program would have entered the workforce in this dataset vis a vis the 1992 dataset.

The above assumption is not used in the construction of the instrument using the GHS datasets as the sample contains more of the younger cohorts. These cohorts had true exposure to the UPE as more primary schools were available to these cohorts (see Yoloye (1999)[64] for information on schools expansion in the late 70s). Besides, the dataset is very large and the potential effects on the instrument stemming from the earlier cohorts previously mentioned, would be attenuated.

Lastly, a possible issue that could arise when using these instruments on the present data is migration. This potential problem exists because the data set does not contain information on where individuals were born or went to school but on individual's present location. Individuals could possibly be located in places different from where they went to school and the instrument potentially could be inaccurate for this group of people. In that scenario, our instruments might be weak. However, this is not the case in Nigeria. Most movements are within states from rural to urban areas and not across states which could affect the validity of our instrument. As was explicitly documented in FOS (1999) [23], and FOS (2000)[24], 95.3% and 95.8% of people were still living in the state where they were born. Moveover, the 4.2% who migrate mostly move within the same region. Hence, potential effects on the instrument should be negligible.

7 Estimation and Results

7.1 Estimation of returns to education

Equipped with instruments described in the previous section, the returns to education were estimated following the empirical strategy outlined in section 6. First, a standard Mincer equation like equation (9) was estimated using OLS. Apart from standard variables in the mincer equation, other controls such as cohorts, sector, higher powers of age, sex and state were used in the estimation[20]. However, instead of using imputed experience which is usually computed using a standard formula, I instead use age. The rationale for doing this is linked to the implicit flaws in using the standard formula for calculating experience especially in developing countries[21]. Besides, using age is consistent with most of the recent relevant literature[22]. Examples of papers using age instead of experience include Angrist and Kruger (1991)[3], Harmon and Walker(1995)[30], Maluccio (1997)[40], Ashenfelter and Rouse (1999)[9]. Furthermore, age is a good proxy for individuals' experience and is usually accurately measured in the data.

Table 3 is a summary of the results of the estimation process using both OLS and 2SLS. Estimation was carried out including all controls, clustering by age and correcting for potential heteroskedacity. The first stage result points to the impact of the program on attainment. The reduced form estimates point to the direct impact of the program on wages. What is striking from Table 4 is the low return to education. In 1997/98 the return to an extra year of schooling is 3.7%

[20]Slight variation in the cohort and state dummies were made in some of the estimations to avert serious collinearity problems.

[21]Standard Experience formula =(age-years of schooling - 6)

[22]In many papers, age is used instead of experience when actual experience is not in the data. In Card (1999)[14], the author summarizes in tables the recent studies on estimating returns to education. More than half of the studies use age rather than experience

while in 1998/99 the return to schooling is 3.0%. Using the pooled data from both years the return to an extra year of schooling is 2.7% . Another quite unexpected finding is that the OLS and IV estimates are very similar. OLS is downward biased when using the individual years and upward biased when pooling the two years together but the estimates are not statistically different.

The above results do not categorically establish the returns to education to be very low for everyone in Nigeria for the years in question. This is because returns to education can be heterogenous. Recall that all that is being estimated is the average for the entire labor force. Hence, it might be useful to try to break down the population into groups to see if the results would change drastically or if the low returns to education can be isolated for a subgroup in the population. In the next section, returns to education will be estimated for subgroups of the population as both a robustness check on the results and to relate the results to particular groups in the country. However, prior to these checks, the question of time differences in returns to education is investigated.

7.2 Result of Estimation using the NCS: Time differences in returns to education

Table 4 summarizes the OLS and IV estimates of returns to education using the NCS dataset. Like in the above estimation, corrections where made for potential heteroskedacity and controls where included. Recall that for 1985, returns to education is only estimated for the urban sector of a subgroup of the population. The estimates for 1992 and 1996 are averages for the whole population. The results in table 4 are striking. One can notice that for all three years, the estimate of the impact of UPE is similar. However, the returns to education is at an all time high of 11% for an extra year of schooling in 1985, falls to zero in 1992 and rising again to 5.1% in the mid 90s. It is possible to argue that the estimates for the 1985 are not comparable to the other two years because it provides estimates for a particular subgroup of the population. This is not the case as computed estimates for returns to education for this particular subgroup in 1992 is negative and more importantly insignificant (p-value of 0.8) and estimate of returns in 1996 for this group is 0.031 which though less than 0.051 (the return for the whole population), is not statistically different from it. Thus, similar trend in returns exist even when looking at this subgroup over time. To test the null hypothesis, the estimates of returns to schooling are compared using t-tests. The null hypothesis of no time differences in returns to education is rejected at a 5% significance level. The average magnitude of the difference in returns to education between 1985 and 1996 is about 8% for every extra year of schooling which is sizeable.

The pooled estimate for the impact of the UPE on schooling is similar to the individual years. However, the estimate for returns to schooling is insignificant and negative . This is expected since the whole 1985 dataset is included in this pooled estimate and as earlier mentioned in section 6.3, the instrument by construction does not do a good job of capturing true exposure to free education in the mid 80s unless for a specific subgroup.

It is also of note that the OLS estimate is upwardly biased or downwardly biased compared to the 2SLS estimates depending on the year being considered (see table 4). However, the estimates using OLS and IV are quite similar in terms of magnitude but are statistically different in all cases. This finding points again to the randomness of OLS estimates in the presence of omitted variables and endogeneity.

Also, though it seems that returns to schooling have fallen post 1996 when looking at estimates using the GHS for 1997-1999, it is important to note that the NCS data and GHS datasets are not directly comparable as one dataset considers household heads and the other contains information on the entire labor force.

The zero returns to education in 1992 would seem implausible to anyone not knowledgable about the peculiarities of the Nigerian scenario. However, it is important to note that we are estimating the average returns for the whole population of households heads. Results for a specific group in the population could be more informative. Besides, the year 1992 was in a period marked with positive oil shocks and it is common knowledge in Nigeria that benefitting from the oil boom wealth during the military rule depended more on social networks than on educational attainment[23]. In fact, sociologist and political scientist have written consistently of the undue importance of social networks, regional control and corruption on wealth distribution in Nigeria during the military rule, especially in periods of oil booms [24]. Moreover, zero returns is not peculiar to Nigeria alone. Glewwe (1996) [27] also finds zero returns to schooling among private workers in Ghana when using a maximum likelihood estimator

8 Robustness checks

One issue one could raise, based on the above results, is centered on gender. In Nigeria, many claim that gender affects wages and it is possible that males and females have different returns to education. Also in Nigeria, the sector of the economy where an individual dwells and works can affect earnings. Hence, individuals in the rural and urban areas could have different returns to their education. Be-

[23] Prior to democracy in Nigeria, social networks and educational attainment were not correlated.

[24] Some recent books touching on issues like these are Soyinka (1997)[57], Suberu(2001)[60].

sides, the literature clearly documents the difficulty in estimating income in the rural areas because people work mainly in the informal sector (farming, fishing, animal rearing) and it is very hard to isolate wages for individuals in these households. This problem of getting precise wage estimates for individuals in the rural areas is one reason to estimate returns separately for rural and urban areas and focus more attention on the average returns to education in the urban areas. Using both OLS and the IV estimator, returns to education were estimated by gender and sector, using the pooled data and including all appropriate available controls. Table 5 provides a summary of the returns to education pooled estimates using the GHS dataset by sector and gender with robust standard errors. This table provides some interesting results. First, the impact of the program on men's and women's school attainment was the same. Second, the return to education for men is twice the return for women. This is an interesting result since based on economic arguments, return to education for women is expected to be higher than return for men. However, some gender studies on Nigeria have shown that many women earn less than their counterpart with similar education in the workplace. Also, it is commonplace in Nigeria to see educated married women settling for jobs with lower pay but more flexible work hours just to make room for household responsibilities. Looking at the second part of table 5 one also notices higher returns to education in the urban areas than in the rural areas, which is expected. Despite the differences across groups, these results are compatible with earlier results. Returns to education in Nigeria was still below a 5% increase in income for every extra year of schooling in the 90s. These estimates are clearly on the low side relative to estimates from other countries.

Another argument that can be made is that estimating the returns to education across sectors, or solely focusing on the urban sector, does not fully deal with the problem of precisely estimating individual income which is necessary for a valid estimate on the returns to schooling. Many people in the urban areas are still involved in the informal sector, and for these individuals accurately estimating their earnings accounting for family free labor could be prone to error.[25] Hence as a robustness check, the return to education was estimated for households containing a single individual. Here the problem of possibly overestimating the returns to education because of inability to adequately untangle individual earnings is removed. Table 6 column two and three is a summary of the returns to earnings for the single-individual households using the pooled data from the GHS. The impact of the instrument on schooling is similar to the previous analysis. Also, the return to education for this group is higher than the average for the population but not statistically different. Again the main results still holds. The returns to schooling in the late 90s were below 5% for every extra year of schooling. Similar robustness checks were carried out using the NCS dataset to check that the time differences result earlier noted holds in at least a subgroup of the population, noting similar trend.[26]

In line with the question of accurately identifying the returns to education in the late 90s in Nigeria, another robustness check is to re-estimate the returns to education dividing the sample into wage earners and self employed[27]. The argument is that returns to education can only be properly estimated for wage

[25]It is important to note that for both the GHS and the NCS surveys, survey staff are trained to tackle this problem of measuring individual income in the informal sector using standard computations. However, these computations may still be prone to errors.

[26]Returns to education was estimated for male household heads in 1985, 1992 and 1996 and across sectors for each years using applicable controls noting similar trends of highest returns in 1985, drop in 1992 to insignificant levels and rise in returns to education by 1996.

[27]This exercise could not be carried out for the NCS survey years as one is unable to clearly identify wage earners in this data set.

earners as wages are to a large extent a measure of productivity. The second part of Table 6 is a summary of the returns to education estimate for wage earners and self-employed, pooling the 1997/98 and 1998/99 dataset together[28]. The first interesting finding is that though return to education is higher for wage earners than self employed, there is not much difference in returns to education between the wage workers and the self employed. This finding is contrary to the theory that education basically serves as a signal and really does not embody human capital[29]. The results are again consistent with earlier results showing low returns to education less than 0.05 in the late 90s in Nigeria.

Another robustness check is to estimate the returns to education by cohorts. The argument is that individuals are at different stages of their life cycle and it is possible returns to education differ. These potential cohort effects were controlled for in earlier regressions noting no significant effect of cohorts on returns to schooling. However, for completeness, the returns to education was estimated for different birth cohorts exposed to the UPE. The cohorts are constructed based on decade of birth. As considering each cohort separately reduces the precision of the estimates as the sample size drops significantly, the instrument is constructed similar to when using the NCS dataset.[30] The estimates of the returns by cohorts is summarized in Table 7. These results are similar to the previous results in Table 3. More importantly, return to schooling is similar across cohorts (not statistically different in most cases) although returns are highest for the youngest

[28]Due to the small sample size of wage earners, problems mentioned earlier in the paper with respect to the instrument when sample size is small can crop up. To get around this potential problem, estimates are derived only for states where impact of the 1976 phase in of the program would be strongest.

[29]The estimates for wage earners and self-employed are not significantly different.

[30]In constructing the instrument in this case, non-exposure is assigned to all rural areas in states were schools were limited in the past based on information from FOS and focus is placed only on regions with exposure for earlier cohorts.

cohort. One can infer from this finding that any possible difference in quality of education across cohorts over time had minimal impact on returns.

Another possible form of bias that can affect precisely estimating the returns to schooling is selectivity. This issue would be addressed in the next section.

The findings in this section have confirmed that average return to education in Nigeria in the 90s is low. In addition, this low return is driven more by women and rural households. These results also indicate that OLS estimates of returns to education are biased downward in most cases for the subgroup analysis. This finding is not peculiar as many authors have found higher IV estimates of returns to schooling than OLS (See Card 1999[13]. What is interesting is that this difference in IV and OLS estimates are not apparent when looking at the whole population (see table 3). Also, all these subgroup estimates of returns to schooling are in general lower than what some past researcher have reported to be characteristic of Africa and developing countries in general (see Psacharopoulos and Patrinos 2002[52]).

8.1 Correcting for Selectivity

For the question for which precise estimates of returns to education was sought, a potential source of bias, common when estimating earning equations, is self-selection bias. That is, if individuals can choose whether to be within the work force based on individual self-selection, then the schooling variable will be a dependent rather than independent variable. Thus, ordinary least squares (OLS) estimates of schooling will be inconsistent. One way to check and correct for selection bias based on the pioneering work of Heckman (1974 and 1979)[31] is to calculate the inverse Mills ratio, add it as an additional regressor in the earnings equation and run a simple OLS to see if its coefficient is significant[31]. This simple test of self selection was carried out and the coefficient on the inverse mill ratio was significant in 1998 and the pooled regression but not in 1997. Similar results were obtained when including the Mills ratio in the second stage of a 2SLS analysis using the instrument. However, in all cases the coefficient on schooling did not change significantly from its previous value without the correction see Table 8.

The above method has come under criticism for relying on unverifiable assumptions about the unobservable and functional form of the selection model to obtain identification. In addition, there are arguments that there are other potential sources of self selection not captured via this means. For example when estimating the wage equation, log of earnings ($logy_i$) is observed only for those working ($w_i = 1$). Hence, a correlation can exist between the instrument M_i and the error term for those working when conditioning on the instrument if the probability of being employed is correlated with schooling and hence the instrument (Angrist,1997[2]).

To address this potential problem and ensure identification, the propensity score was used. A general control for selection bias requires only the existence of a function $f(M_i)$, such that the error term of the outcome equation (ϵ_i) is independent of the instrument, conditional on working w_i and $f(M_i)$ (Angrist 1997[2]). However, for the propensity score to serve as a conditioning variable in the presence of selection bias, (ϵ_i) and selection status are assumed to be jointly independent of the instrument and also ϵ_i is independent of M_i.[32] This correction mechanism allows the population to be stratified according to their propensity

[31] Here one assumes that the error terms are jointly normal and independent of the instruments

[32] To see why these assumptions are sufficient to control selection bias when conditioning on propensity score see Angrist (1997), pp 106 [2]. Recent literature has highlighted that these assumption are restrictive.

scores so that the mean outcomes for each of the identified strata can be compared.

The implementation of this procedure requires three steps

1. First, estimate the propensity score of working as the fitted value of w_i regressed on covariates. I make use of both a probit and a linear model in this selection model estimation.

2. The next step is to derive the predicted value of schooling, using equation(8).

3. Then estimate equation(9) with other covariates, the propensity score and predicted value of schooling.

Table 8 shows the estimates of schooling correcting for selectivity using the mle with a Heckman correction model, Heckman two step estimation procedure and the propensity score correction with a linear and a probit model. These results support the results of the test of selectivity mentioned earlier. Selectivity is not an important issue in this analysis as comparisons between the 2SLS estimates of returns to schooling with controls are very similar to estimates after correcting for potential selectivity with most of the different models.

Identification is sought through the propensity score estimation using a probit model. Therefore the preferred estimate of average returns to education in Nigeria was 3.6% for every extra year of schooling in 1997/98 and 3.0% in 1998/99[33]. These estimates of average returns to education in Nigeria are lower than other estimates for other African countries including Aromolaran's(2002)[6] estimates for Nigeria.

[33]The pooled regression estimate was lower than the estimates for the cross-section. However, the estimates are not significantly different.

8.2 Importance of controls

The above estimation included appropriate controls in the wage equation. As an experiment, return to education is estimated on the standard mincer wage equation without controls. In this standard equation, income is a function of years of schooling, age and the square of age only.

Table 9 shows 2SLS and OLS results using the GHS data on a simple mincer equation. Comparing this results to earlier results with controls, it is clear that not including controls biases estimates. First, the impact of the program on school attainment is more than twice the estimates with controls (see Table 3). Similarly, the returns to education is about 1.5 percentage points higher without controls than with controls. This finding is also apparent if considering the estimation of returns to education without controls using the NCS dataset (see table 10) in comparison to earlier results in table 4. However, unlike in the GHS data result where the bias caused by not controlling is similar across years, the bias does not seem to be apparent in 1985 of the NCS and is huge in the pooled estimate and in 1996. This finding is important because if for example one had considered household heads solely in 1996 using the NCS dataset without any controls, one would conclude that returns to education are very high in Nigeria (0.13% for every extra year of schooling). Similarly, if one had pooled the NCS data together and estimated the returns for education without controls, one would have concluded that returns over 1985-1996 are relatively high at 0.088 for every year of schooling. This result would have been misleading. Also, if I used the GHS dataset without controls the conclusion would have been that returns to education in Nigeria are average (approximately 5.3% increase in income for every extra year of schooling). This conclusion of average returns is also different from the true result of low returns to education in Nigeria.

8.3 Comparison to other estimates for Africa

The returns to education in Nigeria are low but the question is whether this is a Nigeria phenomenon or there is a possibility returns are being over estimated for other African countries. Earlier on in this paper, recent papers estimating returns to education in most parts of Africa were highlighted. These papers had returns to education typically over 6% increase in income for every extra year of schooling. In fact most of these papers had returns well over 10%. To state specifically a few example, Jones (2001)[34] estimates returns in Ghana at 8.1% and Lassibille and Tan (2005) [36] for Rwanda estimates of returns range between 19.4 to 33.4%. Siphambe (2000)[56] estimates for Botswana are mostly in the range of 7% to 14% depending on the specification. Chirwa and Zgovu (2001) [65] estimate returns for Malawi noting returns between 5% and 10% for every year of schooling depending on group analyzed and Psacharopoulos (1994)[51] estimates the returns to education for several African countries all with returns over 8% for every year of schooling. These few examples from the literature are in contrast with the preferred result in this paper. However, these results are similar to the result in Table 9 and 10 where a simple Mincer equation with no controls was used.

A careful analysis of the above mentioned papers and others in the literature review revealed that minimal controls were used in most papers[34]. For example in estimating returns in Malawi, no controls for gender and location where included. Similarly for Rwanda and Botswana. It is important to note that most of these

[34]Glewwe (1996)[27] is one paper that uses adequate controls. Interestingly, He finds low returns to education, which is similar to the finding of this paper.

papers put in some type of control variable though minimal [35].

Adequate controls are needed to be able to attenuate omitted variable bias in estimating returns to education. Wages in most parts of Africa are affected by gender and the area of the country a person lives and as much as possible, these other factors affecting income need to be controlled for. Estimations using OLS and other methods that do not adequately deal with the endogenous nature of schooling coupled with inadequate controls will likely lead to biased estimates of returns to schooling.

9 Implications and Conclusions

9.1 Implications of significant time differences and low returns to education in Nigeria

The above results point to significant time differences in returns to education in Nigeria (See Figure 5). More importantly, the results showed that average returns to education were extremely low. Why do we care about these results?

First, low returns to education and marked differences in private returns to education over short periods of time can discourage investment in education (in terms of time differences, individuals may perceive investment in education as risky and invest less). This is crucial if education has large social returns and externalities. Furthermore, if education investments positively affects human capital and growth, then less investment in education cannot be beneficial. A clear indicator that individuals are investing less in education was reflected in falling enrollment rates and also a decline in quality of education noted in Nigeria over the 90s (see

[35]As mentioned in the literature review, many papers try to correct for selectivity but some lack basic controls like location, sector and gender.

Malik(1997)[39] and FOS 2000[22])[36].

Second, low returns to education in Nigeria can lead to individuals finding alternative investments (leading to fall in school enrollment). It can also lead to individuals who already have invested in education seeking seeking international markets where there are higher returns to their education or switching to rent-seeking activities. These three reactions to low returns to education were common place in Nigeria and many other countries in Africa in the 90s[37]. According to a study by the Geneva-based intergovernmental body, the International Organization for Migration (IOM), and the UN's Economic Commission for Africa (ECA) Africa lost 60,000 professionals (doctors, university lecturers, engineers, etc) between 1985 and 1990 (see Aredo(1998)[5]. Even though this is not a large chunk of professionals within Africa, it is still significant. Moreover, this form of emigration can only be a road block to the growth and development of a country. Hence, continued low returns to education in Nigeria compared to elsewhere is a sure stimulus for more of this kind of emigration if unrestricted.[38]

Lastly, as these results indicate, returns to education within the range of 2-5% for Nigeria, and most previous papers have estimated returns to education for other

[36]Although gross enrollment rose over the 90s at all levels of education, the enrollment rates for both primary and secondary education dropped significantly in the mid 90s and dropout rates rose dramatically. The decline in the quality of education over the 90s was linked to many factors amongst which are incessant strikes and school closing, a rise in teacher student ratios, change in secondary education system and inadequate school input, political instability and declining government allocation to education as the military government did not see education a priority. This downward trend has slowly been reversed with the change to civilian rule since mid 1999.

[37]Although it is widely documented that many Nigerian emigrated in the 90s, specific data relating to emigration from Nigeria is not available.

[38]Immigration to most of the western world from developing countries especially Africa became more difficult in the 90s and restriction by receiving countries have only tightened over time(less than 5% of visa applicants to developed countries especially North America and Europe are granted). These restriction have curbed emigration from Africa significantly. However, high visa applications up until now indicates individuals preference for immigrating.

African countries in the range of 5-15% (Psacharopoulos and Patrinos (2002)[52]) using OLS and other similar estimation techniques with few controls, there is a possibility that returns to education are being overstated for some other countries in Africa. This could explain why other Africans also question the economic value of their education despite high reported returns in other parts of Africa. Also, preference to immigrate is not just a Nigerian phenomena but a Sub-Saharan phenomena.

9.2 Conclusions

From the above analysis, it has been established using the unique instrument (UPE) that significant differences in average returns to education did exist in Nigeria between 1980 and 2000. More importantly, the estimates for average returns to an extra year of schooling recently in Nigeria are 3.0% and 3.6% for 1997/98 and 1998/99 respectively. Meaning that for every extra year of schooling, there is less than a 4% increase in wages. The results also indicate that return to education for men is twice that of women. These low estimates of returns are robust to other specifications (meaning estimates are not significantly different) and are lower than other estimates for Nigeria and other African countries using OLS and other estimation techniques.

The results also suggest that returns to education does differ substantially across sector and gender. Furthermore, the importance of including controls when estimating returns is highlighted in the results. OLS is biased but the direction and extent of the bias varies from year to year. However, estimates using the GHS dataset generally point to minimal bias in OLS estimates although estimates are statistically different in some cases.

I also find quite similar returns to education across wage workers and self-employed workers, in contrast to Aromolaran's (2000)[6]. However as the latter rightly noted, it is difficult to measure income for those in the informal sector as earnings attributable to physical capital or return for bearing risk might not be excluded when reporting income. Finally, I find similar results across cohorts, which suggests that the fall in quality argument cannot be the primary reason returns have fallen over time.

This paper contributes to the literature by providing more reliable estimates of returns to education in a west African country using the instrumental variable approach. Furthermore, the results show returns to education estimates in Nigeria that are lower than what is thought to be characteristic of Africa. This paper also provides evidence of time differences in returns to education in Africa. Time differences has not been extensively considered prior to now, but are substantial and merit further investigation. The results also emphasizes the importance of including control when estimating the mincer wage equation and the inadequacies of the OLS estimation of returns to schooling. Finally, several explanations have been sought for the changing demand for education, the increase shift to rent seeking activities and increased emigration rates from Nigeria over the 90s. The low returns to education in Nigeria suggest a reasonable explanation for these phenomena. These findings highlight the need to find instruments and re-estimate returns to education in other African countries.

The work presented here has limitations. The returns to education estimates are averages for the population or sub-groups in the population. As mentioned in the literature review, recent work points to heterogeneity of returns across individuals which has not been accounted for in this paper. Also, some of the results presented are based on estimates using the NCS dataset which contains information solely on household head and imputed years of schooling. It is also important to note that even though the instrument used in this analysis had very large effects on schooling and affected a wide group of people, as Angrist and Imbens 1999[4] highlighted, returns to education estimates using a treatment may only capture a weighted average of the returns to education for those affected by the instrument. Another limitation of this analysis is the assumption of a linear relationship between wages and schooling.

Finally, in terms of policy recommendation, the present Nigerian government should focus on understanding why returns to education are low and fluctuating. One way of doing this, is to sponsor further surveys and analysis aimed at understanding these findings.[39] In addition, the finding that the return to education for men is twice that for women raises important policy questions. Policy makers might consider whether programs to encourage educated women to actively take part in the workforce at their level of expertise would be useful. For example, a program subsidizing child-care for educated women or flexible work hour programs for women with children might lead to more women applying and actively seek jobs they are qualified for.

The question of why returns to education are quite low in Nigeria (especially so for women) and also, reestimating returns to education in other African countries using the IV strategy and accounting for heterogeneity, are interesting areas for further research.

[39] In another paper in my dissertation I address the roles of government and other institutions in explaining the low returns to education.

References

[1] Adesina, S. (1988) "The Development of Modern Education in Nigeria". Heineman Education Books, Ibadan.

[2] Angrist, J.(1997),"Conditional independence in sample selection models", Economic letters,1997,54, 103-112.

[3] Angrist, J. and A. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings", Quarterly Journal of Economics, 106, 979-1014.

[4] Angrist, J. and G. Imbens (1995)"Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity" Journal of the American Statistical Association, Vol. 90, No. 430, 431-442.

[5] Aredo, D. and Zelalem, Y., 1998, 'Skilled labor migration from developing countries: An assessment of brain-drain from Ethiopia', in Proceedings of the Seventh Annual Conference on the Ethiopian Economy, The Ethiopian Economic Association.

[6] Aromolaran, A. October (2002),"Private wage returns to schooling in Nigeria 1996-1999" Centers Discussion Paper No 849 Economic Growth Center, Yale University.Now in AFrican development Bank REview 2004

[7] Arrow, K.J, (1973) "Higher education as a filter" Journal of public economics Vol 2.

[8] Ashenfelter, O. and C.rouse, (1998), "Income, schooling and ability; evidence from a new sample of identical twins", QJE VOl 113 253-284.

[9] Ashenfelter, O. and C. Rouse. (August 1999),"The Payoff to Education", European Summer Symposium in Labour Economics, CEPR IZA, Ammersee, Draft.

[10] Bhagwati,J. and T.N Srinivasan 1977."Education in a job ladder model and the fairness in hiring role". Journal of public economics, Vol 7.

[11] Becker, G.S., (1964), "Human capital:a theoretical and empirical analysis with refeence to education". 2nd edn (Columbia University press New York) NBER.

[12] Carneiro, P. (2002), in James J. Heckman Xuesong Li,(2003) "selection bias, comparative advantage and heterogenous returns to education: evidence from China" National bureau of economic research, Working Paper 9877.

[13] Card, D.(1995), "Earnings, schooling and ability revisted", in; Solomon Polachek,ed.,Research in labor economics,Vol.14 (JAI Press,Greenwich,CT)pp,23-48.

[14] Card,D. (1999), "The causal effect of education on earning" In the Handbook for labor economics volume 3A Elsevier.

[15] Casapo, M. (1983). "Universal Primary Education in Nigeria: Its Problems and Implications." African Studies Review 26 (1): 91-106.

[16] Dabalen, A. (1998). "Returns to Education in Kenya and South Africa: Instrumental Variable Estimates." University of California, Berkeley (mimeo).

[17] Deaton A., (1985), "Panel data from time series of cross-sections". Journal of Econometrics, 30,109-126.

[18] Duflo, E. (2001) "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment" American Economic Review , vol. 91, issue 4, pages 795-813.

[19] Easterly, W.(2001) "The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics" , MIT Press, July 2001.

[20] Fafunwa, B.A.(1974) "History of Education in Nigeria." London: Allen and Unwin, (New edition in 1991).

[21] Federal Office of Statistics (FOS) (1999) Poverty Profile for Nigerians, 1980-1996, FOS: Lagos.

[22] "Annual statisitcal bulletin and digest",several years Federal office of statistics.

[23] "Annual abstract of statisitcs", (2000) Federal office of statistics.

[24] "Annual abstract of statisitcs", (1999) Federal office of statistics.

[25] Federal Republic of Nigeria ,(1981) National Policy on Education. Lagos: NERDC.

[26] Glewwe,P.(2002), "Schooling and skills in developing countries: education policies ans socioeconomic outcomes" Journal of econmic literatue VolXL, pp436-483.

[27] Glewwe, P.W.(1996). "The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment.". Journal of Development Economics.

[28] Griliches,Z. (1977), "Estimating the returns to schooling: Some econometrics problem" Econometrica Vol 45.

[29] Haas,M.E, Sunal, C.S., Sunal, D.W, Rufai,R., Inuwa, A., (2003) “Perceptions of Unequal Access to Primary and Secondary Education: Findings from Nigeria” African Studies Review, April.

[30] Harmon, C. and Walker,I (1995), “Estimates of the economic retuns to education for the united kingdom” American Economic Review 85;1279-1286.

[31] James, J. H., Xuesong, L.(2003), “Selection bias,comparative advantage and heterogenous returns to education:evidence from China” National bureau of economic research, Working Paper 9877.

[32] Hogan, V. and Roberto R.(2003) “Using Unobserved Supply Shocks to Estimate the Returns to Education” A previous version of this paper, “Using Heteroscedasticity to Estimate the Returns to Education” is available as NBER working paper 9145.

[33] Hovey, D.C, Knight J. B, Sabot,R.H.(1992) “Is the Rate of Return on Primary Schooling Really 26 Per Cent?” Journal of African Economies, vol. 1, issue 2, pages 192-205.

[34] Jones, P. (2001) “Are educated workers really more productive’ Journal of Development Economics 64: 67-79.

[35] Kahyarara, G. & F. Teal & A.Wambugu & M.Söderbom, (2004). “The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing,”Development and Comp Systems 0409041, Economics Working Paper Archive at WUSTL.

[36] Lassibille, G. and Tan, J. (2005) “The Returns to Education in Rwanda, Journal of African Economies, vol. 14, issue 1, pages 92-116.

[37] Kazianga, H. (2002), “Schooling Returns for Wage Earners in Burkina Faso: Evidence from the 1994 and 1998 Living Standard Measurement Surveys”, duplicated, Economic Growth Center, Yale University.

[38] Mazonde,I.N. (1995) ”culture and education in the development of Africa”.

[39] Malik, M. (1997). “Education and the issue of quality”. Education Today, 6 (1), 25-31.

[40] Maluccio, J.,(1998) “Endogeneity of schooling in the wage function: Evidence from the rural Philiphines”Food Consumption Nutriton Divison (FCND) discussion paper NO. 54.

[41] Mincer, J. (1974), “Schooling experience and earnings” New York NBER.

[42] Miller, P. ; C. Mulvey; N.Martin,(Jun, 1995)"What Do Twins Studies Reveal About the Economic Returns to Education?"A Comparison of Australian and U.S. Findings The American Economic Review, Vol. 85, No. 3. , pp. 586-599.

[43] Mkpa, M. A. (2000) " Overview of educational development:precolonial to present day" excerpts from article available on line at *http : //www.onlinenigeria.com/links/eduadv.asp*.

[44] Mohammed, H. (1984). "Attitudes of Rural People Towards Women's Education in Kano State with Particular Reference to Birnin Kudu LG" Kano, Nigeria: Bayero University Press.

[45] Mwabu, G. and P. Schultz,(1996), "Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa,"American Economic Association Papers and Proceedings, 86(2), , pp. 335- 339.

[46] Nielsen, H.S. and N. Westergard-Nielsen,(2001), "Returns to Schooling in Less Developed Countries: New Evidence from Zambia,"Economic Development and Cultural Change, January , pp. 364-394.

[47] Nwachukwu, A.E. (1985). "An Historical Analysis of the Roots of Universal Public Primary Education in Nigeria" (1900-1980) University of Kansas Ph. D. dissertation.

[48] Ozigi, A., & Ocho, L. (1981). "Education in northern Nigeria". Winchester, MA:

[49] Okojie, C., March (2000), "Gender and Education as determinant of household poverty in Nigeria" (WIDER) Disc. 2002/37.

[50] Osili,U. and Long, B.T. (2003). "Universal Primary education and fertility: A Nigerian Experience" NEUDEC conference paper 2003.

[51] Psacharopoulos, G. (1994). "Returns to Investment in Education: A Global Update." World Development 22(9): 1325-43.

[52] Psacharopoulos, G. and Patrinos, H.A. (2002)"Returns to Investment in Education: A Further Update"World Bank Policy Research Working Paper 2881.

[53] Ram, R. and Ram D. S.,(1988), "Farm Households in Rural Burkina Faso: Some Evidence on Allocative and Direct Return to Schooling, and Male-Female Labor Productivity Differentials," World Development, 16(3), pp. 419-424.

[54] Schultz,P. (2004) “Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education” J Afr Econ.2004; 13: 95-148.

[55] Ram, R. and Ram D. S.,(1988), “Farm Households in Rural Burkina Faso: Some Evidence on Allocative and Direct Return to Schooling, and Male-Female Labor Productivity Differentials,” World Development, 16(3), pp. 419-424.

[56] Siphambe, H.K.,(2000), “Rates of Return to Education in Botswana”, Economics of Education Review, 19, pp. 291-300.

[57] Soyinka, W. (1997) “The Open Sore of a Continent: A Personal Narrative of the Nigerian Crisis”(The W.E.B. Dubois Institute series)

[58] Spence, A.M, (1973) “Job Market signaling”, QJE, Vol 87.

[59] Staiger, D. and Stock, J.(1997), “Instrumental Variable Regression with Weak Instruments”, Econometrica, vol. 65, pp. 557-586.

[60] Suberu, R.T. (2001) “Federalism and Ethnic Conflict in Nigeria” United States Institute of Peace: Washington, DC.

[61] Tekaligne, G.(1997), “Returns to Education in the Manufacturing Sector in Zimbabwe:Some Empirical Evidence”South African Journal of Economics, 65(1), March , pp. 99-113.

[62] Willis, R. (1986), “Wage determinants: A survey and reinterpretation of human capital earnings functions”, Handbook of Labor Economics, volume 1A Elsevier.

[63] World Bank, (1996), Nigeria - “Poverty in the midst of plenty : the challenge of growth with inclusion” : a World Bank poverty assessment Document Type: Sector Report

[64] Yoloye, E.A. (ed) 1989: “Resources for Science, Mathematics and Technical Education in Nigerian Secondary Schools A National Survey. A Report submitted to the Ministry of Science and Technology Lagos.”

[65] Zgovu E. and Ephraim W.C. (2000)“Does the Return to Schooling Depend on the Type of Employment? Evidence From the Rural Labour Market in Malawi.”

Appendix

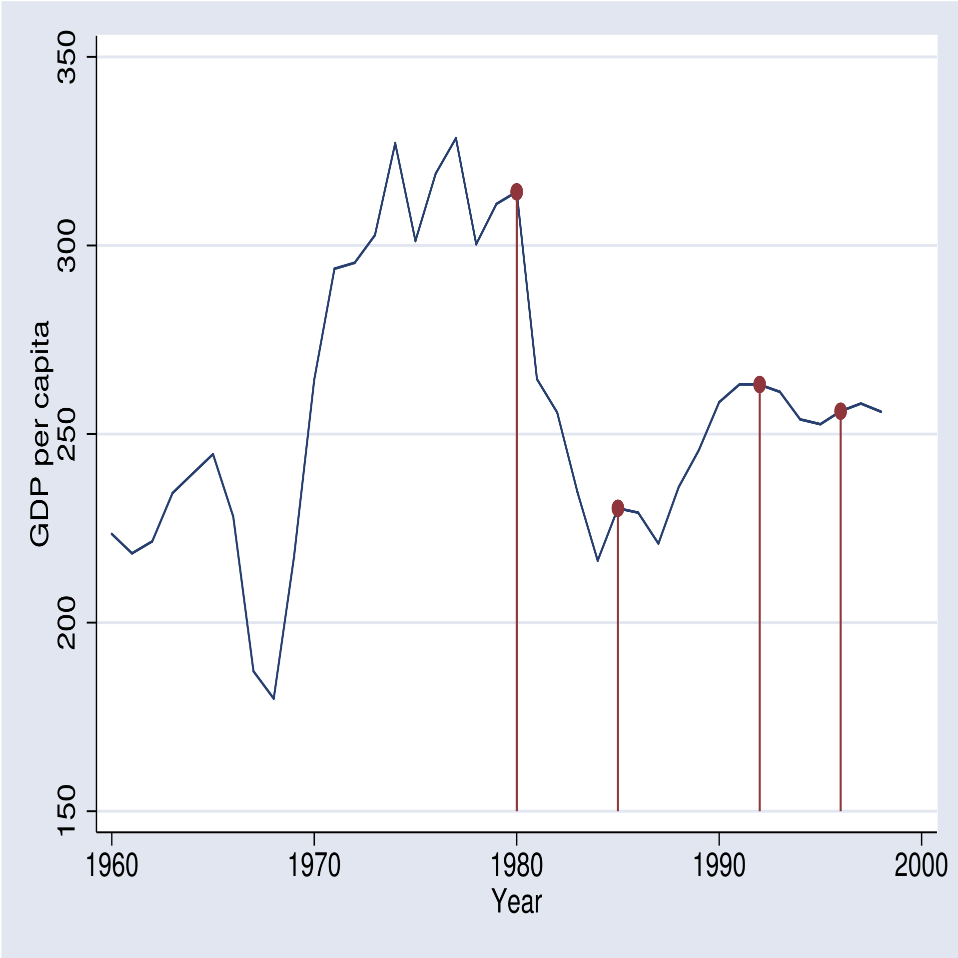

Figure 1: Trends in GDP per capita in Nigeria, 1960-1998
[*vertical lines show 1980, 1985, 1992 & 1996]

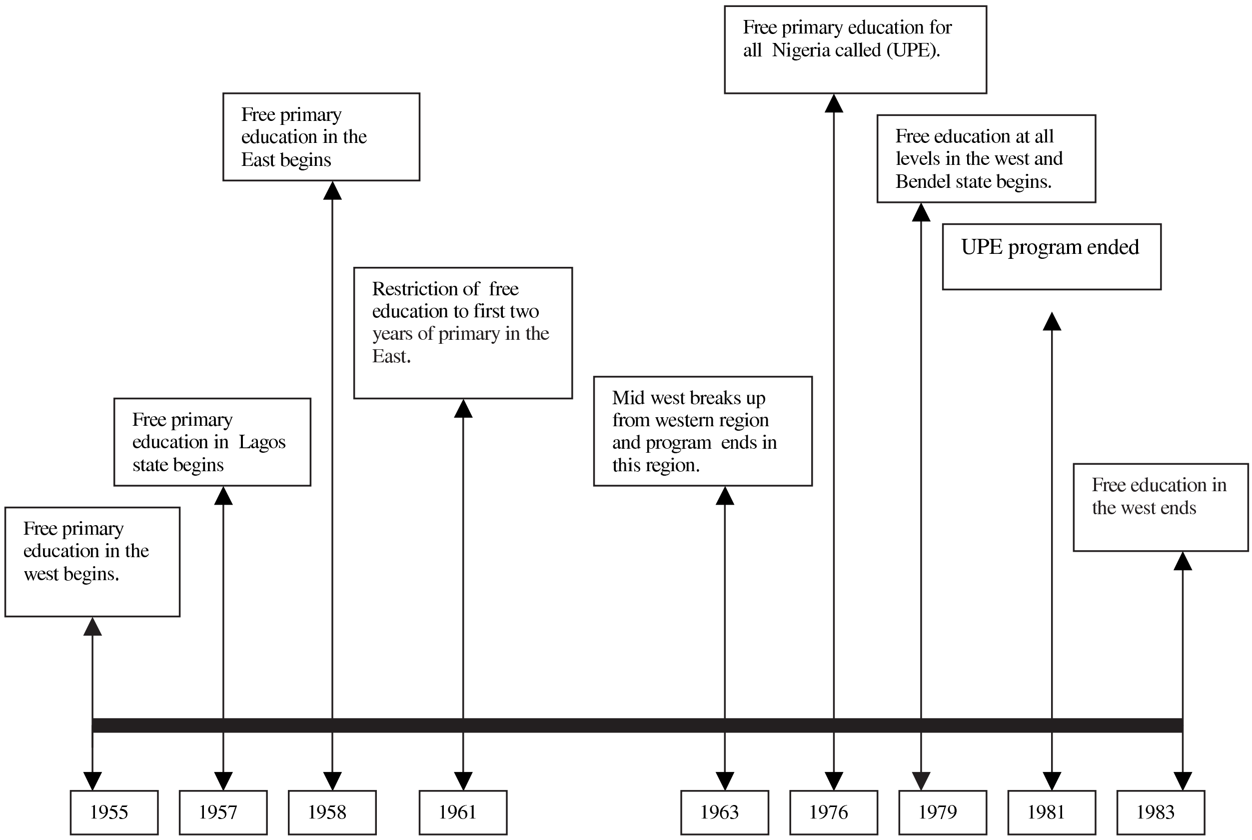

Figure 2: Timeline of free education in Nigeria

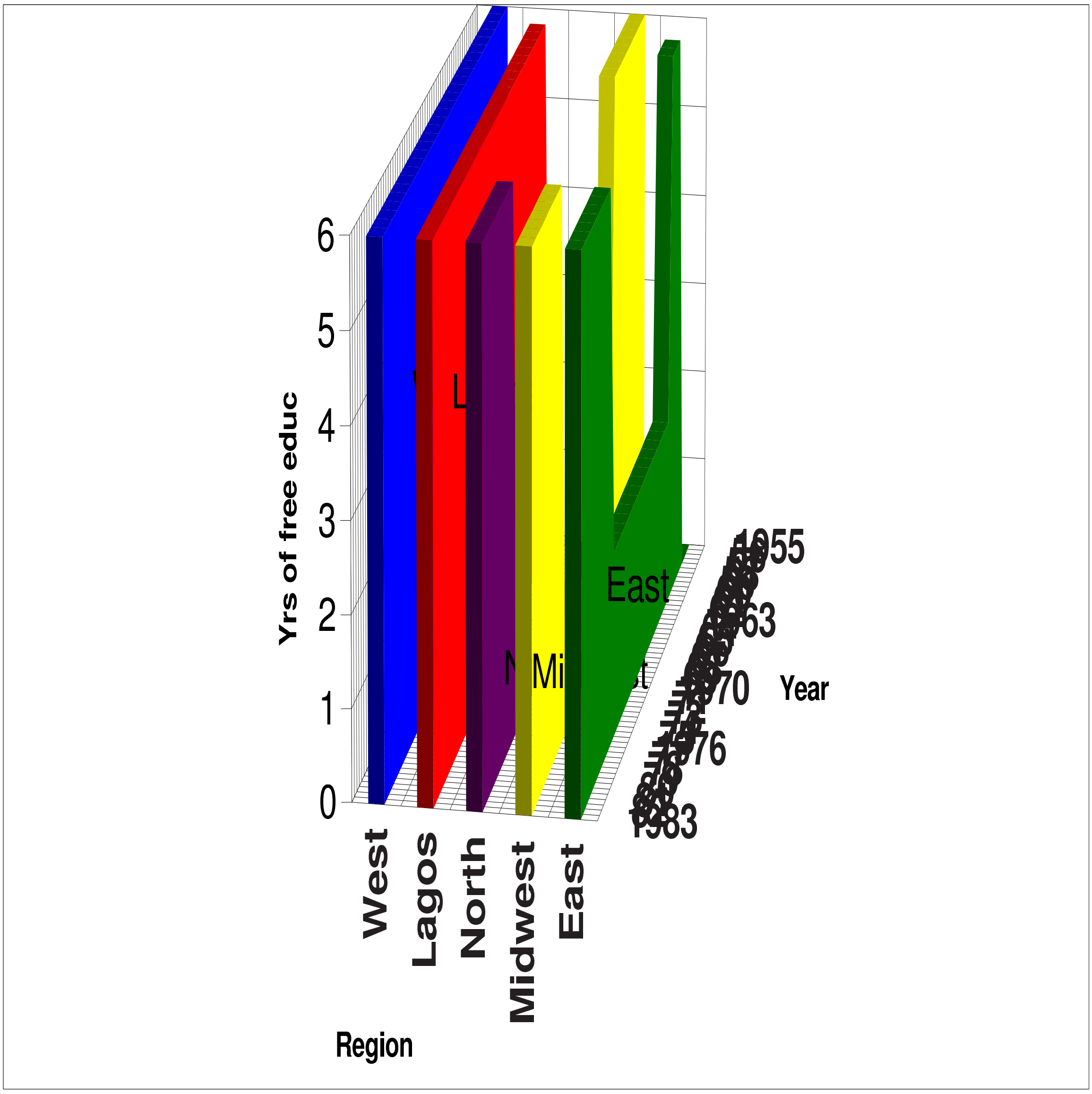

Figure 3: The free education program in Nigeria

Table 1: **Summary Statistics**

Year	1985 (NCS)	1992 (NCS)	1996 (NCS)	1997/98 (GHS)	1998/99 (GHS)
Observations	9,308	9,675	14,383	131,477	106,325
Age	43.22	44.27	44.64	23.486	23.32
	(13.68)	(14.04)	(13.33)	(18.05)	(18.21)
Sex	0.848	0.85	0.861	0.523	0.516
(male=1)	(0.36)	(0.36)	(0.35)	(0.5)	(0.50)
Sector	0.566	0.41	0.211	0.241	0.236
(urban=1)	(0.5)	(0.5)	(0.48)	(0.43)	(0.43)
Years sch	2.89	3.82	3.49	4.17	4.14
	(4.26)	(4.94)	(4.79)	(5.08)	(5.14)
HH size	5.015	5.225	4.469	6.12	6.337
	(4.26)	(3.7)	(2.74)	(3.34)	(3.5)
Income	165.51	176.36	107.86	92.67	93.73
	(201.93)	(282.14)	(214.58)	(298.30)	(158.7)

*Note 1985-1996 data is from the National consumer survey (NCS) and 1997/98 and 1998/99 is from the General household survey (GHS). Standard deviation in bracket.

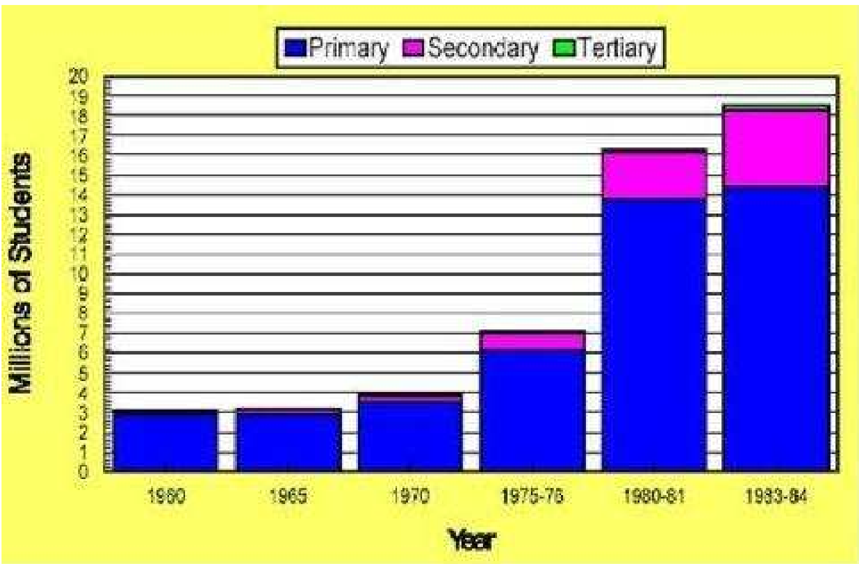

Figure 4: Impact of free primary education on enrollment

Table 2: **Real mean income over time by education level (GHS)**

Education	1997/98		1998/99	
	N	Mean(SE)	N	Mean(SE)
No Education	19890	79.17	15526	76.81
		(2.18)		(1.3)
Some Primary	1843	112.58	1590	91.13
		(18.43)		(3.15)
Full Primary	9787	94.13	7391	97.67
		(1.41)		(1.72)
Full Secondary	5346	111.47	4208	120.15
		(2.0)		(1.98)
Tertiary	1706	156.03	1527	174.56
		(4.34)		(5.90)

Real mean income over time by education level (NCS)

Education	1985		1992		1996	
	N	Mean(SE)	N	Mean(SE)	N	Mean(SE)
Less than Primary	6000	128.37	5571	161.78	8710	81.57
		(2.25)		(3.77)		(2.05)
Complete Primary	2161	193.49	2080	184.85	3033	137.61
		(4.56)		(5.86)		(3.66)
Complete Secondary	865	273.37	1468	203.40	1923	151.08
		(7.37)		(7.88)		(6.24)
Complete Tertiary	282	410.57	556	219.27	717	185.51
		(16.89)		(11.87)		(10.36)

Table 3: **Summary of 2SLS results OLS vs IV 1997-1999**

Schooling	1997/98		1998/99		pooled	
	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)
	Using the length of exposure instrument					
1st Stage results						
UPE exposure	NA	0.135*	NA	0.18*	NA	0.146*
		(0.014)		(0.014)		(0.011)
R^2	NA	0.36	NA	0.36	NA	0.36
2nd Stage results						
Yrs of sch	0.026*	0.037*	0.027*	0.030*	0.026*	0.027*
	(0.001)	(0.014)	(0.001)	(0.01)	(0.001)	(0.011)
Reduced form est.						
UPE exposure	0.005*		0.005*		0.004*	
	(0.002)		(0.002)		(0.001)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes

[** 5% and *1% significance levels. Other variables included in first and second stage results not shown in table (control variables include age and higher powers of age, cohort sex and location). F stats always above 20.]

Table 4: **Summary of OLS vs 2SLS results with controls for 1985,1992 and 1996**

Schooling	1985		1992		1996		All	
	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)
	Using the length of exposure instrument							
1st Stage results								
UPE exposure	Na	0.19*	NA	0.17*	NA	0.14*	NA	0.18*
		(0.05)		(0.03)		(0.02)		(0.016)
R^2	Na	0.12	NA	0.30	NA	0.32	NA	0.27
2nd Stage results								
Yrs of Sch.	0.07*	0.13*	0.027*	0.022	0.028*	0.053**	0.040*	0.013
	(0.003)	(0.05)	(0.04)	(0.06)	(0.002)	(0.03)	(0.001)	(0.03)
Reduced form est.								
IV	0.024*	NA	0.004	NA	0.008**	NA	0.002	NA
	(0.01)		(0.01)		(0.005)		(0.004)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

[** 5% and *1% significance levels. Other variables included in first and second stage results not shown in table (control variables include age and higher powers of age, cohort sex and location). F stats always above 20. NA- not applicable]

Table 5: **Robustness checks: 2SLS estimate of returns to education by gender and sector**

Year	MEN		Women		Rural		Urban	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	First stage: estimate of IV impact on schooling							
UPE	NA	0.17*	NA	0.18*	NA	0.17*	NA	0.10*
		(0.013)		(0.016)		(0.013)		(0.017)
	Second stage: estimate of return to schooling							
RTE	0.024*	0.048*	0.030*	0.024*	0.024*	0.027*	0.030*	0.042*
	(0.001)	(0.014)	(0.001)	(0.01)	(0.001)	(0.012)	(0.001)	(0.02)

[* 5% significance levels]

Table 6: **Robustness checks:Estimate of returns to education for sub-groups(pooled estimate)**

Year	Single households OLS	Single households IV	Work for Profit OLS	Work for Profit IV	Wage Worker OLS	Wage Worker IV
	First stage estimate of UPE					
UPE	NA	0.23* (0.03)	NA	0.29* (0.019)	NA	0.23* (0.03)
	Second stage estimate of RTE					
RTE	0.035* (0.001)	0.040* (0.02)	0.020* (0.001)	0.030* (0.011)	0.024* (0.002)	0.42** (0.02)

[*5%, **10% significance levels]

Table 7: **Robustness checks: Pooled estimate of returns by cohorts**

Year of Birth	1997-1999 OLS	1997-1999 IV
1941-1950	0.033* (0.001)	0.032 (0.02)
1951-1960	0.031* (0.001)	0.022* (0.010)
1961-1970	0.027* (0.001)	0.035* (0.008)
1971-1980	0.022* (0.002)	0.056* (0.011)

[* 5% significance levels
First stage results not included as similar to previous results with UPE being highly significant. Earlier birth cohorts not relevant for instrument.]

Table 8: **Returns estimates with controls after correcting for selectivity**

	OLS	2SLS	Heckman	Heckman2	pscorel	pscore2
IV (Length of exposure)						
1997/98	0.026*	0.037*	0.035*	0.036*	0.035*	0.036*
	(0.001)	(0.014)	(0.017)	(0.02)	(0.014)	(0.014)
1998/99	0.027*	0.030*	0.027*	0.027*	0.030*	0.030*
	(0.001)	(0.01)	(0.011)	(0.011)	(0.01)	(0.01)
pooled	0.026*	0.027*	0.027*	0.029*	0.026*	0.027*
	(0.001)	(0.011)	(0.013)	(0.013)	(0.01)	(0.01)

** 5% and *1% significance levels
pscore1-propensity score estimation with linear probability model and pscore2 -propensity score calculation with probit model.
Heckman- maximum likelihood and Heckman2- two step consistent estimates.
[First stage results not included as similar to previous results with UPE being highly significant. Slight changes were made in terms of controls used for the different regression to avert potential multicollinearity problems.]

Table 9: **Results for IV 1997-1999 without controls**

Schooling	1997/98	1998/99	pooled
	(IV)	(IV)	(IV)
UPE instrument	0.542*	0.554*	0.548*
	(0.012)	(0.012)	(0.008)
Constant	7.83*	7.77*	7.75*
	(0.259)	(0.230)	(0.194)

OLS vs IV results for 1997/98 and 1998/99 without controls

log y	1997/98		1998/99		pooled	
	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)
Age	0.03*	0.032*	0.032*	0.034*	0.031*	0.032*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Age sq.	-0.0002*	-0.0002*	-0.0002*	-0.0003*	-0.0002*	-0.0002*
	(0.000)	(0.000)	(0.000)	(0.00)	(0.00)	(0.00)
Yrs of sch	0.040*	0.051*	0.044*	0.055*	0.042*	0.053*
	(0.001)	(0.001)	(0.001)	(0.003)	(0.000)	(0.002)
Const.	3.17*	3.06*	3.09*	2.97*	3.14*	3.016*
	(0.031)	(0.040)	(0.045)	(0.045)	(0.023)	(0.03)
Year Dummies	No	No	No	No	Yes	Yes

[* 5% significance level IV= length of exposure to UPE]

Table 10: **Summary of first stage IV results without controls**

Sch(y)	1985	1992	1996	pooled
	(1V)	(IV)	(IV)	(IV)
UPE	0.41*	0.464*	0.627*	0.52*
	(0.047)	(0.024)	(0.019)	(0.013)
Constant	6.71*	8.136*	2.42*	5.14*
	(0.751)	(0.497)	(0.429)	(0.191)

[Other variables in first stage reduced form like age excluded in summary. IV is length of exposure to free education]

OLS vs 2SLS results for earnings equation without controls

Variables	1985		1992		1996		pooled	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Age	0.001*	0.009*	0.024*	0.021*	0.031*	0.04*	0.026*	0.031*
	(0.006)	(0.008)	(0.006)	(0.007)	(0.003)	(0.004)	(0.002)	(0.003)
Age sq.	-.0001*	-.0001*	-.0003*	-.0002*	-.0003*	-.0003 *	-.0002*	-.0002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of school	0.093*	0.126*	0.038*	0.026	0.064*	0.13*	0.062*	0.088*
	(0.003)	(0.022)	(0.002)	(0.017)	(0.002)	(0.007)	(0.001)	(0.006)
Year dummies	no	no	no	no	no	no	Yes	Yes
Constant	4.53*	4.22*	3.92*	4.07*	3.17	2.58*	3.92*	3.68*
	(0.159)	(0.274)	(0.147)	(0.182)	(0.082)	(0.090)	(0.047)	(0.050)

[* 5% significance level.]

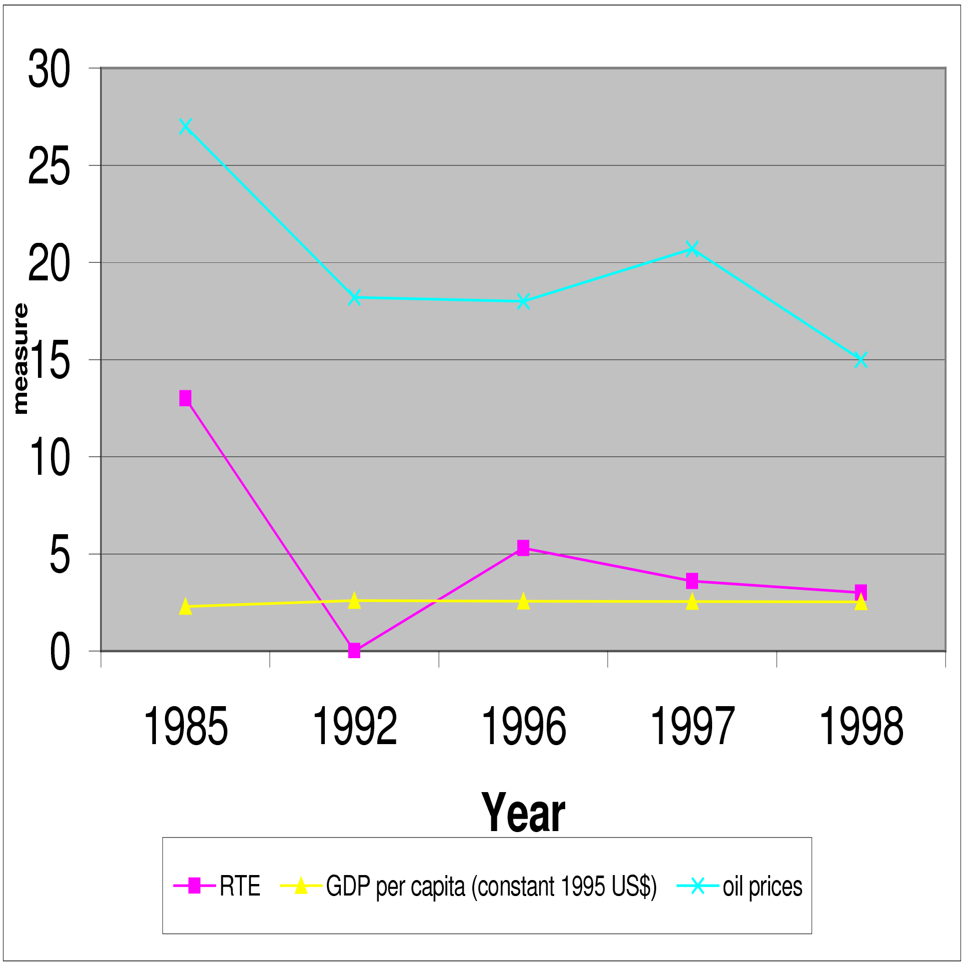

Figure 5: Comparing returns to education(rte), GDP per capita and oil prices over time

Note: Returns to education on the y axis is in % increase for every extra year of schooling and GDPC is GDP per capita. GDPC, RTE and oil prices are in different units.