Africa's Education Enigma? The Nigerian story.*

Ruth Uwaifo Oyelere†

August 30, 2005

Abstract

In the last two decades, the social and economic benefits of formal education in sub-Saharan Africa has been debated. This ongoing debate has created a need to accurately estimate the returns to investments in this sector on the micro level. Anecdotal evidence point to time varying and low returns to education in Africa. Unfortunately, there has been little econometric evidence to support these claims. Here I focus on Nigeria, a country that holds 1/5 of Africa's population, and using instrumental variables based on the timing of the implementation of free primary education in this country, returns to education are precisely estimated at different time periods in the 80s and 90s. In addition, claims of time differences in returns are investigated. The results show that there have been significant changes in returns to education over short time periods. More importantly, the average returns to education are particularly low in contrast to conventional wisdom (precisely 3.6% for every extra year of schooling). This new evidence provides an explanation for both the changes in demand for education in Nigeria and the flight of skilled human capital that characterized the 90s in many African countries.

(**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 at University of California Berkeley and 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 technique used in these estimations were prone to bias because of measurement error and unobservables correlated to 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 200 ethnic groups and 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, though the massive increase in human capital investment via education. Also contributing to this controversy was the fall in living standards and real income of many well-educated Nigerians between 1983 and 1998, in contrast 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 would be addressed in this paper.

The research questions I consider are, first, 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. Second, I address the question: what were the returns to education late in the 90s in Nigeria? The answer should not only provide estimates of the overall returns to education in an African country were the economic value of education is the subject of debate, but can also help us understand the evolution of demand for education in Nigeria. Furthermore, the recent debate led by Lant Pritchett[53] on "Where has all the education gone in Africa?" is important and this study on Nigeria a sizeable portion of Africa could provide alternative explanations to the answers presently available. [2]

Following the endogenous approach to schooling, the returns to education was estimated using two stage least squares (2SLS). The instruments used in this analysis are based on a free primary education program designed to increase educa-

[2]Samer Al-Samarrai and Paul Bennell[10] are some authors who have also written about this issue.

tional attainment, exploiting differences in the periods of implementation across states/regions overtime in Nigeria.

The instrument can be constructed in two ways. First, just as a simple dummy variable indicating whether an individual was exposed to the free education program or not. On the other hand, one can consider the length of exposure. The argument here is the longer an individual is exposed to free education the higher the school attainment. I would be using both instruments separately as combining them does not satisfy overidentifcation restrictions at all reasonable levels. However, I argue in the paper that the length of exposure to the UPE program makes a better instrument and all conclusions and stated estimates are based on this instrument.

To highlight the importance of including appropriate controls in the estimation, the 2SLS estimation of returns to education was carried out with or without controls. Furthermore, as a bench mark to compare these estimates, the OLS technique is also used to estimate the returns to education. Using these techniques, I reject the the null hypothesis of no time differences in returns to education. In fact the average returns to education were insignificant in 1980, rose to over 3% for every extra year of schooling in the mid 80s, fell back to insignificant levels by 1992, and rose to 1.3% by 1996. More striking are the low overall returns estimates. Precisely, I find a 3.6 percent increase in income for every extra year of schooling late in the 90s in Nigeria. This estimate is 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 OLS estimates in other sub-saharan countries [3]. In Nigeria in particular, these estimates suggest upward

[3]See Psacharopoulos and Patrinos (2002)[54] in particular for a review of worldwide estimates. They report average returns to education in Africa of 11.7%. See Schultz (2003)[55] for review on

bias in Aromolaran's (2002)[5] estimates of returns to education, which did not correct for potential sources of bias. Several robustness checks were carried out including correcting for potential sources of selectivity and the above results still hold.

The present study therefore makes an important contribution to the literature on schooling by providing the first estimates of returns to education, using a good instrument, in a West African country. Furthermore, the results draw attention to two major problems with education outcomes in Africa: low returns to education and time differences in returns to education. Time differences has not been considered prior to now, but are important since fluctuating returns make investment in education risky. Finally, several explanations have been sought for the brain drain from Africa in the 90s and the changing demand for education. The low returns to education in Nigeria provide a very reasonable 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 preliminary descriptive 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.

selected African countries. It should be mentioned that my estimates are not directly comparable to studies that estimate returns at each level of education.

2 Literature Review

According to economic theory, earnings are affected by 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)[59] and Arrow (1973)[6] in the 1970s points to education as a signal or a screening process of innate ability. This view is linked to the "sheep-skin effect hypothesis". On the other hand Bhagwati and Srinivasan(1977)[8] view education as a tool for job competition in a distorted labor market. The most common approach to looking at education came from Becker's (1964)[9] 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 if they had 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)[2], Ashenfelter et al (1998)[50], 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 this 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)[54] and Card (1999)[14]). One of the earliest papers using the IV technique in a developing country was Duflo (2001)[18] on Indonesia. Since this paper, other attempts have been made in developing countries but there has been little progress in this regard in considering countries in Africa (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 account for or 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)[5] for Nigeria and Siphambe (2000)[57] for Botswana. Other authors still maintained the OLS framework but went 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 like Glewwe

[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 government sector workers in Ghana, Westergard and Nielsen(2001)[46] for Zambia and most recently Lassibille and Tan (2005) [35]for Rwanda.

(1996)[27] made 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 on African countries, involves estimating returns based on survey of firm based employees rather than households. (See for example Jones(2001)[34] for Ghana, Tekaligne,(1997)[62] for Zimbabwe, and Kahyarara et al (2004) for Kenya and Tanzania [36].) As noted in Psacharopoulos and Patrinos (2002)[54], 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. Besides, firm-based employees are 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) for Kenya and Tanzania [36] and Dabalen (1998) for Kenya and South Africa [16]. Both papers made use of some similar instruments like distance to school and parents education. However, results could still be biased because of common issues of exogeneity of some of the instruments used and problems with the dataset for Dablen (1998).

As with Dabalen (1998)[16], many papers using the instrumental variable (IV) approach has also been critiqued by several authors recently. Staiger and Stock(1997)[60] argued that many studies using IV have weak instruments which led to even more imprecise estimates of returns to education. Carneiro and Heckman(2002)[12] argued along similar lines, stating that most of these instruments are correlated with unobservables like ability and hence lead to inconsistent estimates of returns to education [6].

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

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. In this paper, I will explore a good instrument for schooling to derive more precise estimates of returns to education.

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 for future pay.

The human capital model of household or individual decision-making has its roots in Becker's 1964 [9] 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 will receive if he or she acquires schooling level S_i. Furthermore, I assume individuals

have utility function $U(.,.)$, that are functions of their level of schooling S_i and average earnings, y_i. I also assume individuals want to 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, only considering 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)[63] 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 returns 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 in setting up the estimating equations, 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 term and β is the return to education/schooling.

The Mincer earning function is compatible with equation (7) as the $C\prime s$ could simply contain variables like experience, $(experience)^2$ and other exogenous factors affecting earnings, standard to the Mincer functional form. I intend on using different 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 data on are 1980, 1985, 1992 and 1996. These surveys cover 10280 households in 1980, 9317 households in 1985, 9697 household 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 19 states in Nigeria in 1980 and 1985, 31 states in 1992 and 37 states in 1996. The NISH sampling design is a two-stage replicate sample method, which is a common random sampling procedure. Data from these four surveys are comparable as the same sampling procedure was used in the four 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. Moreover, 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. 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 and not in years of household head's education [10].

[10] The potential problem of overstating amount of schooling, when level of education is reported was checked using data from the General Household Survey (GHS) of 1996/97, indicating no such problem.

To ensure that the data are comparable overtime 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[11]. Finally, to improve survey estimates, a 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 coverage. The federal office of statistics in Nigeria conducts this survey yearly and data are collected from randomly selected households during the four quarters of the year[12]. A drawback of the survey is that different households are surveyed in each survey year. The survey periods I am going to be using are 1997/1998 and 1998/1999. I have data on 32024 households in 1997/98 and 24889 households in 1998/99[13]. 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 3 consecutive years, has an edge over the NCS dataset because it offers information not on household heads alone, but also on all other members of

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

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

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

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 second question.

5 Preliminary Data Analysis

Before highlighting the empirical strategy used to answer each of the questions earlier stated, it is useful to review through some descriptive and summary 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 would differ substantially for some variables like sex, age and so on. Also, one cannot help noticing the massive drop in income overtime. However, the sharp downward trend in household income from 1980 to 1985 is consistent with the downward trends in GDP per capita over the same time period in Nigeria (see Figure 1).

Furthermore, this drastic fall in income, from by about two thirds of the 1980 level by 1996, is consistent with the finding by the World Bank(1996)[64] and Okojie (2002)[49] of an over 300% increase in poverty incidence (from 12% in 1980 to over 50% in 1996).

Table 2 in the appendix 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, even though mean income has been

Table 1: **Summary Statistics**

Year	1980 (NCS)	1985 (NCS)	1992 (NCS)	1996 (NCS)	1997/98 (GHS)	1998/99 (GHS)
Observations	10,265	9,308	9,675	14,383	131,477	106,325
Age	40.84	43.22	44.27	44.64	23.486	23.32
	(12.78)	(13.68)	(14.04)	(13.33)	(18.05)	(18.21)
Sex	0.897	0.848	0.85	0.861	0.523	0.516
(male=1)	(0.30)	(0.36)	(0.36)	(0.35)	(0.5)	(0.50)
Sector	0.457	0.566	0.41	0.211	0.241	0.236
(urban=1)	(0.5)	(0.5)	(0.5)	(0.48)	(0.43)	(0.43)
Years sch	2.402	2.89	3.82	3.49	4.17	4.14
	(3.96)	(4.26)	(4.94)	(4.79)	(5.08)	(5.14)
HH size	3.802	5.015	5.225	4.469	6.12	6.337
	(3.29)	(4.26)	(3.7)	(2.74)	(3.34)	(3.5)
Income	478.31	165.51	176.36	107.86	92.67	93.73
	(481.2)	(201.93)	(282.14)	(214.58)	(298.30)	(158.7)

*Note 1980-1996 data is from the National consumer survey and 1997/98 and 1998/99 is from the General household survey. Standard deviation in bracket

falling at all levels of education over time, if you compare 1980 to 1996, the higher the education level the less drastic the fall in mean income. Hence, education may have served as a form of insurance during the economic recession. Despite this, there is still no clear trend downward or upwards in terms of differences in mean income between the most educated and least educated comparing all the years we have data on in the 90s.

Another point worth nothing is the strange inverted trend in mean income in 1980 in which the most educated have the least mean income. 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, determinants of wage would have less to do with education than with governmental connections and

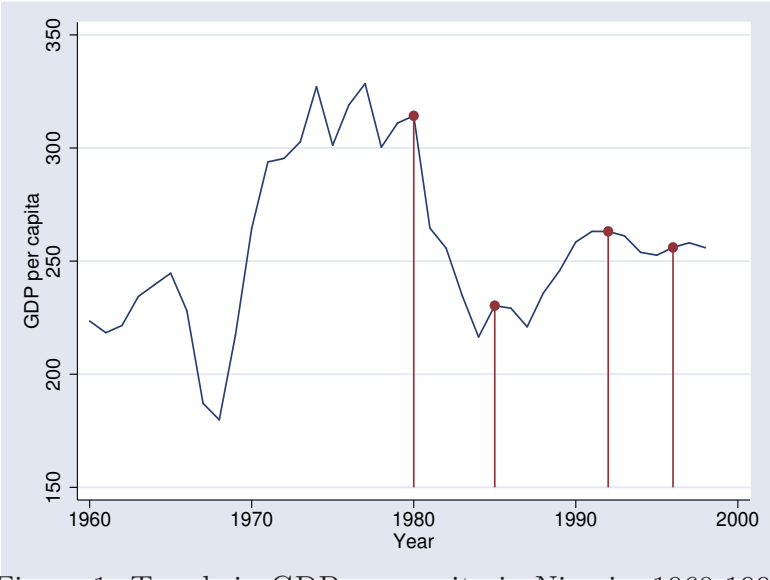

Figure 1: Trends in GDP per capita in Nigeria, 1960-1998

*vertical lines show the four NCS survey years.

social networks. However, one can note that immediately after the oil boom ended in 1981 and the economy collapsed, this trend was reversed and mean income increased with education level. These preliminary results are interesting and call for further investigation using econometric techniques.

6 Estimation techniques

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

Question 1

To test the hypothesis that there are no time differences in returns to education in Nigeria, both the NCS data and GHS datasets were used. Furthermore, some simplifying assumptions on the endogenous schooling model were imposed. These

are:

1. Log earnings are linear in schooling.

2. There are individual variations 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 earning that would still make OLS biased even in this simplified case).

As a bench mark, returns to education was first estimated using OLS on a simple Mincer-type earnings function like 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 individual i.

Subsequently, the return to schooling are estimated by combining equation(9) and (8) using instrumental variables(IV). 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 would be used in a two stage least squares (2SLS) estimation of returns to schooling. For completeness, year 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 22 do not change the results.

The annual estimates of returns to education are then compared for significant differences. If estimates are significantly different, the null hypothesis is rejected.

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.

Question 2

To answer the question on what are the returns to education in Nigeria, only the 1997/98 and 1998/99 GHS data was used. This was because the data covers the whole labor force and contains more information than the NCS. Also, the data are more recent and wage and schooling information is more precisely stated. The empirical strategy for this question is similar to that used to address question one. However, as precise estimates of returns to education are sought, issues of potential selectivity are addressed post estimation for this question.

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 instruments for schooling are exposure or length of exposure to free primary education.[14] 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 old Western region, which is different from the new regional classification. The

[14]The idea of using exposure to the UPE as an instrument originated from Osili and Long (2003)[52] paper on the impact of education on fertility in Nigeria. Using a difference in difference approach similar to Duflo (2001) [18], Osili and Long(2003) identifies clearly significant impact of the program on primary school attainment over the period of its implementation.

regional implementation of this program was not linked to this region's riches or being most prone toward more education, but determined by a choice of policy by its regional leader. Hence, the policy reflected his own preference and not the preference of the populace of the region as in a democracy. 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. However, 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.[15] 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[16]. Figure 2 is a timeline of the program implementation and Figure 3 is a snapshot of the variation in free education across region over time caused by the program. It is this variation in cohorts exposed to free education over time and regions that I exploit as an instrument for school attainment.

[15]In this paper the instrument would be addressed as UPE

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

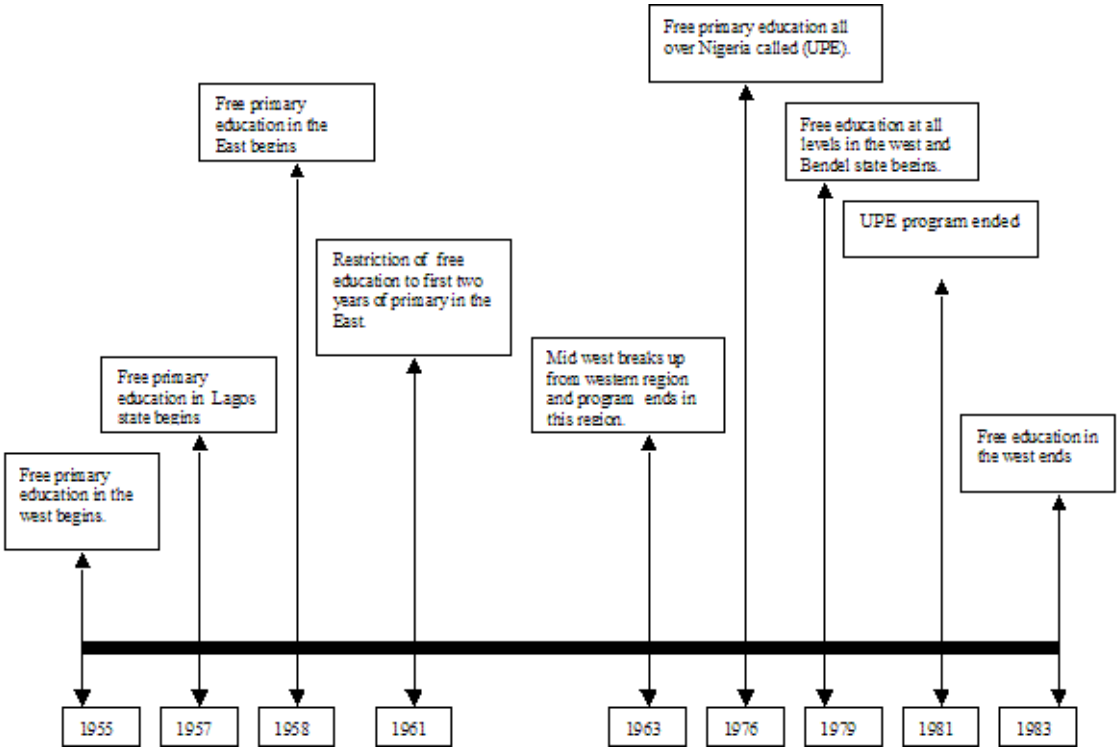

Figure 2: Timeline of free education in Nigeria

6.2 Why the UPE makes a good instrument

Does the program constitutes a good instrument? We know that any good instrument must satisfy three characteristics and the UPE program meets these criteria.

First, a good instrument must be relevant. The relevance/importance of the free primary education program on school attainment and education development in Nigeria has been documented extensively by several authors. For example, Nwanchukwu (1981)[47], Casapo(1983)[15] and Osili and Long (2003)[52] successfully highlight the impact of the UPE program on school attainment. Other descriptive data evidence pointing to the impact of the program is as follows. By

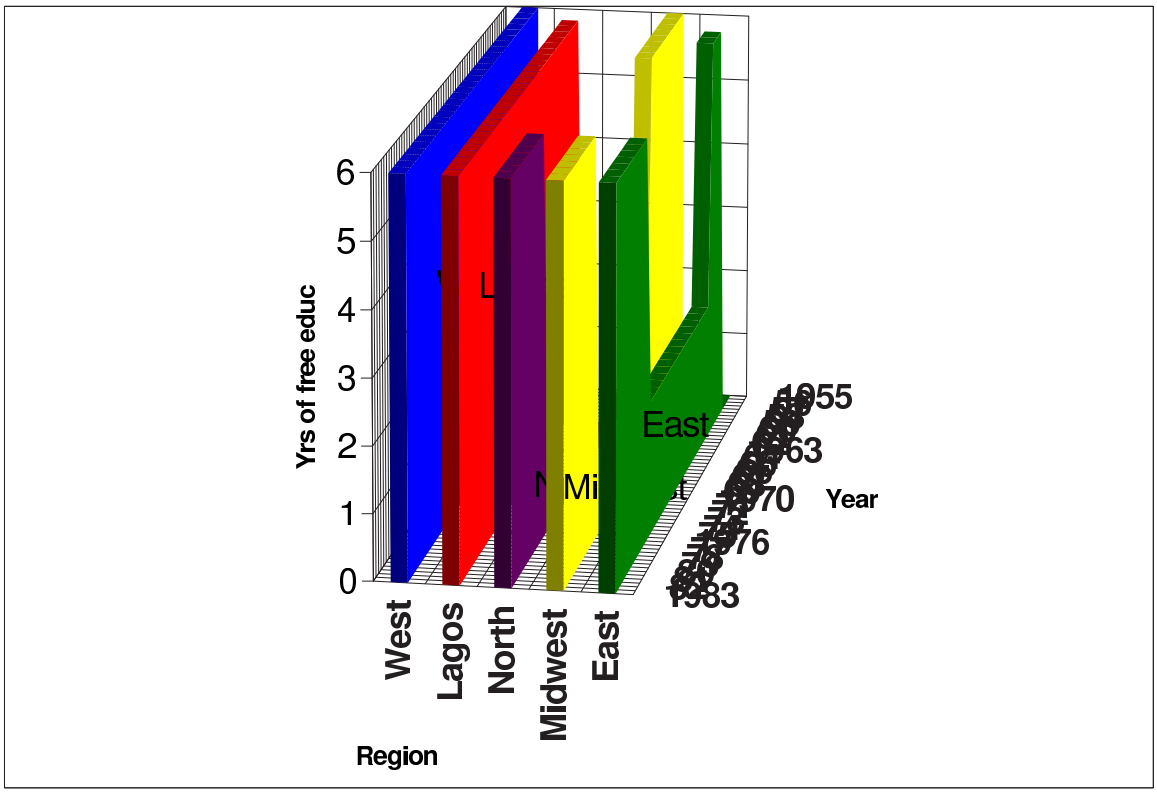

Figure 3: The free education program in Nigeria

1947, the Eastern region 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 (see figure 4). This evidence provides further support for the impact of this program, especially as population growth of school age children was quite consistent over this period. 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.

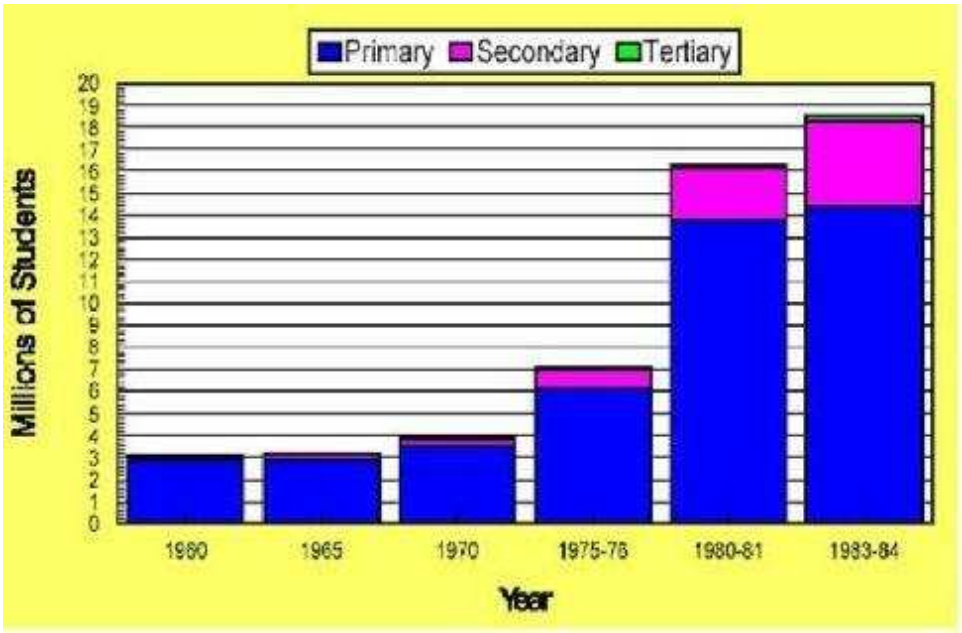

Figure 4: Impact of free primary education on enrollment

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[17].

[17]The scenario where program implementation affected teachers' present income was investi-

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 individuals' preferences. As the program was implemented in a colonial and military setting, program implementation across region and time reflects various commanders' preferences[18]. 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 took over power. Though the program was scrapped at the end of his regime, he has once again implemented the program in 1999, over 20 years later, 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. Moreover, detailed documentation on the history and administration of the program confirm that timing of implementation was random and not influenced by regional/political factors. This means the choice of location for the initial implementation and length was not linked to non-random regional factors. Based on the above arguments and other research[19] into the program implementation, I argue the UPE instrument is exogenous.

Finally, as mentioned earlier, recent studies have critiqued the instrumental

gated, noting no such relationship

[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

[19]The possibility of the temporary fall in quality of education during the initial phase in period of the program affecting an individual's 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 overtime.

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. Staiger and Stock (1997)[60] showed, that when the F-test from a regression of the endogenous variable on the instruments is less than 5, the instruments are weak, but in the case of our instrument the F is much higher than 5 in all cases. Furthermore, the instrument is significant even at the 99% confidence level. Moreover, as exposure to the free primary education has no link with ability nor is it correlated with it, estimates should not be inconsistent.

6.3 Construction of the instrument

As stated in the introduction, the UPE instrument can be constructed in two ways. First, it can be a simple dummy variable indicating whether an individual was exposed to the program or not. The argument is that just being exposed to free schooling was enough to change one's school attainment. On the other hand, one can consider the length of exposure. The argument here is the longer an individual is exposed to free education, the higher the school attainment. I use both instruments separately, as combining them does not satisfy overidentifcation restrictions at all reasonable levels. However, I argue that the length of exposure to the UPE program makes a better instrument because 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 education, as was the case in Nigeria, 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.[20] 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 alternatives 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)[52]. 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. 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.[21]

The instrument is constructed based on year of birth. 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

[20]For most of the analysis I would be showing the results using both instruments. However, I would base all my conclusions and estimates on the results using the length of exposure instrument which proved to be a better instrument.

[21]The estimate of UPE impact (0.65) I tried to replicate was from table four of Osilli and Long 2003[52]). My estimate was 0.54 which is different from 0.65 but one can expect to find slight differences as different datasets are being used. They combine 1990 and 1999 of the DHS while I am using 1997-1999 of the GHS and they also have control variables like religion which are not in the GHS dataset.

from different cohorts in different areas of the country being exposed to free education for different lengths of time. In addition, the instrument was constructed slightly differently when using the different datasets based on the knowledge of the program implementation over 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. In Hass et al(2003)[29] they explicitly stated 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 southeastern or southwestern state in Nigeria. The reason for this bias was culture and the location of most primary schools in selected urban areas.

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 few people truly exposed to free education. 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 would 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 and hence did not

try schooling. In this scenario, the instrument would be weak.

The issue raised above is relevant to this analysis since the 1980, 85 and 92 data years of the NCS naturally contains a higher proportion of observations who were in school during the early phases of the program when true exposure was limited and also have relatively smaller sample sizes compared to the GHS. To get around potential problem when using this dataset, as I do not know exactly which towns in the regions did not have schools, I assume true exposure to the program results in trying school except for when cultural barriers prevent this. This assumption is credible 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 and also culture for girls (see Fafunwa (1974)[20], Ozigi & Ocho, (1981)[48] and Mazonde I(1995)[38]). With my assumption, I cannot isolate those who did not try schooling because of culture from those who were truly not exposed to the program because of accessibility. However, good controls that capture general culture like sex, region and state are included in the regression analysis.

It is well known that most of those prevented from going to school when it was accessible, pre 1980 were conservative muslim girls in the north. As noted by Casapo 1983[15], "The national commitment to UPE was endorsed by many Nigerians. However, rural parents and conservative Muslim parents were less likely to support it for their daughters. These parents were also likely to be concerned about the influences of the public schools' Western style of education on their children because of the culture". Also, Haas et al (2004)[29] made reference to the survey by Mohammed (1984) [44] who found that 58 percent of the one hundred rural Muslim parents she interviewed in northern Nigeria were concerned with

their daughters' marriageability if they go to school. Incidentally, many rural areas and areas with a concentration of conservative muslims in the north did not have schools in many villages and towns until the late 70s/early 80s. This implies these girls were truly not exposed to the UPE program and were rightly treated as such by construction.

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 were really exposed to the UPE as more primary schools were available to these cohorts (see Yoloye (1999)[65] for information on schools expansion in the late 70s). Besides, the dataset is very large and the potential effects on the instrument previously mentioned would be attenuated. However, the results are not significantly different when using the whole sample of the GHS dataset to construct the instruments with or without this assumption.[22].

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. This implies that individuals could possibly be located in places different from where they went to school and the instrument potentially could be irrelevant/invalid 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

[22]Using the GHS data years, the returns to education was estimated using the instrument constructed with or without the assumption, noting no significant difference. This simple analysis implies that making this assumption does not bias estimates

state where they were born.

7 Estimation and Results

7.1 Naive Estimation

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) with no controls was estimated using OLS [23]. However, instead of using imputed experience which is usually computed using a standard formula, I instead use age. The rationale for doing is linked to the implicit flaws in using the standard formula for calculating experience especially in developing countries[24]. Besides, using age is consistent with most of the recent relevant literature[25]. Examples of papers using age instead of experience include Angrist and Kruger (1991)[2], Harmon and Walker(1995)[30], Maluccio (1997)[40], Ashenfelter and Rouse (1999)[7]. Furthermore, age is a good proxy for individuals' experience and is usually accurately measured in the data. Table 3 is a summary of the OLS results using the NCS survey data. These OLS estimates of returns to education would serve as a bench mark for other estimates like the 2SLS.

Table 4 and 5 shows 2SLS results using the NCS data and both potential instruments separately in a naive earnings equation with no controls. Column 1 presents results using length of exposure and column 2 uses the dummy for exposure or non-exposure. Table 6, shows the summary of the first stage results. One

[23]For 1980-1996 the years of schooling is imputed based on education level and schooling system.

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

[25]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

easily notices the significant impact of the UPE instruments, both in the pooled regression and estimates for each year. However, the standard errors in 1980 are higher vis a vis other years and could mean the estimates are less precise. Table 5 shows the 2nd stage results using both instruments. There are a few things one easily notices when comparing these naive 2SLS estimates with its OLS counterpart in Table 3. First, the 2SLS results indicate a positive significant impact of education on earnings in 1980 vis a vis the OLS estimates that are insignificant. Next, the estimate of returns to schooling based on the UPE dummy instruments are higher than those by the length of exposure instrument. However, in comparison to the OLS estimates both UPE instruments give smaller estimates of returns to education than OLS in 1992 (OLS upwardly biased) but higher estimates for 1980, 1996 and the pooled regression (OLS downward biased). For 1985, the estimated return to education using the dummy instrument is higher than the OLS estimate while it is lower using the length of exposure instrument. These divergent results for the dummy instruments in 1985 could be a reflection of its lower precision and is a further argument against it vis a vis the length of exposure instrument.[26] Assuming these results are valid, the finding is a good confirmation of the inability to assume similar biases in OLS over time as in one year the measurement error could be more or less than the omitted variable bias and OLS could be upwardly or downward biased. Considering the first question I am trying to answer, the null is clearly rejected as significant time differences in returns to education exist. Using the naive results above one notices that the returns to education rose from 1980 to the mid 80s, then fell significantly from the mid 80s to the early 90s and

[26]However, one has to be cautious with the 1980 results because a good portion of those affected by the instrument in some regions are still too young to be earning income and are inherently not being considered.

has risen again by the mid 90s to a level similar to the mid 80s.

Turning to the second question, regarding the returns to education in Nigeria in the late 90's, the estimation process above is repeated using the GHS data set. Table 6 and 7 summarizes the IV results of returns to education in Nigeria in the two consecutive periods using both instruments. Based on these preliminary results, it seems the OLS estimates are not really biased as IV estimates and OLS estimates are not significantly different in most cases. The average returns to an extra year of schooling seems to be constant at about 5.0% over these consecutive years.

As only a wage equation with no controls was estimated above, one cannot yet make concrete inferences from these first estimates. The next step is to put in the necessary controls in the wage equation, re-estimate returns to education, and do some other robustness checks to see if the estimates above hold out. As stated earlier, the length of exposure to UPE is a better instrument than an exposure to UPE dummy instrument and the inferences and results presented in this paper would be based on the estimate of this instrument. However, the results based on the dummy instrument are also in most cases reported.

7.2 Estimating the wage equation with controls

There are factors that could potentially affect earnings and some of these factors could be country specific. These variables have to be included in the wage equation to prevent an omitted variable bias problem. Furthermore, to confirm the validity of the estimates, standard controls have to be added to the wage equation. Examples of the controls used in this analysis are dummy variables for region and state an individual lives, dummies for cohort of birth of individuals and higher powers

of the age variable. The potential factors affecting earnings included are gender and location- rural or urban (see figure 5 to further buttress the need to control for sector and region).

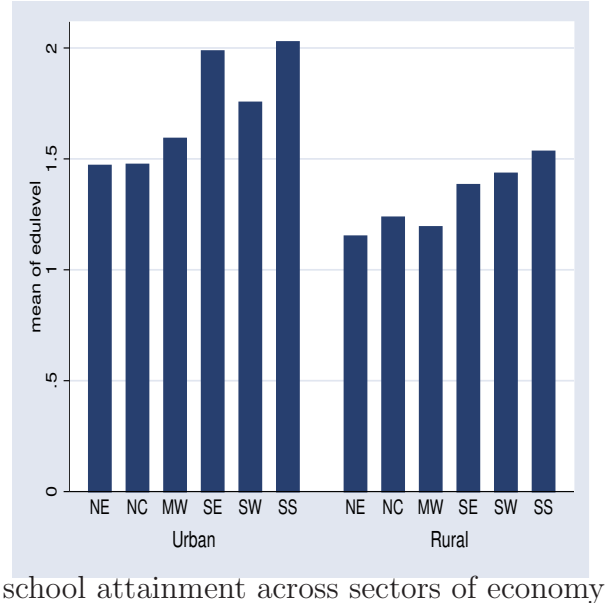

Figure 5: Mean school attainment across sectors of economy and region,1985

As a check to see if these controls and other variables affecting wages significantly change returns estimates, the wage equation was re-estimated using 2SLS and OLS including some of these controls sequentially. Table 8 shows the estimates of the wage equation under difference scenarios including robust standard errors corrected for potential heteroskedacity. The NCS data from 1985 was used for this analysis. Columns 1-4 gives the OLS estimates while columns 5-8 give the 2SLS estimates. This exercise yields some interesting results. First, the inclusion of higher powers on age (age cubed) has no effect on the estimate of returns to schooling for both OLS and 2SLS estimates.[27] Second, we note that the estimates

[27]In most future regressions this higher power is ignored and not added to the estimation equation

are robust to the addition of a gender variable to the earnings equation. Similarly, the addition of the sector variable only reduces the estimates slightly. However, the inclusion of state and regional dummies seems to have significant effects on the 2SLS estimates[28]. In comparison to the naive estimates with no controls, the wage equation including other variables and controls yields IV estimates of returns to education over 3.5% points lower. This finding indicates that initial estimates without controls using 2SLS suffered from omitted variable biases, so that the returns to education were overstated. Similarly the returns to education estimate using OLS in column 4 are higher than the 2SLS estimates using similar variables unlike in all other previous cases where estimates for 2SLS and OLS were quite close, further emphasizing the problems and randomness of the OLS estimates in the presence of endogeneity and omitted variables[29]. A similar exercise was performed for all the other years of the NCS and GHS datasets with similar dramatic drops with the inclusions of the regional or state dummy controls.

Table 9 is a summary of the final estimates of this exercise (this is the case where gender, sector and other control variables are included in the estimation equation). The results in Table 9 are interesting and somewhat unexpected. First, when all variables and controls are included, the OLS estimate is upward biased for each year of the NCS sample. However, the magnitude of difference between the OLS and 2SLS estimates is small. Second, the returns to education are extremely low, and insignificant in the 2SLS in 1980 and 1992. These zero returns to education in 1980 and 1992 would seem implausible to anyone not knowledgable

[28]Even though cohort, state and region dummies are included for column nine more of the change is stemming from the inclusion of state and regional dummies

[29]The potential of multicollinearity in the first and second stage of the regression explaining the fall in the returns to education estimates in the 2SLS estimates has been ruled out using common tests for multicollinearity.

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 1980 and 1992 were periods 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[30]. Recall from Table 2 that in 1980, the mean income for the uneducated was surprisingly higher than all other levels of education. 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 [31] For the other years analyzed, returns are low. They are highest in 1985 at 3.2%, fall to insignificant levels in 1992 and then rise to 1.3% in 1996. These estimates are compared and the null hypothesis is rejected using the t test. This implies that the returns to education, when comparing the mid 80s to the early 80s and 90s, were not constant. However, the magnitude of these differences in returns to education is not too large, at about 3 percentage point. These findings suggest time differences in returns to education over short periods of time.

With respect to the second question on estimating the returns to education recently in Nigeria, a similar analysis as above was carried out. Table 10 is a summary of the results of the estimation process. Once again the estimate are low, at 3.1% in 1997, 3.6% in 1998 and 2.7% for the pooled estimate[32]. If we

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

[31]Some recent books touching on issues like these are Soyinka (1997)[58], Suberu(2001)[61] and Osaghae (1998)[51].

[32]Some of the results using the dummy variables are insignificant for no apparent reason and

compare the 1996 estimates to those more recently, a clear upward trend can be inferred even though the populations being considered in both regressions are slightly different. Also, in contrast to the NCS sample results, the OLS estimates for each GHS sample year is downward biased though OLS and IV estimates are similar.

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 sample, whether household heads or 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 were 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.

8 Robustness checks

One of the first issues 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 affects earnings. Hence, individuals in the rural and urban areas could have different returns to their education. Besides, 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 house-

this further indicates that the length of exposure instrument is preferable.

holds. This problem of getting precise wage estimates for individuals in the rural areas is one reason to focus more attention on the average returns to education in the urban areas. Using both OLS and 2SLS with both IV estimators, returns to education were estimated by gender and sector including all appropriate available controls. Table 11 provides a summary of the returns to education estimates for each year of available data by sector and gender with robust standard errors[33]. This table provides some interesting results. First, the zero return to education in 1980 still holds across gender and sector. However, note that the 2SLS zero average return to education estimate in 1992 is driven by men. Women have a return to education of over 5 percentage points for every extra year of schooling but as they represent only about 10% of the sample in the NCS, on average the returns is still near zero and insignificant[34]. Second, differences in returns to education for men and women have decreased over time, from about 1.3 percentage point difference in 1985 to 0.1 percentage point in 1998/99. The average returns to education by sector follow similar trends as the earlier analysis. The returns in 1980 and 1992 are insignificant and nearly zero and the other periods are characterized by low returns. Furthermore, apart from 1985, where returns to education seem higher in the rural areas, the estimates for the other survey years show returns to be higher in the urban areas. However, the differences in returns across sector are minimal (the greatest difference was 1.6 percentage point in 1985 and 1998/99). The results from Table 11 are to a large extent similar to the average returns for the whole sample estimated earlier (Table 9-10) and more importantly follow similar trends. The null hypothesis of no time differences in returns to education is rejected and

[33]In the subgroups analysis for easy comparison, the instrument is constructed with earlier stated assumption for both the NCS and GHS datasets.

[34]The women in the NCS dataset are women household heads and are generally not a representative of working women in general.

returns to education in Nigeria was still below a 5% increase in income for every extra year of schooling in most cases which is 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 [35] Hence as a robustness check, the returns 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 12 is a summary of the returns to earnings for the single-individual households. Once again the trend is similar with insignificant estimates for 1980 and 1992 with near zero returns in 1980 and above zero returns in 1992. The returns are lower for the single-individual households compared to the whole sample in both surveys but not significantly different. However, the main results still holds. There are time differences in returns to education and the returns to schooling in the late 90s was below 5% for every extra year of schooling.

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 the returns to education differ. These potential cohort effects were controlled for in all the above regressions. However, for completeness, the returns

[35]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.

to education was estimated for the cohorts exposed to the UPE [36]. Table 16 is a summary of the estimates of the returns. It is evident from these estimates that there is not much difference in returns to education between cohorts even though the returns for the 1941-50 cohort is slightly higher in some years. The trend reflected in all the above analysis still holds across cohorts with insignificant returns in 1980 and 1992 and low returns to education over the period in general. These estimates are lower than what some past research have reported to be characteristic of Africa and developing countries in general (see Psacharopoulos and Patrinos 2002[54]).

In line with the second 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 restricting the sample to wage earners[37]. Table 14 is a summary of the returns to education estimate for wage earners/employees and selfemployed/own business in 1997/98 and 1998/99 [38]. The first interesting finding is that there is not much difference in returns to education between the wage workers and the self employed, contrary to the theory that education basically serves as a signal and really does not embody human capital[39]. The results are again consistent with earlier results showing low returns to education within the range of 0.02-0.04 in the late 90s in Nigeria[40].

[36]The cohorts are constructed based on decade of birth.

[37]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

[38]the break down of wage earner or employee and self employed/own business was based on the question frame in the survey and give slightly different responses and so both were used in the analysis

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

[40]As an additional robustness check in estimation, observation were clustered by age or region with no change to significance of estimates

8.1 Correcting for Selectivity

For the second 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[41]. 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 15.

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[1]).

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

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[1]). 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 [42]. This correction mechanism allows the population to be stratified according to their propensity 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 15 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. Only the length of exposure instrument is used for this analysis. 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

[42]To see why these assumptions are sufficient to control selection bias when conditioning on propensity score see Angrist (1997), pp 106 [1]

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 and the length of exposure instrument. Therefore the preferred estimate of average returns to education in Nigeria was 3.6% for every extra year of schooling both in 1997/98 and 1998/99[43]. The estimates of average returns to education in Nigeria are much lower than other estimates in other African countries. Similarly these estimates are lower than Aromolaran's (2002)[5] estimates using OLS for male and female wage workers over similar periods (about 2 percentage points lower).

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 6). More importantly, the results showed that average returns to education were extremely low. Why do we care about these results?

First, the issue of marked differences in returns to education over short periods of time can lead to lower investment in education as individuals may perceive investment in education as risky. This is crucial if human capital is the engine of growth and less investment in education constrains attainable growth rate. 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

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

over the 90s (see Malik(1997)[39] and FOS 2000[22])[44]. It is important for returns to schooling to be relatively high and follow quite a stable trend to encourage future investment in education that would generate the human capital needed to stimulate growth.

Second, low returns to education lead to either individuals finding alternative investments or individuals who already have invested in education seeking higher returns to their education in other labor markets or basically switching to rent-seeking activities. The phenomenon where educated/skilled labor immigrates to other countries especially the developed world to find better returns to skills, is known as "brain drain", was highly pervasive in Nigeria over the 90s[45]. 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, and has been losing an average of 20,000 annually ever since. It has also been estimated that the total cost of this flight of human capital is as much as 4 billion US dollars per annum (see Aredo 2002[4]). This is undoubtedly a constraint on development. One of the worst examples of the brain drain cited by the IOM is Zambia. A few years ago the country had 1,600 doctors, but there are now only 400 in practice. Zambian doctors have migrated to Europe, the US and other places, lured by higher salaries. It is estimated that about 20,000 Nigerian

[44] 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 low declining government allocation to education as government did not see education a priority. This downward trend has slowly been reversed with the change to civilian rule since mid 1999

[45] The brain drain/rent seeking phenomenon and its strong link to low returns to schooling is discussed in a separate paper

academics are now employed in the USA alone which is a sizeable percentage of Nigeria's better trained academicians. There is also evidence that a large chunk of these skilled labor left Nigeria between 1989 and 1995. Several authors have written on the negative effect of brain drain on growth and development of a country, and continued low returns to education in Nigeria, are a sure stimulus for more and more brain drain. Also, brain drain could lead to fall in human capital stock in Nigeria, which could have significant impacts on growth.

Lastly, as these results indicate, returns to education within the range of 2-4% for Nigeria, and most previous papers have estimated returns to education for other African countries in the range of 5-15% (Psacharopoulos and Patrinos (2002)[54]) using OLS and other similar estimation techniques with few controls, there is a possibility that returns to education is 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.

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. The average returns to education were insignificant in 1980, rose to over 3% for every extra year of schooling in the mid 80s, fell back to insignificant levels by 1992 and rose to 1.3% by 1996. Second, precise estimates for average returns to an extra year of schooling recently in Nigeria is about 3.6%. Meaning that for every extra year of schooling, there is a 3.6% increase in wages. These estimates 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. This finding highlights the need to find instruments and re-estimate returns to education in other African countries.

The results also suggest that returns to education does not differ substantially across sector and gender. Furthermore, the importance of including controls like region and state when estimating returns is highlighted in the results.[46]. I also find quite similar returns to education across wage workers and self-employed workers, in contrast to Aromolaran's (2000)[5]. However as the latter rightly noted, it is difficult to disentangle 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 overtime.

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 provide an explanation for changes in demands for education in Nigeria in the 90s. Second, they also provide an explanation for two wide spread phenomena in Nigeria over the 90s: massive brain drain and the move of the educated into rent seeking activities. These phenomena can be attributed to low returns and provide an alternative explanation or response to the question of where has all the educated gone.

The seminal work of Pritchett 1999[53] on this debate offered three explanations to this question. Here I offer an alternative answer.[47] Hence, in response to the

[46]The question of regional differences in returns is being looked into in another paper.

[47]These two explanations are first, the governmental environment is perverse enough that the accumulation of educational capital leads to lower growth. Second, marginal returns to education fell rapidly as the supply of educated labor expanded and demand remained stagnant. The third explanation based on the premise that education quality has fallen drastically so that "years of schooling" have created no human capital cannot be entirely true in Nigeria based on my cohort results above and also because there is substantial evidence showing quality of education

question of where has all the education gone, inferences from my results may suggest a better explanation for Nigeria and possibly 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 which can be limiting. 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[3] 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 has to focus strongly on understanding why returns to education is so low and fluctuating before it can find ways of dealing with this issue. One way of doing this, is to sponsor academic research aimed at understanding these findings.[48] Understanding why returns are low and also, reestimating returns to education in other African countries using the IV strategy and accounting for heterogeneity are interesting areas for research.

fell temporarily during the late 70s when the UPE was phased in Nationwide and was not an issue again until 90s. Besides, the brain drain from Nigeria is a clear indicator that the quality of education was not the main issue affecting returns. If quality of education were very low in Nigeria it would be impossible for skilled capital in Nigeria to successfully transfer their human capital and skills to other countries as they have.

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

References

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

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

[3] Guido Imbens Joshua. Angrist (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

[4] 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

[5] Aromolaran Adebayo. B, 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

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

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

[8] Bhagwati and Srinivasan T.N 1977."Education in a job ladder model and the fairness in hiring role". Journal of public economics, Vol 7. analysis of Panel Data" chapter 10 (John Wiley and sons ltd.).

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

[10] Samer Al-Samarrai and Paul Bennell (August 2003) Where has all the education gone in Africa? IDS (institute of development studies) Africa Project Report

[11] Boissiere M.Knight JB and Sabot R.H.,(1985), "Earning, Schooling, ability and cognitive skills". American Economic Review, Vol 75.

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

[15] Casapo, Margaret. 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 Esther., 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] William Easterly (2001) "The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics" , MIT Press, July 2001.

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

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

[27] Glewwe, Paul William. (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] Sunal, Cynthia Szymanski, Sunal, Dennis W, Rufai, Ruqayyatu, Inuwa, Ahmed, Haas, Mary E 2003 "Perceptions of Unequal Access to Primary and Secondary Education: Findings from Nigeria" African Studies Review, Apr

[30] Harmon, Colm and Ian Walker, (1995), "Estimates of the economic retuns to education for the united kingdom" AER 85;1279-1286.

[31] 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

[32] Hogan Vincent, Roberto Rigobonn(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] J B Knight, R H Sabot and D C Hovey Is the Rate of Return on Primary Schooling Really 26 Per Cent? Journal of African Economies, 1992, 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] Gérard Lassibille and Jee-Peng Tan "The Returns to Education in Rwanda,, Journal of African Economies, 2005, vol. 14, issue 1, pages 92-116

[36] Måns Söderbom & Francis Teal & Anthony Wambugu & Godius Kahyarara, (2004). "The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing,"Development and Comp Systems 0409041, Economics Working Paper Archive at WUSTL

[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] Isaac N Mazonde (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 John,(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] Paul Miller; Charles Mulvey; Nick 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] M. A. Mkpa (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, Hamid. (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, Germano. and T. Paul 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, Helen Skyt and Niels 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 Christiana, March (2000), Gender and Education as determinant of household poverty in Nigeria (WIDER) Disc. 2002/37

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

[51] Eghosa E. Osaghae (1998) Crippled Giant: Nigeria Since Independence 1

[52] Una Osili and Bridget Terry Long (2003). Universal Primary education and fertility: A Nigerian Experience NEUDEC conference paper 2003.

[53] Lant Pritchett (December 1999) Where has all the education gone? World Bank Working paper 1581

[54] George Psacharopoulos and Harry Anthony Patrinos(2002)"Returns to Investment in Education: A Further Update"World Bank Policy Research Working Paper 2881, September 2002

[55] Paul Schultz (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.

[56] Ram, Rati and Ram D. Singh,(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.

[57] Siphambe, Happy Kugwa,(2000), "Rates of Return to Education in Botswana", Economics of Education Review, 19, pp. 291-300.

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

[59] Spence, A.M, (1973) "Job Market signaling", QJE, Vol 87.

[60] Staiger, Douglas and Stock, James (1997), "Instrumental Variable Regression with Weak Instruments", Econometrica, vol. 65, pp. 557-586.

[61] Rotimi T. Suberu (2001) Federalism and Ethnic Conflict in Nigeria

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

[63] Willis, Robert (1986), "Wage determinants: A survey and reinterpretation of human capital earnings functions", Handbook of Labor Economics, volume 1A Elsevier.

[64] 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

[65] 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.

Appendix

Table 2: **Real mean household income over time by education level**

Education	1980		1985		1992		1996	
	N	Mean(SE)	N	Mean(SE)	N	Mean(SE)	N	Mean(SE)
Less than Primary	7146	314.33	6000	128.37	5571	161.78	8710	81.57
		(4.7)		(2.25)		(3.77)		(2.05)
Complete Primary	2063	294.96	2161	193.49	2080	184.85	3033	137.61
		(8.44)		(4.56)		(5.86)		(3.66)
Complete Secondary	876	287.04	865	273.37	1468	203.40	1923	151.08
		(11.6)		(7.37)		(7.88)		(6.24)
Complete Tertiary	185	262.12	282	410.57	556	219.27	717	185.51
		(20.52)		(16.89)		(11.87)		(10.36)

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)

Table 3: **OLS results for earnings equation without controls**

Variables	1980	1985	1992	1996	pooled
Age	0.013*	0.020*	0.024*	0.031*	0.02*
	(0.003)	(0.004)	(0.006)	(0.003)	(0.002)
Age sq.	-0.0001*	-0.0002*	-0.0003*	-0.0003*	-0.0002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of school	0.002	0.085*	0.038*	0.064*	0.050*
	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)
Year dummies	no	no	no	no	Yes
Constant	5.63*	4.07*	3.92*	3.17	5.363*
	(0.060)	(0.083)	(0.147)	(0.082)	(0.047)

* 1% significance level. R^2range between (0.1-0.2) for all regressions. Year dummies for pooled data regression.

Table 4: **Summary of first stage IV results with no controls**

Sch(y)	1980		1985		1992		1996		pooled	
	(IV1)	(IV2)	(1V1)	(IV2)	(IV1)	(IV2)	(IV1)	(IV2)	(IV1)	(IV2)
UPE	1.40*	6.62*	1.01*	5.64*	0.888*	6.16*	1.15*	6.92*	1.09*	6.58*
	(0.052)	(0.177)	(0.031)	(0.134)	(0.023)	(0.118)	(0.017)	(0.080)	(0.012)	(0.056)
Const.	-0.611*	-1.32*	5.23*	4.0 *	6.027*	4.40*	1.30*	1.68*	2.17*	1.69*
	(0.347)	(0.339)	(0.387)	(0.375)	(0.451)	(0.421)	(0.362)	(0.329)	(0.191)	(0.179)

Dependent variable is years of schooling and relevant independent variable is the instrument (UPE). Other variables in first stage reduced form like age excluded in summary. IV1 is length of exposure to free education and IV2 is dummy for exposure. * 1% significance level

Table 5: **IV result using UPE instruments without controls**

loginc(y)	1980		1985		1992		1996		pooled	
	(IV1)	(IV2)	(1V1)	(IV2)	(IV1)	(IV2)	(IV1)	(IV2)	(IV1)	(IV2)
UPE										
Age	0.011*	0.012*	0.016*	0.02*	0.021*	0.024*	0.034*	0.035*	0.021*	0.022*
	(0.003)	(0.003)	(0.004)	(0.004)	(0.007)	(0.007)	(0.004)	(0.004)	(0.002)	(0.002)
Age sq.	-0.0001*	-0.0001	-0.0002*	-0.0002*	-0.0002*	-0.0002*	-0.0003*	-0.0003*	-0.0002*	-0.0002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Yrssch	0.025*	0.025*	0.070*	0.088*	0.023*	0.037*	0.081*	0.087*	0.057*	0.067*
	(0.006)	(0.006)	(0.004)	(0.005)	(0.009)	(0.007)	(0.003)	(0.003)	(0.003)	(0.002)
Const.	5.60*	5.60*	4.21*	4.05*	4.10*	3.94*	3.03*	2.97*	5.31*	5.24*
	(0.059)	(0.059)	(0.098)	(0.093)	(0.182)	(0.168)	(0.090)	(0.089)	(0.050)	(0.049)

Year dummies included for pooled data regression. * 1% significance level

Table 6: **First stage results for IV 1997-1999 without controls**

Schooling	1997/98		1998/99		pooled	
	(IV 1)	(IV 2)	(IV 1)	(IV 2)	(IV 1)	(IV 2)
UPE instrument	0.542*	3.35*	0.554*	3.34*	0.548*	3.34*
	(0.012)	(0.062)	(0.012)	(0.071)	(0.008)	(0.047)
Constant	7.83*	7.59*	7.77*	7.64*	7.75*	7.56*
	(0.259)	(0.259)	(0.)	(0.230)	(0.194)	(0.193)

Other variables in first stage reduced form like age excluded in summary. * 1% significance level
IV 1 = instrumental variable length of exposure to UPE and IV 2= instrumental variable UPE dummy.

Table 7: **OLS vs IV results for 1997/98 and 1998/99 without controls**

log y	1997/98			1998/99			both		
	(OLS)	(IV1)	(IV2)	(OLS)	(IV1)	(IV2)	(OLS)	(IV1)	(IV2)
Age	0.03*	0.032*	0.031*	0.032*	0.034*	0.033*	0.031*	0.032*	0.032*
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Age sq.	-0.0002*	-0.0002*	-0.0002*	-0.0002*	-0.0003*	-0.0003*	-0.0002*	-0.0002*	-0.0002*
	(0.000)	(0.000)	(0.00)	(0.000)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Yrs of sch	0.040*	0.051*	0.045*	0.044*	0.055*	0.050*	0.042*	0.053*	0.047*
	(0.001)	(0.001)	(0.004)	(0.001)	(0.003)	(0.003)	(0.000)	(0.002)	(0.002)
Const.	3.17*	3.06*	3.12*	3.09*	2.97*	3.039*	3.14*	3.016*	3.08*
	(0.031)	(0.040)	(0.039)	(0.045)	(0.045)	(0.03)	(0.023)	(0.03)	(0.03)

*1% significance level

Table 8: **Robustness checks:IV estimates adding controls (1985)**

log income	OLS		OLS		IV		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1st Stage results								
lenght exp.					1.01*	1.02*	1.01*	0.84*
					(0.032)	(0.032)	(0.031)	(0.034)
2nd stage results								
Age	0.08	0.081	0.081*	0.058*	0.081*	0.79*	0.081*	0.057*
	(0.002)	(0.002)	(0.013)	(0.022)	(0.013)	(0.013)	(0.013)	(0.022)
Age sq.	-0.002	-0.002	-0.001*	-0.001*	-0.002*	-0.002*	-0.001*	-0.001*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of school	0.085*	0.083*	0.075*	0.067*	0.071*	0.074*	0.071*	0.032*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.007)	(0.005)	(0.006)	(0.007)
Age cube	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sex	No	0.215*	0.21*	0.217*	No	0.222*	0.212*	0.254*
		(0.025)	(0.025)	(0.025)		(0.026)	(0.025)	(0.026)
Sector	No	No	0.386	0.426*	No	No	0.393*	0.488*
			(0.025)	(0.019)			(0.020)	(0.023)
other controls	No	NO	NO	Yes	No	NO	NO	Yes
Constant	3.14*	2.99*	2.74*	2.82*	3.25*	3.06*	2.77*	2.90*
	(0.201)	(0.207)	(0.199)	(0.373)	(0.211)	(0.211)	(0.202)	(0.376)
1st Stage results			**UPE**	**dummy**				
UPE dummy					5.64*	5.68*	5.55*	4.79*
					(0.134)	(0.134)	(0.131)	(0.149)
2nd Stage results								
Age	0.08	0.081	0.081*	0.058*	0.084*	0.082*	0.084*	0.057*
	(0.002)	(0.002)	(0.013)	(0.022)	(0.013)	(0.013)	(0.129)	(0.022)
Age sq.	-0.002	-0.002	-0.001*	-0.001*	-0.002*	-0.002*	-0.001*	-0.001*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.000)
Years of school	0.085*	0.083*	0.075*	0.067*	0.088*	0.089*	0.083*	0.048*
	(0.002)	(0.002)	(0.002)	(0.022)	(0.005)	(0.005)	(0.006)	(0.006)
Age cube	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sex	No	0.215*	0.21*	0.217*	No	0.210*	0.202*	0.237*
		(0.025)	(0.025)	(0.025)		(0.025)	(0.025)	(0.026)
Sector	No	No	0.386	0.426*	No	No	0.371*	0.459*
			(0.025)	(0.019)			(0.020)	(0.021)
other controls	No	NO	NO	Yes	No	NO	NO	Yes
Constant	3.14*	2.99*	2.74*	2.82*	3.11*	2.94*	2.68*	2.86*
	(0.201)	(0.207)	(0.199)	(0.373)	(0.21)	(0.21)	(0.202)	(0.373)

** 5% and *1% significance levels. R^2 for first stage between 0.22-0.25 for column 9

Table 9: **Robustness checks: Summary of OLS vs 2SLS results with controls for 1980,1992 and 1996**

Schooling	1980		1992		1996		All	
	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)	(OLS)	(IV)
	Using the length of exposure instrument							
1st Stage results								
UPE exposure	NA	1.405*	NA	0.889*	NA	0.889*	NA	0.924*
		(0.053)		(0.034)		(0.02)		(0.013)
R^2	NA	0.11	NA	0.35	NA	0.41	NA	0.29
2nd Stage results								
RTE using IV 1	-0.001	0.011	0.027*	0.01	0.028*	0.013*	0.033*	0.025*
	(0.002)	(0.006)	(0.004)	(0.013)	(0.002)	(0.005)	(0.001)	(0.004)
Reduced form est.								
IV (1)	0.015	NA	0.007	NA	0.011*	NA	0.023*	NA
	(0.009)		(0.010)		(0.05)		(0.004)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Using the UPE dummy instrument							
1st Stage results								
UPE dummy	NA	6.14*	NA	5.31*	NA	5.74*	NA	5.79*
		(0.185)		(0.135)		(0.097)		(0.063)
R^2	NA	0.14	NA	0.40	NA	0.46	NA	0.34
2nd Stage results								
RTE using IV 2	-0.001	0.004	0.027*	0.021*	0.028*	0.02*	0.033*	0.034*
	(0.002)	(0.005)	(0.004)	(0.009)	(0.002)	(0.004)	(0.001)	(0.003)
Reduced form est.								
IV (2)	0.023	NA	0.126*	NA	0.116*	NA	0.197*	NA
	(0.031)		(0.05)		(0.024)		(0.017)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

**5% and *1% significance levels. IV1 and IV2 defined as above. Reduced form estimate of the instrument is derived by estimating a wage equation with the instrument instead of years of schooling using OLS. F statistic always high.
NA- not applicable

Table 10: **Robustness checks: Summary of 2SLS results using both instruments 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	0.127*	NA	0.167*	NA	0.146*
		(0.014)		(0.016)		(0.011)
R^2	NA	0.36	NA	0.33	NA	0.36
2nd Stage results						
Yrs of sch (using IV 1)	0.025*	0.031*	0.031*	0.036*	0.030*	0.027*
	(0.001)	(0.013)	(0.001)	(0.014)	(0.001)	(0.009)
Reduced form est.						
UPE exposure	0.004*		0.006*		0.009*	
	(0.2)		(0.002)		(0.002)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Using the Dummy instrument					
1st Stage results						
UPE dummy	NA	0.303*	NA	0.343*	NA	0.41*
		(0.074)		(0.086)		(0.053)
R^2	NA	0.36	NA	0.33	NA	0.37
Reduced form est.						
UPE dummy	0.012		0.001		0.016	
	(0.01)		(0.012)		(0.010)	
2nd Stage results						
Yrs of sch (using IV 2)	0.025*	0.048*	0.031*	0.004	0.030*	0.038*
	(0.001)	(0.025)	(0.001)	(0.034)	(0.001)	(0.017)
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. F stats always above 20. State control dropped in 1998 because of multicollinearity issues.

Table 11: **Robustness checks: 2SLS estimate of returns to education by gender and sector with controls**

	MEN			Women			Rural			Urban		
Year	OLS	IV1	IV2	OLS	IV1	IV2	OLS	IV1	IV2	OLS	IV1	IV2
1980	-0.001	0.010	0.003	0.005	-0.010	-0.000	-0.0037	0.001	-0.006	0.003	0.013	0.009
	(0.002)	(0.007)	(0.005)	(0.006)	(0.016)	(0.014)	(0.003)	(0.01)	(0.007)	(0.002)	(0.009)	(0.007)
1985	0.069*	0.037*	0.054*	0.073*	0.024*	0.048*	0.066*	0.055*	0.074*	0.071*	0.039*	0.053*
	(0.002)	(0.008)	(0.007)	(0.0067)	(0.018)	(0.012)	(0.013)	(0.01)	(0.009)	(0.003)	(0.01)	(0.009)
1992	0.025*	-0.007	0.014	0.038*	0.055*	0.033*	0.025*	0.010	0.007	0.029*	0.014	0.038*
	(0.004)	(0.015)	(0.011)	(0.010)	(0.021)	(0.017)	(0.006)	(0.014)	(0.011)	(0.005)	(0.022)	(0.016)
1996	0.030*	0.010***	0.024*	0.038*	0.029*	0.029*	0.03*	0.015*	0.028*	0.032*	0.020*	0.018**
	(0.002)	(0.006)	(0.005)	(0.005)	(0.012)	(0.009)	(0.002)	(0.006)	(0.005)	(0.003)	(0.010)	(0.009)
97/98	0.027*	0.024*	0.024*	0.032*	0.026*	0.025*	0.026*	0.022*	0.023*	0.034*	0.030*	0.026*
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.02)	(0.002)	(0.001)	(0.003)	(0.002)
98/99	0.028*	0.024*	0.023*	0.036*	0.025*	0.023*	0.027*	0.017*	0.02*	0.039*	0.033*	0.028*
	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.003)

** 5% and *1% significance levels
First stage results not included as similar to previous results with UPE being highly significant. .

Table 12: **Robustness checks: 2SLS estimate of returns to education by single household with controls**

	Single households		
Year	OLS	IV1	IV2
1980	-0.001	0.000	0.003
	(0.003)	(0.008)	(0.008)
1985	0.062*	0.026*	0.038*
	(0.005)	(0.012)	(0.01)
1992	0.024*	0.011	0.034*
	(0.009)	(0.022)	(0.017)
1996	0.019*	0.005	0.016**
	(0.005)	(0.010)	(0.008)
1997/98	0.035*	0.022*	0.029*
	(0.002)	(0.004)	(0.003)
1998/99	0.036*	0.029*	0.027*
	(0.002)	(0.006)	(0.004)

** 5% and *1% significance levels
First stage results not included as similar to previous results with UPE being highly significant.

Table 13: **Robustness checks: 2SLS estimate of returns to education by cohorts with controls**

YOB	1980		1985		1992		1996		1997/98		98/99	
	OLS	IV1	OLS	IV1	OLS	IV1	OLS	IV1	OLS	IV1	OLS	IV1
1941-50	-0.001	0.008	0.08*	0.046*	0.026*	-0.001	0.037*	0.022*	0.033*	0.027*	0.033*	0.023*
	(0.003)	(0.009)	(0.004)	(0.012)	(0.008)	(0.02)	(0.004)	(0.01)	(0.002)	(0.004)	(0.002)	(0.005)
1951-60	0.004	0.001	0.061*	0.048*	0.029*	0.019	0.036*	0.019*	0.029*	0.023*	0.034*	0.028*
	(0.004)	(0.008)	(0.004)	(0.008)	(0.007)	(0.012)	(0.004)	(0.009)	(0.001)	(0.002)	(0.001)	(0.003)
1961-70	0.013	0.064	0.040*	0.003	0.014***	0.016	0.025*	0.017*	0.026*	0.023*	0.028*	0.023*
	(0.029)	(0.05)	(0.01)	(0.01)	(0.008)	(0.014)	(0.004)	(0.007)	(0.001)	(0.002)	(0.002)	(0.002)
1971-80	NA	NA	NA	NA	0.011	0.067	0.012	0.019***	0.024*	0.023*	0.020*	0.023*
	NA	NA	NA	NA	(0.012)	(0.043)	(0.008)	(0.011)	(0.02)	(0.002)	(0.020)	(0.004)

***10%, ** 5% and *1% significance levels
First stage results not included as similar to previous results with UPE being highly significant earlier cohorts not relevant for instrument. YOB- year of birth NA- Not applicable as cohort to young to be working at that time.

Table 14: **Robustness checks: 2SLS estimate of returns to education by wage worker/own business with controls**

Year	Work for Profit			Wage Worker		
	OLS	IV1	IV2	OLS	IV1	IV2
1997/98	0.025*	0.023*	0.023*	0.032*	0.032*	0.016*
	(0.001)	(0.002)	(0.001)	(0.002)	(0.008)	(0.008)
1998/99	0.026*	0.022*	0.021*	0.034*	0.023*	0.022*
	(0.001)	(0.002)	(0.002)	(0.002)	(0.011)	(0.009)
	Own business			Employee		
1997/98	0.025*	0.023*	0.024*	0.033*	0.039*	0.02*
	(0.001)	(0.001)	(0.001)	(0.002)	(0.009)	(0.007)
1998/99	0.027*	0.023*	0.022*	0.035*	0.029**	0.031*
	(0.001)	(0.002)	(0.002)	(0.002)	(0.014)	(0.011)

** 5% and *1% significance levels
First stage results not included as similar to previous results with UPE being highly significant.

Table 15: **Returns estimates with controls after correcting for selectivity**

	OLS	2SLS	Heckman	Heckman2	pscorel	pscore2
IV (Length of exposure)						
1997/98	0.025*	0.031*	0.030*	0.030*	0.035*	0.036*
	(0.001)	(0.014)	(0.014)	(0.015)	(0.013)	(0.013)
1998/99	0.031*	0.036*	0.032*	0.035*	0.036*	0.036*
	(0.001)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)
pooled	0.030*	0.027*	0.020*	0.021*	0.026*	0.027*
	(0.001)	(0.009)	(0.010)	(0.001)	(0.009)	(0.009)

** 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.

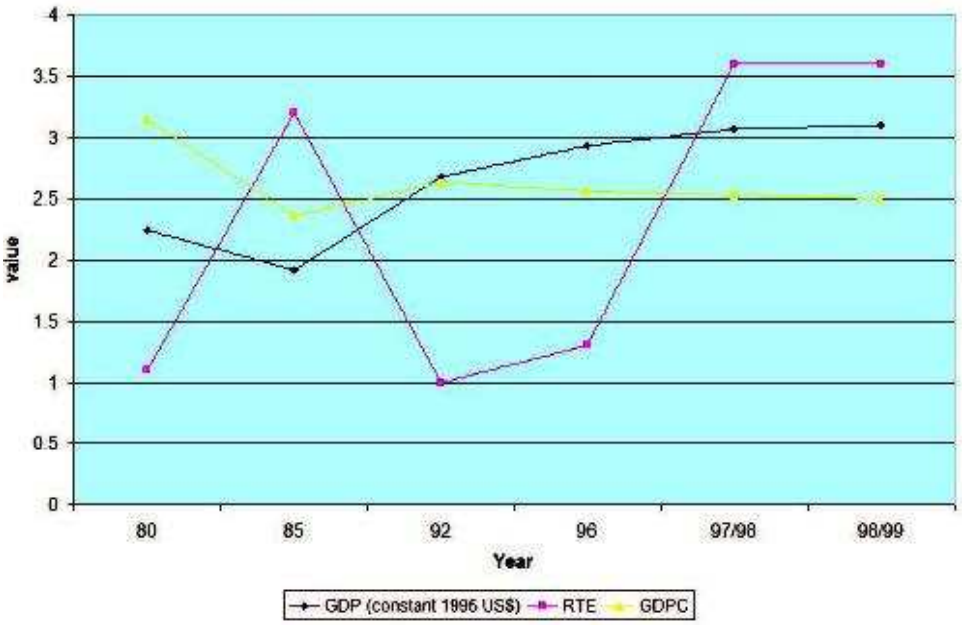

Figure 6: Comparing returns to education(rte), GDP per capita and GDP in Nigeria over time

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