African Development Bank Group

Working paper series

N° 145- February 2012

Assessing the Returns to Education in the Gambia

Jeremy D. Foltz and Ousman Gajigo

Working Paper Series

African Development Bank
Angle des l'avenue du Ghana et des rues
Pierre de Coubertin et Hédi Nouira
BP 323 -1002 TUNIS Belvédère (Tunisia)
Tel: +216 71 333 511
Fax: +216 71 351 933
E-mail: afdb@afdb.org

The Working Paper Series (WPS) is produced by the **Development Research Department** of the African Development Bank. The WPS disseminates the findings of work in progress, preliminary research results, and development experience and lessons, to encourage the exchange of ideas and innovative thinking among researchers, development practitioners, policy makers, and donors. The findings, interpretations, and conclusions expressed in the Bank's WPS are entirely those of the author(s) and do not necessarily represent the view of the African Development Bank, its Board of Directors, or the countries they represent.

Working Papers are available online at
http:/www.afdb.org/

Correct citation: Foltz, Jeremy D.; and Gajigo, Ousman (2012), Assessing the Returns to Education in the Gambia, Working Paper Series N° 145, African Development Bank, Tunis, Tunisia.

AFRICAN DEVELOPMENT BANK GROUP

Assessing the Returns to Education in the Gambia

Jeremy D. Foltz and Ousman Gajigo[1]

Working Paper No. 145
February 2012

Office of the Chief Economist

[1] Jeremy D. Foltz and Ousman Gajigo are respectively Associate Professor, University of Wisconsin, Madison, and Economist, African Development Bank respectively. The authors are grateful to the Central Bureau of Statistics for giving access to the household data sets and the Department of State for Education for access to the school data. They are also indebted to Abu Camara and Lamin Fatty at the Central Bureau of Statistics for the innumerable ways they helped with the data. This work has also benefited from suggestions by Laura Schechter, Thi Minh Ngo, seminar participants at Wisconsin and Yale, participants at the 4th Minnesota International Economic Development Conference, the editor and two anonymous reviewers.

Abstract

Using three nationally representative surveys from the country, we estimate the private rates of returns to education in The Gambia. To obtain consistent estimates, we exploit exogenous variation in school availability in the country at the district level at the time current wage earners where born. Our results show that the private rates of returns to education are quite high, although heterogeneous across regions of the country. The high rates of returns are robust to alternate formulations.

Keywords: Gambia; Schooling; Returns to Education; Wage

JEL Codes: C36**,** I21, I25, J24,

I. Introduction

The importance of education in development is a perennial topic in economics especially in the context of sub-Saharan Africa's development experience. The connection is not surprising since the region stands out both in its low level of schooling and its low historic average rate of economic growth. In the macroeconomic growth literature, Krueger and Lindahl (2001) showed that education is positively associated with economic growth, a result that accords well with many previous studies. Micro-level research on private rates of returns to education has shown disparate estimates in sub-Saharan Africa in the private benefits to education. Our work focuses on private returns to education in The Gambia[2], a small country in West Africa with very low levels of schooling. Like other countries in the region, it also has achieved little economic growth since independence in 1965. It is therefore not surprising that the country is not on schedule to achieve one of the Millennium Development Goals: universal primary education by 2015.

This work adds to the large literature that provides a range of estimates on the private rate of returns to education in Africa. Psacharopoulos' 1994 review of the literature found that the average private rate of returns to education for sub-Saharan Africa were around 13%, though with significant variation in estimates across countries. Many studies on returns to education in sub-Saharan Africa have improved on the estimation methodology of the works cited by Psacharopoulos (1994). For example, Glewwe (1996) was able to directly address school quality and ability, problems that

[2] The Gambia's official name is "The Gambia" including the capitalized article, so we use that each time that way in the text.

have largely been unaddressed in the earlier literature. That work estimated the rate of returns to education in Ghana to be in the range of 3% to 6%. In a more recent and equally rigorous work, Oyelere (2010) estimates returns to schooling that are slightly under 4% in Nigeria. And in another study about a West African country, Kazianga (2004) also estimated returns to education in Burkina Faso, finding that the returns to education are between 9% and 17% at the primary school level and 13% and 20% at the secondary school level in the private sector. In the public sector, the returns range between 0% to 6% at the primary level and 10% and 11% at the secondary level. Schultz (2004) found estimates for Ivory Coast ranging from 3.8% to 28%. In a different region of sub-Saharan Africa, Siphambe (2000) estimates the rate of returns to education in Botswana to be in the range of 12% to 18%.

The literature therefore provides a wide range of estimates of the rate of return to education in sub-Saharan Africa. It could be the case that there are indeed very large differences between countries in the rates of returns to education since there has been very little replication of estimates within a single country. Part of the difference in estimates may also be due to the use of improved econometric techniques among recent papers. Some of these new approaches have addressed issues such as ability bias and selection - problems that were not always addressed in many earlier papers.

Another possibility is that differences in estimation strategies can also produce different results since the estimates may be specific to only a subset of the population in a given country. Specifically, the estimates from using an instrumental variable approach may not be comparable across different studies that employ different instruments since such an estimation strategy produces the local average treatment effects (Card, 2001;

Imbens and Angrist, 1994). Typical estimates using instrumental variables, in which the most common instrument measures access to schooling, provide measures of the returns to schooling for those who would have continued in school but did not have access to schooling. Given that in the African context there is great variation across countries, ethnic groups, and religions in the proclivity of parents to send their children to school even when it is available and affordable, one should also expect great variation in estimates of returns based on that population.

This work contributes to the literature by providing the first estimates of the private rate of returns to education for The Gambia and among its regions. Our estimates rely on the exploitation of the exogenous variations in the availability of schools across the country at the district level and its interaction with year of birth of individuals to control for ability bias. In addition, we use exogenous rainfall shocks to control for selection bias. Like many instrumental variables, ours are not perfect. We discuss the possible violations of the exclusion restriction and provide further robustness checks to mitigate against them.

Our study uses three nationally representative household surveys from 1992, 1998 and 2003 that provide a very high coverage rate for the overall population of The Gambia. The results show high and significant private rate of returns to education for individuals in the wage sector. The results also suggest large significant differences in the rate of returns to education across regions.

This work proceeds as follows. In section II, we describe the data set, which includes the historical education data of The Gambia. Section III discusses our estimation strategy. Our main results on private rate of returns are presented in Section IV, where we also present our robustness checks. Section V concludes the paper.

II. Data description

Like other countries in the Sub-Saharan Africa region, The Gambia is very poor by world standards, predominantly rural, and has an agricultural-based economy. Since independence in 1965, economic growth has been nearly non-existent, averaging about 0.7% per year between 1965 and 2009. The GDP per capita (PPP) in 2009 was $1,285 in constant 2005 US dollars (World Bank 2010).

The data set we use includes three household surveys (1992, 1998 and 2003) carried out by the Central Bureau of Statistics in The Gambia. These surveys, the Household Poverty Survey, cover all the seven regional administrative areas (regions 2-6 are commonly known as Divisions) and most districts[3]. The surveys are repeated cross-section and are carried out approximately every five years. The numbers of households sampled in years 1992, 1998 and 2003 were 1,387; 1,923 and 4,672 respectively, making a pooled sample size of 7,982 households. This household coverage results in 62,538 sampled individuals. Out of this sample, approximately 13,780 individual are wage earners.

Table 1 presents the summary statistics of key variables for individuals in the labor market. The three time periods are very similar in most of the variables listed. As expected, the average years of schooling (*S*), at 2.92 years, is very low in the sample. Surprisingly, the average of this variable is lower in the 1998 and 2003 samples than in

[3] The country is divided in to roughly six administrative areas: five divisions plus the capital and its surrounding area called the Greater Banjul Area. Within the five divisions are districts numbering close to 40. The 1992 and 1998 surveys covered most but not all districts. The 2003 survey covered all districts.

1992, but virtually all of that difference can be accounted for in the differences between the rural samples of the two sets of periods. While wage earners are similar to the general sample in average schooling, they are on average 7 years older relative to the general population. Similarly, women have significantly fewer years of schooling than men, attaining 2.34 years of schooling on average relative to men's average of 3.68 years.

In constant 2003 values, annual wages have increased by 49% between 1992 and 2003. Figures 1 and 2 provide the wage distributions by gender and location. Unlike wages, total household income has not grown over this time period. In fact, it fell by 14% over that time period most likely because of the drought of 2003. This is consistent with macroeconomic figures since the average annual GDP per capita growth rate was -0.05% from 1992 to 2003 (World Bank 2010).

Figure 1: Kernel Densities of Male and Female wages.

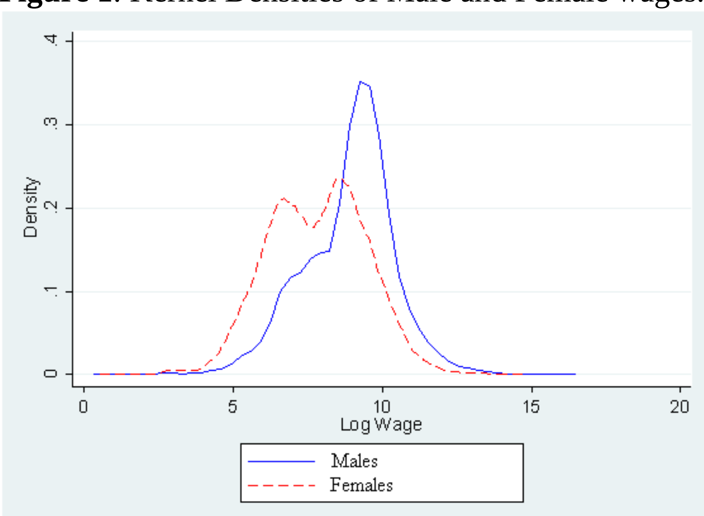


Figure 2: Kernel Densities of Urban and Rural wages.

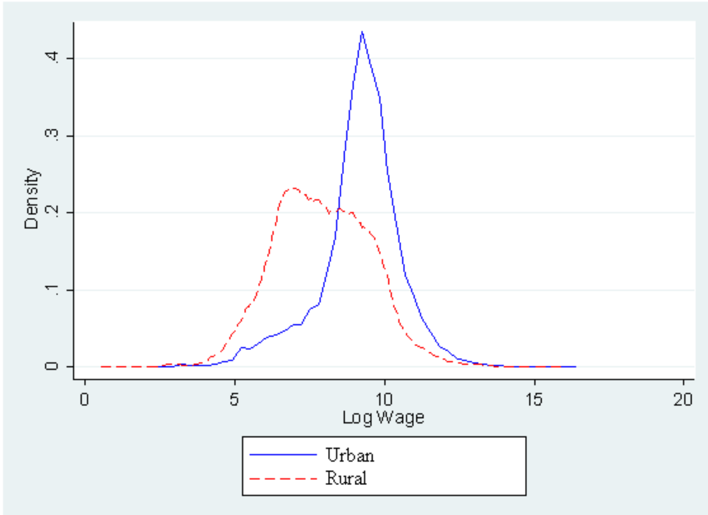


Table 1: Summary statistics of some key variables (standard deviations are in parentheses). This summary is restricted to wage workers we use in our analysis.

	Pooled		1992		1998		2003	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	Obs.	Mean (Standard Deviation)	Obs.	Mean (Standard Deviation)	Obs.	Mean (Standard Deviation)	Obs.	Mean (Standard Deviation)
Log-wage‡	13,779	8.48 (1.66)	3,048	7.82 (1.67)	3,013	8.29 (1.72)	7,718	8.82 (1.53)
Schooling (*S*)	13,507	2.92 (4.14)	3,036	3.63 (3.70)	2,956	2.52 (4.37)	7,515	2.87 (4.15)
Age	13,775	29.39 (16.50)	3,048	36.81 (13.26)	3,009	38.65 (14.18)	7,718	27.01 (16.45)
Experience[4] (*E*)	13,775	19.28 (13.38)	3,048	18.91 (13.12)	3,009	20.78 (13.96)	7,718	18.84 (13.21)
Female	13,775	0.47 (0.50)	3,048	0.43 (0.50)	3,009	0.39 (0.49)	7,718	0.48 (0.50)
Rural	13,779	0.56 (0.50)	3,048	0.53 (0.50)	3,013	0.55 (0.50)	7,718	0.57 (0.50)
No Schooling	13,592	0.60 (0.49)	3,048	0.79 (0.42)	2,983	0.73 (0.45)	7,561	0.56 (0.45)
Primary School	13,592	0.18 (0.38)	3,048	0.05 (0.21)	2,983	0.08 (0.27)	7,561	0.21 (0.41)
Secondary Schooling	13,592	0.20 (0.40)	3,048	0.13 (0.33)	2,983	0.17 (0.38)	7,561	0.22 (0.41)
Tertiary Schooling	13,592	0.02 (0.12)	3,048	0.03 (0.16)	2,983	0.02 (0.15)	7,561	0.01 (0.11)

‡***Log-wage*** is the natural log of annual wage in 2003 Dalasis ($1=27 Dalasis in 2003).
No Schooling=1 if no schooling; ***Primary School***=equals 1 if the highest level of schooling is the primary level; ***Secondary School***=equals 1 if the highest level of school is the secondary level; and ***Tertiary Education***=equals 1 if the individual has tertiary level education. The variables ***No Schooling, Primary School, Secondary School*** and ***Tertiary Education*** sum to 1**.**

School and Education Data

Our analysis relies heavily on the historical education and population data. The data on the dates and location of school constructions comes from the Ministry of Education, which keeps a record of all formal schools ever constructed in the Gambia. For example, the first

[4] Experience is measured in years. We do not have data on real labor market experience acquired. Rather, we constructed it as follows: $E=Max\{0, \text{age-18}\}$. The choice of 18 is admittedly arbitrary. It is possible to acquire labor market experience before reaching 18, especially if one never attended school. We have used age instead of this constructed experience variable and the results are highly similar.

modern school was constructed in The Gambia in 1835. This is far earlier than the earliest date of birth for a worker in our sample, which is 1912. Our time series of district and national population data comes from the Central Bureau of Statistics. Because of its small size (the land size is 4000 square miles), collecting population data for the Gambia even in early 20th century was far less challenging than in many other African countries. Population figures for the-then Gambia colony[5] had been kept by the colonial authorities as early as 1900. In 1900, the population of The Gambia was approximately 110,000, and at independence in 1965 it had reached 407,800 people.

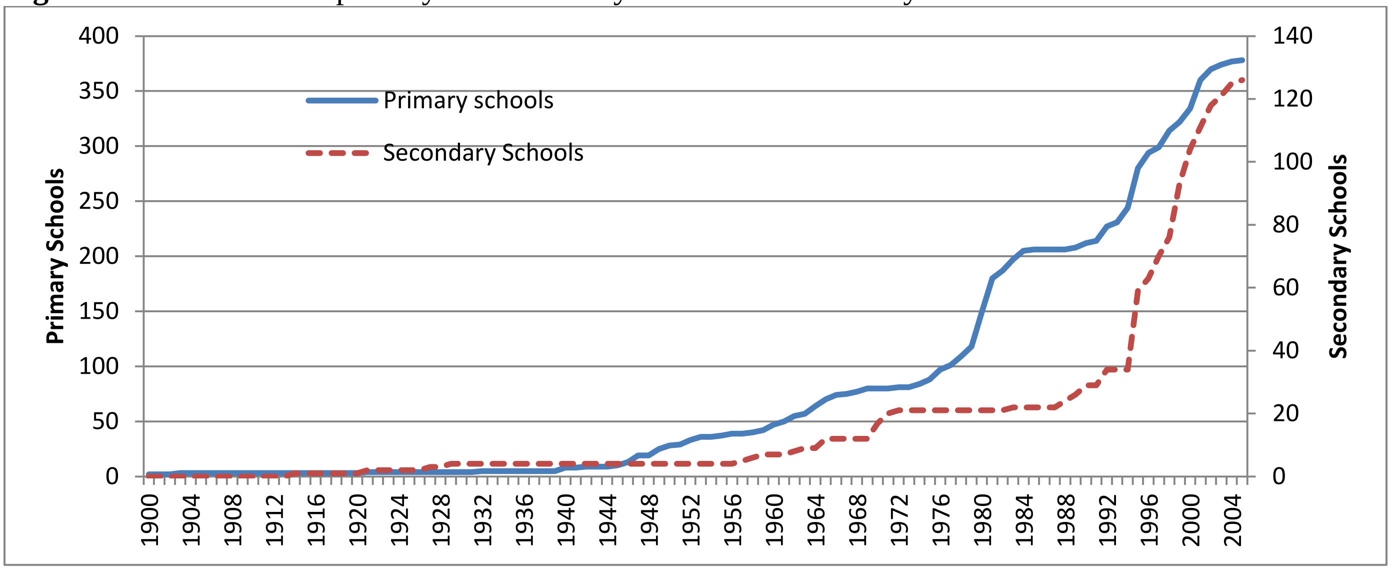
Given the history, schooling is naturally low in levels and even in growth.[6] The adult literacy rate in 1990 was 26%. Net enrollment rate at the primary level in 1991 was 48%. The pupil to teacher ratio increased from 31.3 in 1990 to 32.9 in 1998, due to a slow growth in enrollment combined with little investment in education, schools, or addition of teachers over time (World Bank 2010). While historical school attendance and achievement numbers are difficult to find, data on the dates of construction of every school in the formal education sector in the country does exist. The levels and densities of both primary and secondary schools from 1900 to 2005 are presented in figures 3 and 4.

In addition to the low total number of schools in the country, the distribution of schools has also been very unequal across regions, with a bias toward the capital and other urban areas. Table 2 shows the spatial school density per region. The farther away a region is from the capital, the smaller the number of schools per square kilometer.

[5] The Gambia was a British colony and protectorate. The small territory was divided into two parts: the colony and the protectorate. The colony covered the capital and coastal areas and was ruled directly by colonial administrators headed by a governor appointed from London. This area corresponds to the administrative areas of Region 1 and parts of Region 2. The interior was considered the protectorate, which was ruled indirectly through local chiefs and corresponds to parts of Region 2, and all of Regions 3, 4, 5 and 6 (Hughes and Perfect 2006).

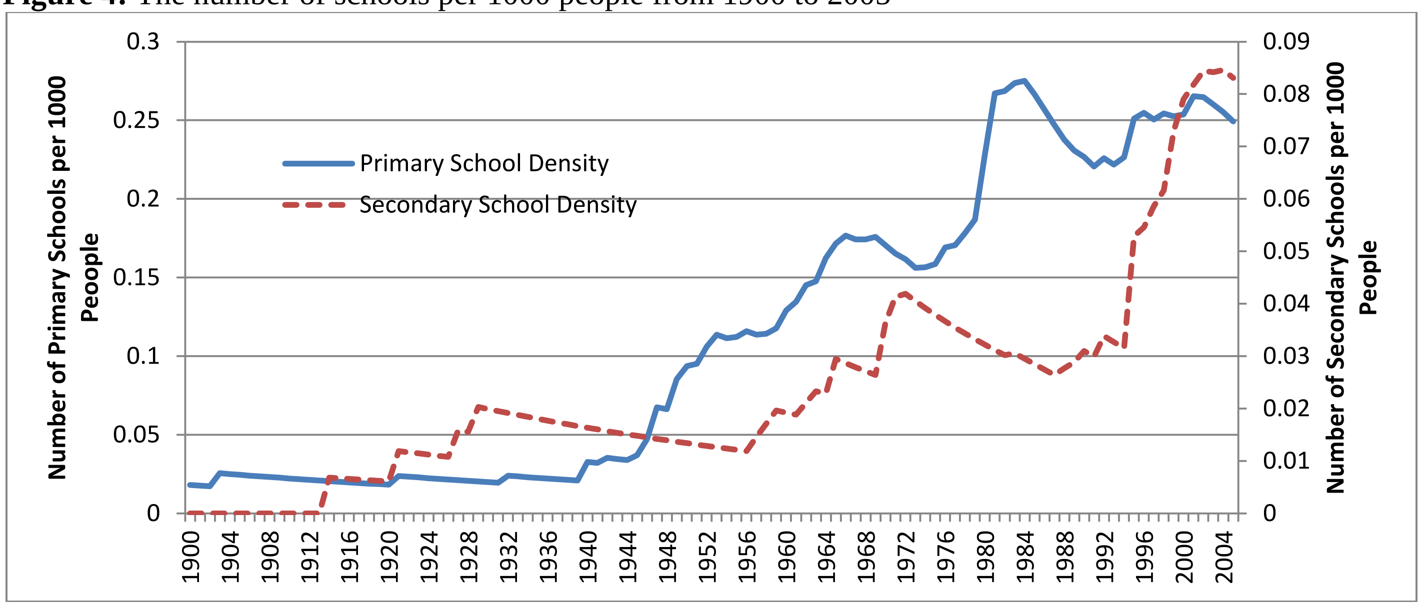
[6] Good data on enrollment, literacy and educational expenditure does not extend earlier than the 1990s.

Figure 3: The number of primary and secondary schools in the country from 1900-2005.



Source: school data comes from the Ministry of Education while population data comes from the Central Bureau of Statistics.

Figure 4: The number of schools per 1000 people from 1900 to 2005



Source: School data comes from the Ministry of Education while population data comes from the Central Bureau of Statistics.

Table 2: Number of Schools per 10 square kilometers

	1965-1969		1970-1974		1975-1979		1980-1984		1985-1989		1990-1994	
	Primary Schools	Sec. Schools	Primary Schools	Sec. Schools	Primary Schools	Sec. Schools	Primary Schools	Sec. Schools	Primary Schools	Sec. Schools	Primary Schools	Sec. Schools
Region 1	2.258	1.828	2.344	1.914	2.882	1.935	3.355	1.935	3.656	1.978	4.108	2.516
Region 2	0.114	0.012	0.123	0.029	0.136	0.029	0.189	0.029	0.205	0.029	0.215	0.043
Region 3	0.116	0.000	0.123	0.018	0.139	0.019	0.239	0.019	0.271	0.019	0.334	0.023
Region 4	0.049	0.000	0.051	0.005	0.067	0.006	0.172	0.006	0.174	0.009	0.175	0.013
Region 5	0.033	0.000	0.033	0.002	0.046	0.003	0.093	0.003	0.105	0.003	0.109	0.003
Region 6	0.021	0.000	0.026	0.000	0.057	0.000	0.125	0.002	0.155	0.005	0.169	0.005

III. Estimation Strategy and Results

To estimate the private rate of returns to education in the country, we start with the following standard Mincer-type equation:

$$\ln(Y_{ik}) = \alpha + \beta S_{ik} + \gamma_1 E_{ik} + \gamma_2 E_{ik}^2 + \chi_{ik}^{'} \varphi + \varepsilon_{ik}, \qquad (1)$$

where $\ln(Y_{ik})$ denotes the natural log of wage income, S_{ik} stands for years of schooling attained and E_{ik} denotes labor market experience for individual i in district k. The vector χ represents other determinants of earnings such as sex, rural residence, and regional location. Also included are the two year dummies (1998 and 2003), with 1992 being the excluded year.

A direct estimation of equation (1) would likely lead to a biased and inconsistent estimate of the rate of returns to education for a number of reasons. The first reason is that individuals in the wage labor market are unlikely to be a representative sample of the population. Specifically, participation in the labor market occurs only if the market wage is equal to or exceeds the individual's reservation wage. Secondly, since unobserved ability of individuals is likely to be correlated with schooling and wages, any estimate of β from equation (1) would likely suffer from omitted variable bias.

In order to solve the first potential bias with our data, let the labor market participation decision of individual i be given by:

$$L_{ik} = R_{ik}^{'}\psi + \chi_{ik}^{'}\varphi + \nu_{ik} \qquad (2)$$

where $L_{ik} = 1$ if an individual is in the labor market and 0 otherwise, while R_{ik} is a vector of variables that affect an individual's decision to enter the labor market through their effect on the reservation wage but do not directly affect the market wage. In this formulation, $L_{ik} = 1$ when we observe labor market participation indicating that the market wage exceeds the individual's reservation wage. In addressing the selection problem, the challenge is to identify variables in *R*, which we do by taking advantage of rainfall risk in the area since agriculture is the primary economic activity. Kijima, Matsumoto and Yamano (2006) in Uganda, Rose (2000) in India, Cameron and Worswick (2003) in Indonesia and Ito and Kurosaki (2006) in India all show significant responses of labor supply to rainfall risk in agriculturally dominated areas.

We use four different rainfall variables in *R*: rainfall shock (in district) in survey year, rainfall shock the year before, rainfall shock two years before and coefficient of variation of rainfall (summary statistics of these variables are provided in Table 3). We define a rainfall shock at the district level as the deviation of the year's rainfall from the five-year average. The coefficient of variation is the ratio of the standard deviation to the mean of rainfall in district over the previous five years. The rainfall shock variables provide aggregate risk to agricultural production (at the district level) that is unlikely to be mitigated by risk sharing among households because they face spatially covariant rainfall distributions. While the rainfall shock variables provide transient risk, the coefficient of variation of rainfall provides a relatively more permanent measure of risk (Rose 2000). Overall the rainfall variables are likely to affect the reservation wages of individuals through their effect on agricultural income, but unlikely to have a significant effect on current wages. Therefore, their effect is a “push” effect – that is, on labor supply.

The direction of the effect of rainfall shocks and coefficient of variation on labor market participation is ambiguous because households in The Gambia, as in many African countries (Barrett, Reardon and Webb 2001), have diversified livelihood strategies. For example, a negative rainfall shock can increase labor supply if the prevailing market wage exceeds the marginal returns to labor on the farm or in the non-farm enterprise operated by the household or individual. Conversely, a positive rainfall shock can reduce labor supply if the returns to farm labor exceed the prevailing market wage. Given that most households and individuals do not exclusively farm (or work exclusively as entrepreneurs or wage laborers), the rainfall effect on labor supply is an empirical question. The validity of the above exclusion variables (rainfall shocks and coefficient of variation) depends critically on their lack of direct effects on wages. In other words, these variables need to have significant effects on labor market participation but no direct effect on wages. In the appendix (table A1), we show that all the rainfall shock variables and the coefficient of variation of rainfall have no direct statistically significant effect on log wages.

There is also a potential “pull” effect – that is, a variable affecting labor demand, which can determine labor market participation and therefore would belong in R. We therefore add the proportion who are self-employed in each district as a measure of labor demand. The appendix, table A1, shows that like the rainfall data, business ownership in the district does not have a statistically significant effect on wages.

Using these "push" and "pull" variables, we estimate equation (2) as a *Probit* model to obtain the Inverse Mills Ratio from labor market participation. The Inverse Mills Ratio can be used to augment equation (1) to address the potential selection bias from only observing wages from the labor market participation. The result in table 4 shows the relevance of the above selection instruments. A rainfall shock in the survey year has a positive effect[7] on labor market participation but it is not statistically significant, while rainfall shocks in the preceding year and two years before both have significant and negative effects on labor market participation. We also found that the coefficient of variation of rainfall decreases labor market participation, which is consistent with Ito and Kurosaki (2006). And finally, higher numbers of businesses in the district is correlated with higher labor force participation. Table 4 shows that years of schooling (*S*) has a U-shaped relationship with labor force participation. Without adding a quadratic term for the number of years of schooling as we do in column 2, the result in column 1 alone would have implied a counter intuitive relationship.

Table 3: Summary Statistics of identifying variables in equation in *R* in equation (2).

Variable	**Mean**	**Standard Deviation**
	1	2
Coefficient of Variation of Rainfall	0.24	0.13
Current Year Rainfall Shock	-16.90	240.06
Rainfall Shock in Preceding year	-166.16	146.61
Rainfall Shock in 2 Years Earlier	-41.90	126.34
Proportion of Business Owners in district	0.06	0.13

[7] In other words, an above average rainfall in district is associated with employment in the wage sector.

Table 4: Relevance of Selection Instruments. Below are the *probit* results (marginal effects). The dependent variable is whether the individual is in the labor market. Robust and clustered standard errors are in parentheses. Note that the number of observations here far exceeds the sample in the wage regression because we use the whole sample.

	Whole Sample	
	1	**2**
Schooling (*S*)	-0.002*** (0.0005)	-0.020*** (0.002)
Schooling Squared (S^2)		0.002*** (0.0001)
Age	0.041*** (0.0004)	0.037*** (0.001)
Age sq.	-0.0004*** (0.0000)	-0.0004*** (0.00001)
Female Dummy	-0.109*** (0.004)	-0.112*** (0.004)
Rural Dummy	-0.027*** (0.005)	-0.039*** (0.006)
1998 Dummy	-0.026* (0.015)	-0.021* (0.012)
2003 Dummy	-0.027* (0.016)	-0.017* (0.009)
Coefficient of Variation of Rainfall	-0.152*** (0.032)	-0.581*** (0.091)
Current Year Rainfall Shock	0.00001 (0.0000)	0.0001*** (0.00002)
Rainfall Shock in Preceding year	-0.0001** (0.00003)	-0.0004*** (0.00005)
Rainfall Shock in 2 Years Earlier	-0.0002*** (0.00003)	-0.0002*** (0.0001)
Proportion of Business Owners in district	0.292*** (0.046)	0.307*** (0.089)
Observations	50633	50633
Log Likelihood	-20253	-20334

***significant at 1% level; **significant at 5% level; *significant at 10% level. The excluded year dummy is 1992.

In order to address the possibility of an omitted variable bias due to unobserved ability, we use an instrumental variable estimation method that exploits the variation in access to schooling at the time of an individual's birth. We accomplish this by interacting the primary and secondary school densities in district with year of birth of each individual[8]. Our first stage equation is

$$S_{ik} = \theta_1 + \theta_2 P_{ik} + \theta_3 P_{ik}^2 + \theta_4 M_{ik} + \theta_5 M_{ik}^2 + \chi_{ik}^{'}\varphi + \mu_{ik} \tag{3}$$

where P_{ik} and M_{ik} respectively denote the densities of primary and secondary schools in district k the year individual i was born[9]. As is evident from equations (1) and (3), the excluded instruments are the densities of primary (*P*) and secondary schools (*M*) and their quadratic terms. For these variables (*P* and *M*) to serve as proper instruments, they most be highly correlated with *S* and be uncorrelated with ability in ε in equation (1)[10].

The first requirement of our identification strategy concerns the relevance of the instruments. The proximity to schools, which partially proxies the cost of schooling,[11] is likely to directly influence the probability of parents sending their children to school. This requirement is also directly testable. First, we show that the school proximity is indeed a significant determinant of schooling as shown in the results in Table 7. Both measures of school density and their quadratic terms show significant effects on educational attainment among wage earners.

[8] In this way, the instruments capture both the effects of school access as well as any year effects. They are similar in spirit to the instruments used in Duflo (2001) and Oyelere (2010)

[9] Our findings in the results section are unchanged if we instead use school densities in districts at the time individual was 6 years old - that is, a year before students can be enrolled in primary school.

[10] This second stage equation (equation 1) does not control for age. This is because age and experience are highly correlated in our sample (the correlation coefficient is 0.9). However, in a later section, we estimate the wage equation for different age cohorts.

[11] In areas without schools, parents who want to send their children to school will foster them out to other families, sometimes related, sometimes not, in other towns where there are schools. This child fostering inevitably imposes additional costs relative to keeping a child at home and sending them to a nearby school.

Furthermore, in the instrumental variable estimation of equations (1) and (3) jointly, we provide the F-test results for the joint statistical significance of $\theta_2, \theta_3, \theta_4$ and θ_5. These coefficients are jointly significant statistically, showing that our instruments are highly correlated with the level of schooling attained.

Another condition for the consistent estimation of β is that these instruments (*P* and *M*) are uncorrelated with ability, which is relegated in ε_{ik}. In other words, this exogeneity assumption implies that school density in the district where the individuals were born is uncorrelated with ability and current wages. While this requirement is not directly testable, we make the case that the likelihood of the condition being violated is low in our empirical strategy. Because school density differs from district to district and is systematic while ability is likely to be randomly distributed in the population, it is unlikely that variations in the density of schools are correlated with an unobservable such as ability.[12] As table 2 makes clear, the number and density of schools have been very low in The Gambia. Because the exposure to schooling is very limited in all regions of the country, the low average level of schooling makes it unlikely that only high ability individuals would have access to and attend school.

On the requirement that there should be no correlation between school density in an individual's district and her current wage, we acknowledge that this probability is not as low as that of the requirement that ability and school density having low correlation. We, nevertheless, present some evidence that observed current wages are not likely to be correlated with school density in individual's district at her time of birth. While the correlation of school densities over

[12] It is unlikely that the colonial government was responsive to local needs in terms of where schools should be built. And while it is likely that the post-independence governments will favor urban areas for many public projects, it is hard to see how the distribution of these public projects is correlated with ability.

time is unavoidable, we show in table 5 that there is a convergence in density of schools per region. For example, while Region 3 had the second lowest primary school density in 1960, it ended up having the highest density in 1990. In other words, while regions close to the coast had a relatively higher number of schools initially, they have also experienced tremendous increases in population as shown by their explosion in population density between 1960 and 1990. And during this same period, the number of schools in the more remote regions has started to increase significantly. So while we acknowledge the possibility of correlation between school density in district at the time of the individual's birth and her current wage, any such correlation is likely to be mitigated by rapid population growth.

Table 5: Primary School and Population Density in Regions between 1960 and 1990.

		School Density (per 1000 people)				Population Density (per sq. km)			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
		1960	**1970**	**1980**	**1990**	**1960**	**1970**	**1980**	**1990**
Region1	Primary	0.281	0.134	0.112	0.090	323	840	1427	2906
	Secondary	0.211	0.104	0.064	0.051				
Region2	Primary	0.236	0.299	0.267	0.209	29	53	70	138
	Secondary	0	0.075	0.049	0.039				
Region3	Primary	0.149	0.225	0.272	0.307	24	27	47	42
	Secondary	0	0.024	0.028	0.021				
Region4	Primary	0.126	0.250	0.269	0.270	20	25	32	101
	Secondary	0	0	0.019	0.032				
Region5	Primary	0.201	0.143	0.202	0.258	21	33	41	51
	Secondary	0.017	0.022	0.025	0.020				
Region6	Primary	0.054	0.064	0.154	0.218	27	43	50	78
	Secondary	0	0	0	0.007				

Table 6: Summary Statistics of Excluded Instruments

Variable	Obs.	Mean	Standard Deviation
	1	**2**	**3**
Primary School Density (*P*)	13770	0.222	0.124
Secondary School Density (*M*)	13728	0.092	0.117

Primary School Density (*P*) is the number of primary schools per 1000 people in district when individual was born.
Secondary School Density (*M*) is the number of secondary schools per 1000 people in district when individual was born.

Even after estimating returns to education with the above corrections and adjustments, we are still left with the issue of how to properly interpret the estimate for β in equation (1) after an IV estimation. The consistent estimation of β requires that the marginal returns to education be similar for all individuals (Card 2001). However, if returns to education are heterogeneous across individuals, then the results of our IV estimation would give us the weighted average of the returns to schooling of individuals induced to attend school by the increase in school densities. This latter outcome would be equivalent to local average treatment effect or LATE (Card 2001; Deaton 2010; Imbens and Angrist 1994). In other words, our estimates of the returns to education may be restricted to the sub-sample of the population on the cusp of school attendance or enrollment and would be driven by the elasticity of that enrollment with respect to access. This latter point has some implication for the comparability of estimates of returns to education across different studies using different estimation strategies. As long as different instrumental variables (in this paper it is school construction but could be lower school fees in others) have different elasticities, the comparability of estimated returns to education would be limited. In other words, the LATE will likely differ based on the type of instrument used and its effect on the level or years of education attained. This should be kept in mind in comparing returns to education across countries and studies.

Table 7: Determinants of Schooling Attainment (for wage earners only). The dependent variable is the years of schooling (*S*). Robust and clustered standard errors are in parentheses. See previous the preceding figures in table 6 for summary statistics of variables *P* and *M*.

	OLS
	1
Age	0.013 (0.010)
Age squared	-0.001 (0.0001)***
Rural Dummy	-1.252 (0.095)***
Female Dummy	-1.535 (0.067)***
1998 Dummy	-1.012 (0.092)***
2003 Dummy	-1.289 (0.079)***
Secondary School Density (M)	14.364 (2.085)***
Secondary School Density Sq. (M^2)	-20.136 (4.681)***
Primary School Density (P)	3.069 (1.456)**
Primary School Density (P^2)	-6.546 (3.639)*
Constant	4.445 (0.395)***
Regional dummies	Yes
Observations	13457
F-test p-values§	23.11 (0.000)
R squared	0.19

***significant at 1% level; **significant at 5% level; *significant at 10% level. The excluded year dummy is 1992.

§The F-test is the test of the joint significance of the coefficients on M, M^2, P and P^2.

IV Results:

Table 8 presents our estimates of the private rate of returns to education in the wage sector in The Gambia. This table shows our second stage results from equation (1) and also the first stage results (equation 3). The p-values of the Hansen-Sargan statistics validate our excluded instruments and the high value of the F-test statistic of the joint significance suggests they are highly relevant. The statistical significance of the Inverse Mills Ratio also suggests that the sample selection effect is non-trivial.

The OLS estimate of the rate of returns to schooling for the pooled sample is 6.8% without selection correction and 7.4% with selection correction, which are not significantly different from each other statistically. The IV estimates of the rate of returns to schooling are 24.1%, 19.1%, and 35.8% depending on specification. Our preferred rate of returns to education for the country is 24.1% in column 3 of table 8 because this particular result addresses both selection and omitted variable bias.

The results by gender, in table 8, show that the rate of returns to education is higher for males (35.8%) than for females (16%)[13]. This is surprising because the average level of schooling of males is higher than that of females in the sample. Assuming diminishing marginal returns, then one would expect the returns to be higher for females relative to males. However, the expectation of higher female marginal returns assumes the absence of unobserved gender discrimination. Specifically, societal norms could counteract the expected higher female returns by creating gender barriers in certain higher-return occupations. In an experimental study of

[13] We use interaction terms in columns 4 and 5 of table 8 to allow the marginal returns to education to vary by gender and location. While this procedure forces the coefficients on the other variables to be the same for both gender and rural dummy, this decision is not very limiting since we are not particularly interested in gender differences in variables such as experience, rural and year dummies.

returns to capital among Sri Lankan micro-entrepreneurs, Del Mel et al. (2009) found that the marginal returns to capital is higher for males than females even though the former have larger capital sizes. While they found that this gender difference in marginal returns cannot be explained by differences in ability or risk preferences, it is consistent with differential concentration in industries and relative intra-household bargaining power. It is also worth pointing out that Schultz (2004) also found that the returns to education for males are higher than females across age cohorts in another country in West Africa, Ivory Coast.

The higher marginal returns to education for urban residents (19.1%) relative to rural workers (7.3%) is also counter-intuitive given the relatively higher average educational attainment in urban areas. While we would expect higher marginal rate of returns at lower levels of schooling, the kinds of occupations available in rural areas are unlikely to provide high returns to education relative to those in urban areas. For example, high return government jobs are more likely to be located in urban areas. The significant difference in returns to education between rural and urban areas suggests some heterogeneity across regions. We present the estimations by regions in table 9 and indeed find large heterogeneity. It appears that the high returns to education are driven primarily by 2 out of the 6 regions[14]. For the four other regions, the rate of returns to education ranges from 8% to 16%.

The heterogeneity of the rate of returns within a small country such as the Gambia suggests that the variation in returns within individual countries may be higher than variability between countries (Psacharopoulos, 1994). Our estimate of the rate of returns to education in

[14] The population share of regions 1, 2, 3, 4, 5 and 6 are 26%, 29%, 13%, 5%, 14% and 13% respectively.

four out of six regions is very similar to the single digit estimated values for other West African countries in the region (see for example: Oyelere, 2010; Kazianga, 2004 and Glewwe, 1996).

What accounts for the heterogeneity in our estimated rate of returns to education when we use the IV approach? As can been seen in table 9, our measure of the rate of returns to education with OLS is very similar across regions. The interpretation of our estimated returns to education is affected by the potential heterogeneity of the marginal returns to schooling induced by possibly different responses to school construction and access. Given that different regions of The Gambia had different levels of school densities at any point in time, it is unlikely that changes in school densities would induce identical enrollment in all districts and regions. Given this fact, the appropriate interpretation of our IV estimate is the local average treatment effect (Card 2001; Deaton 2009; Imbens and Angrist 1994). In other words, our estimate provides the returns to education on the subset of the population who attained some schooling but would otherwise not have, had there been no change in construction of schools in their districts. This interpretation of the IV estimate also suggests that there may be limits to comparing estimated returns to education across different countries due to differences in estimation strategies. Even studies carried out on the same population that use different instruments are unlikely to have comparable estimates of the return to an additional year of schooling.

Another factor that can account for the significant differences in returns to education across regions is the violation of the exclusion restriction in our IV estimate strategy. Specifically, historical school densities may affect current wages through channels not restricted to attained schooling (*S*). This could result from differences in levels of regional development or urbanization - initial differences that can persist through time across regions and districts[15]. This is consistent with the low Hansen-Sargan statistics in column 4 of table 8 and the results for regions 4 and 5 in table 9. However, even when we exclude regions 4 and 5, our returns to education estimate for the remaining sample is 24%, almost identical with the value in column 3 of table 8.

[15] We would like to thank an anonymous reviewer for alerting us to this possibility and its implications for our results.

Table 8: Wage regression results. The dependent variable is the natural log of wage income. In obtaining returns to educations by gender and location (columns 4 and 5), note the interaction terms. Robust standard errors (clustered at district level) are in parentheses.

	OLS		**IV**		
	1	**2**	**3**	**4**	**5**
Schooling (*S*)	0.068*** (0.005)	0.074*** (0.004)	0.241*** (0.064)	0.191** (0.074)	0.358*** (0.108)
Experience (*E*)	0.049*** (0.005)	0.174*** (0.053)	0.366*** (0.117)	0.297** (0.119)	0.333*** (0.105)
Experience Sq.	-0.001*** (0.000)	-0.003*** (0.001)	-0.006*** (0.002)	-0.005** (0.002)	-0.005*** (0.002)
Female Dummy	-0.737*** (0.088)	-1.078*** (0.126)	-1.370*** (0.245)	-1.349*** (0.266)	-0.733*** (0.179)
Rural Dummy	-0.603** (0.230)	-0.580** (0.227)	-0.339 (0.258)	-0.263 (0.363)	-0.278 (0.288)
1998 Dummy	0.421*** (0.085)	0.193 (0.139)	-0.044 (0.211)	-0.014 (0.223)	0.082 (0.198)
2003 Dummy	0.899*** (0.138)	0.670*** (0.126)	0.466*** (0.129)	0.511*** (0.160)	0.596*** (0.123)
Schooling*Rural				-0.118 (0.077)	
Schooling*Female					-0.190** (0.083)
Inverse Mills Ratio	No	1.368*** (0.573)	3.460** (1.307)	2.726** (1.327)	3.093** (1.175)
Constant	8.204*** (0.073)	5.914*** (0.972)	1.346 (2.526)	3.001 (2.521)	1.362 (2.467)
First Stage					
Primary School Density (*P*)			3.094 (1.563)**	3.158 (1.611)**	3.156 (1.586)**
Primary School Density sq. (P^2)			-6.530 (3.819)*	-6.546 (3.482)*	-6.539 (3.653)*
Secondary School density (*M*)			14.456 (3.623)***	14.461 (2.175)***	14.377 (3.423)***
Secondary School density (M^2)			-20.069 (6.299)***	-20.072 (6.049)***	-20.123 (5.967)***
R squared	0.39	0.38			
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Observations	13503	13033	12989	12989	12989
Hansen-Sargan† statistic (p-value)			5.59 (0.13)	8.18 (0.04)	4.52 (0.21)
F-test statistic§ (p-value)			23.10 (0.00)	17.08 (0.00)	14.22 (0.00)

***significant at 1% level; **significant at 5% level; *significant at 10% level. The excluded year dummy is 1992.
§This is the test on whether the excluded instruments (which appear only in the first stage) are jointly significant.
†This is Hansen-Sargan test with the null hypothesis that the instruments are valid.

Table 9: The Rate of Returns to Education by Regions. The dependent variable is log of wage income. All controls in Table 8 are also included here.

	Region 1		**Region2**		**Region3**		**Region4**		**Region5**		**Region6**	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Schooling(*S*)	**0.06*** **	**0.08** **	**0.06*** **	**0.48** **	**0.06*** **	**0.14** **	**0.06*** **	**0.16* **	**0.11*** **	**0.43** **	**0.07*** **	**0.14** **
						First Stage						
Primary School density (*P*)		3.42**		3.08***		5.60**		6.68**		3.24**		7.40**
Primary School density sq. (P^2)		-5.61*		-6.45*		-5.80**		-3.887**		-5.96*		-3.16*
Secondary School density (*M*)		15.34**		14.34***		16.64*		16.67**		14.58**		17.52**
Secondary School density sq. (M^2)		-19.83*		-20.19*		-18.27*		-17.33		-17.97		-17.15**
F-Test (p-value)		24.02 (0.00)		23.65 (0.00)		25.53 (0.00)		25.53 (0.00)		23.63 (0.00)		21.08 (0.00)
Hansen-Sargan statistic (p-value)		4.27 (0.23)		3.71 (0.30)		5.70 (0.12)		6.72 (0.08)		9.35 (0.02)		5.48 (0.14)
Uses Controls in Column 3 of Table 8	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3664	3658	4004	3985	1626	1622	496	493	2234	2225	1479	1474
R sq.	0.18		0.21		0.31		0.48		0.24		0.35	

***significant at 1% level; **significant at 5% level; *significant at 10% level.

Robustness Checks on Rate of Return to Schooling

Since the effect of differences in learning in first and second grade is not the same as in sixth and seventh grade, it is also possible that the rates of returns to schooling differ at various levels of schooling. In our main estimation, we did not include higher order terms for the years of schooling (*S*) variable in the second stage estimation (equation 1). To determine if this is the case, we estimate the relationship between years of schooling and log wage with a semi-parametric technique that does not force any linearity assumption between the two variables.

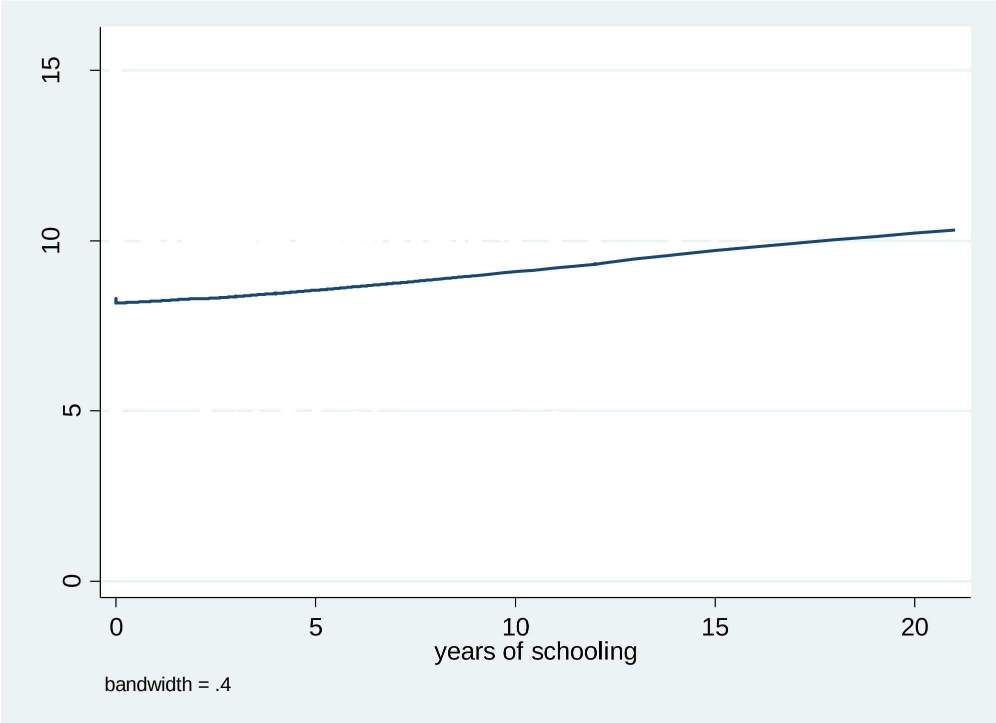
Equation (4) estimates a semi-parametric equation, which admits a flexible function of years of schooling, $f(S_i)$ and allows other relevant individual variables (age, experience, gender, rural, regional and year dummies) to enter linearly (Lokshin 2006).

$$\ln(Y_i) = f(S_i) + X_i^{'}\beta + \mu_i \qquad (4)$$

X represents a vector of individual variables and μ_i is a zero-mean error term. The variables that enter linearly are the same controls in the results presented in column 3 of Table 8 except for the interaction terms and the inverse mills ratio. The semi-parametric procedure estimates $\ln(Y_i) = X_i^{'}\beta + \mu_i$ and then smooths $\ln(Y_i) - X_i^{'}\widehat{\beta} = f(S_i) + \mu_i$ (where $\widehat{\beta}$ is the estimation of β) using locally weighted scatter-plot regression (LOWESS).

Figure 5 presents the non-parametric function, $f(S_i)$ that results from estimating equation (4). The estimated figure is approximately linear and does not show dramatic changes in slope. It is therefore unlikely that our earlier estimation obscures significant heterogeneity in the rate of returns at different levels of schooling.

Figure 5: Semi-parametric estimation of the relationship between schooling and log-wage (N=13,496).



However, the above semi-parametric figure does not address the potential bias that motivated our IV approach. Following the literature, we also estimate returns to education using levels of schooling rather than years of education. This could be important if the attainment of certain educational levels comes with a credential effect. For example, while an individual with 5 years of schooling is only 2 years behind another with 7 years, the differences in the labor market could be significant given the latter is in possession of a primary school leaving certificate.

To estimate returns to education using levels, *S* in equations (1) and (3) is now considered a vector of schooling levels. These levels of schooling, which are dummy variables, are primary (1 to 6 years of schooling), secondary (7 to 12 years of schooling) and tertiary (greater than 12 years of schooling). The excluded dummy is no schooling (0 years of schooling).The average years of schooling for those with primary, secondary and tertiary education levels are 3.4, 9.6 and 14.4 respectively.

Table 10 presents the results of our estimation using levels of schooling. As with our results from using years of schooling, the IV estimates are much higher than the OLS estimates. However, the OLS estimates in tables 10 are similar to those in table 8 following the method by Shultz (2004). For example, calculating the implied returns to an additional year of education for secondary level using the OLS estimates is: (1.173-0.144)/(9.6-3.4)=0.08. Following the same calculation, it can be shown that the returns to a year of education at the primary and tertiary levels for the OLS results are 0.05 and 0.11 respectively. For the IV results, the implied estimates for a year of schooling at the primary, secondary and tertiary levels are 0.24, 0.13 and 0.21 respectively.

Table 10: Returns to Education by Level of schooling. The dependent variable is log wage. The reference education level (omitted category) is no school. Robust standard errors (clustered at district level) are in parentheses.

	OLS	IV
	1	2
Primary	0.144 (0.072)**	0.825 (0.415)**
Secondary	0.647 (0.052)***	1.607 (0.673)**
Tertiary	1.173 (0.064)***	2.633 (1.471)*
Experience	0.052 (0.015)***	0.164 (0.052)***
Experience Sq.	-0.003 (0.001)***	-0.001 (0.0002)***
Female Dummy	-0.386 (0.193)**	-1.070 (0.131)***
Rural Dummy	-0.600 (0.224)**	-0.290 (0.227)
1998 Dummy	0.202 (0.137)	0.260 (0.173)
2003 Dummy	0.673 (0.129)***	0.793 (0.143)***
Inverse Mills Ratio	1.276 (0.568)***	2.845 (1.221)**
Constant	6.149 (0.961)***	6.815 (0.743)***
Observations	13503	12989
R Squared	0.37	

	First Stage		
	Primary Dummy	Secondary Dummy	Tertiary Dummy
Primary School Density (*P*)	0.143 (0.046)***	0.211 (0.114)*	0.165 (0.497)
Primary School Density sq. (P^2)	-0.028 (0.014)**	0.111 (0.080)*	0.012 (0.018)
Secondary School Density (*M*)	0.131 (0.060)**	0.166 (0.083)**	0.198 (0.094)**
Secondary School Density (M^2)	-0.012 (0.009)	-0.201 (0.106)*	-0.172 (0.090)**
F-test statistic§ (p-value)	21.67 (0.000)	18.75 (0.000)	15.83 (0.000)
Hansen Sargan† (p-value)		0.27 (0.59)	

***significant at 1% level; **significant at 5% level; *significant at 10% level. The excluded year dummy is 1992.§This is the test on whether the excluded instruments (which appear only in the first stage) are jointly significant.† This is Hansen-Sargan test with the null hypothesis that the instruments are valid.

There is another potential issue with our estimates of returns to schooling: school quality. It is possible that differences in school quality may affect the rate at which parents send their children to school. If there are significant differences in school quality and it is not controlled for, this could bias our results.

Our data did not include any measure of school quality, such as student teacher ratios, which are commonly used in the literature, and we therefore cannot directly assess the effect of school quality. But because the available evidences suggests few year to year differences in student teacher ratios, we do not suspect that controlling for school quality through student teacher ratios as commonly practiced in the literature would significantly change our results.

Nevertheless, if there were significant school quality changes over time we hypothesize that they would be evident in different returns across student cohorts. If school quality changed significantly over time, it is plausible to suspect that this would be reflected in different rates of returns to education across different age cohorts. Table 11 shows the results of our main wage regression run separately for different age cohorts. Returns to education are still significant and high for all cohorts. Furthermore, none of the estimated coefficients of schooling (i.e. the rate of returns to education) of the different cohorts are statistically different from our main estimate (24.1%) in column 3 of table 8.

Table 11: Returns to Education by Age Cohorts. This regression has the same controls in Table 8 but we only show the coefficient on *schooling* (*S*). Robust standard errors (clustered at the district level) are in parentheses.

	Age 25 or younger	Older than 25 but younger than or equal to 35	Older than 35 but younger than or equal to 45	Older than 45 years
IV				
	1	**2**	**3**	**4**
Schooling (*S*)	0.232** (0.069)	0.331*** (0.098)	0.193** (0.078)	0.241** (0.098)
Observations	2622	4444	3239	3152
Hansen-Sagan[†] statistic (p-value)	1.04 (0.79)	0.84 (0.83)	1.11 (0.77)	1.22 (0.74)
First Stage				
Primary School Density (*P*)	3.541 (1.069)***	3.969 (1.195)***	3.019 (1.003)***	3.518 (1.067)***
Primary School Density sq. (P^2)	-4.377 (2.199)**	-3.721 (1.066)***	-5.091 (1.562)***	-4.380 (2.212)**
Secondary School Density (*M*)	14.905 (6.714)**	15.315 (7.696)**	14.514 (7.367)**	14.899 (6.681)**
Primary School Density sq. (M^2)	-17.914 (10.178)*	-17.684 (8.976)**	-18.327 (9.118)**	-17.908 (10.005)*
F-test[§] Statistic (p-value)	10.64 (0.00)	7.78 (0.00)	6.19 (0.00)	20.37 (0.00)
Uses Controls in Column 3 of Table 8?	Yes	Yes	Yes	Yes

***significant at 1% level, **significant at 5% level, *significant at 10% level
[§]This is the test on whether the excluded instruments (which appear only in the first stage) are jointly significant.
[†]This is Hansen-Sargan test with the null hypothesis that the instruments are valid.

Discussion

All our IV estimates show high rates of returns to education. This leads to the question of why we observe such a high rate of return and yet low school attendance. The first possible answer is that sending a child to school remains very costly for the average Gambian household. The direct cost of schooling involves not only uniforms, books, paper, and other stationery supplies, but also school fees at the middle and high school levels. The indirect cost stems from the loss of child labor on the family farm and as household labor for getting wood and water, which can be substantial for rural households. To the extent that households are credit constrained they are also unlikely to be able to borrow against the future increased earnings to finance current educational expenses.

Another obstacle to schooling is the possibility of perceived low returns to schooling. This may seem surprising since we just documented high returns to schooling in the wage sector. However, the wage sector is relatively small in the country. Farming is the more dominant livelihood in the Gambia. In anticipating future returns to schooling, the agricultural sector and the informal work sectors are likely to be weighted more than the formal wage sector in a parents' expectation of future returns to education.

The preceding paragraph begs the question of what is the reason behind the likely low returns to schooling in the agricultural or informal sectors. One obvious reason to expect high returns to education in the agricultural sector is through its link to increased technology adoption. But there has been no significant technical change in the agricultural sector in The Gambia. Chavas et al. (2005) found that technical efficiency in Gambian farm households is almost 100%. This is what one would expect when there is little or no new technology to learn and thus little heterogeneity in technique or ability to master the technology.

V. Conclusion

We provide the first estimates of the private rate of returns to education in The Gambia. Using three nationally representative household surveys and exploiting the exogenous variations in the availability of schools across regions when individuals were born, we are able to obtain consistent estimates of returns to education. Our IV estimate of the rate of returns to education for an additional year of schooling is 24.1%. However, this figure seems to be masking some significant heterogeneity across regions in the country. In most regions, the rate of returns to education ranges from 8% to 16%. This figure is higher than other estimates of the private rate of returns to education in developing countries in general (Psacharopoulos, 1994) and many recent estimates for West Africa in particular. It is also worth point out that the high inter-regional differences in returns to schooling could be due to the violation of the exclusion restrictions in some districts rather than inherent differences in rates of returns to education.

The combination of high estimated returns to education with low levels of school attendance that are evident in our results suggest that the presence of constraints may prevent households from fully exploiting the high returns to schooling. School attendance is highly correlated with proximity to schools and parents directly list cost as one of the reasons for not enrolling their children. Our results are also consistent with the untested possibility that households discount the high rate of returns to education in the wage sector because it is a very small sector relative to agriculture in the Gambian economy. This effect may be exacerbated by the fact that the agricultural sector in the country has not experienced significant technical change that is likely to reward education.

The results presented here imply that there is a large scope for interventions in the education sector to have significant benefits in The Gambia. Most directly, improving access to schools through construction and staffing of schools as well as reducing direct and indirect costs of schooling can have direct effects on children's propensity to attend and have long-term returns for individuals and the country. In terms of future research, this work raises a number of questions about the returns to schooling in agriculture and the informal sector. It also raises a number of important questions on the tradeoffs of schooling and child labor in The Gambia.

Future work investigating the effects of school access on child labor use in West Africa would also be a welcome addition to the literature.

References

Barrett, C.B., T. Reardon and P. Webb. 2001. "Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics and Policy Implications", *Food Policy*, 26(4): 315-331.

Baum, C. F., M.E. Schaffer and S. Stillman. 2007. "Enhanced Routines for Instrumental Variables/Generalized Methods of Moments Estimation", *Stata Journal*, 7(4): 465-506.

Cameron, L.A. and C. Worswick. 2003. "The Labor Market as a Smoothing Device: Labor Supply Response to Crop Loss" *Review of Development Economics*, 7(2): 327-341.

Card, D. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, 69(5): 1127-1160.

Chavas, J.P., R. Petrie and M. Roth. 2005. "Farm Household Production Efficiency: Evidence from The Gambia" *American Journal of Agricultural Economics*, 87: 160-179.

Deaton, A. 2010. "Instruments, Randomization, and Learning about Development", *Journal of Economic Literature*, 48: 424-455.

Del Mel, S., D. Mckenzie and C. Woodruff. 2009. "Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns", *American Economic Journal: Applied Economics*, 1(3):1-32.

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

Glewwe, P. 1996."The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment." *Journal of Development Economics*, 51(2): 267-290.

Hughes, D. and D. Perfect. 2006. "A Political History of The Gambia, 1816-1994", University of Rochester Press.

Imbens, G. and J. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62(2): 467-476.

Ito, T. and T. Kurosaki. 2006. "Weather Risk and the Off-farm Labor Supply of Agricultural Households in India", Hitotsubashi University.

Kazianga, H. 2004. "Schooling Returns for Wage Earners in Burkina Faso: Evidence from the 1994 and 1998 National Survey ", Economic Growth Center, Yale University.

Kijima, Y., T. Matsumoto and T. Yamano. 2006. "Non-farm Employment, Agricultural Shocks, and Poverty Dynamics: Evidence from Rural Uganda", *Agricultural Economics*, 35(3): 459-467.

Krueger, A. and M. Lindahl. 2001. "Education for Growth: Why and for Whom", *Journal of Economic Literature*, 39:1101-1136.

Lokshin, M. 2006. "Difference-based semiparametric estimation of partial linear regression models" *Stata Journal*, 6(3): 377-383.

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

Oyelere, R.U. 2010. "Africa's Education Enigma? The Nigerian Story", *Journal of Development Economics*, 91(1): 128-139.

Rose, E. 2000. "Ex-ante and ex post Labor Supply Response to Risk in a Low-Income Area", *Journal of Development Economics*, 64(2): 371-388.

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

Schultz, P. 2004. "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education", *Journal of African Economies*, 13(2):95-148.

Siphambe, H.K. 2000. "Rates of Return to Education in Botswana", *Economics of Education Review*, 19(3): 291-300.

World Bank, 2010.World Development Indicators, Accessed online at Woldbank.org.

Table A1: The dependent variable is log wage. The result in this table supports our identification strategy to control for selection. The excluded variables in the first stage selection estimation, equation 2 (*Biz ownership*, *shock_t_0*, *shock_t_1*, *shock_t_2* and *CV_Rainfall*) show no direct significant effect on log wages.

	OLS		
	1	**2**	**3**
Current Year Rainfall Shock	0.0001		0.0001
(*shock_t_0*)	(0.0004)		(0.001)
Rainfall Shock in Preceding year	-0.0003		-0.001
(*shock_t_1*)	(0.001)		(0.010)
Rainfall Shock 2 Years Earlier	0.001		-0.006
(*shock_t_2*)	(0.001)		(0.005)
Coefficient of Variation of Rainfall	0.632		0.628
(*CV_Rainfall*)	(0.971)		(0.966)
Proportion of Business Owners in		0.016	0.021
district (*Biz_Ownership*)		(0.286)	(0.105)
Schooling (*S*)	0.066	0.073	0.072
	(0.006)***	(0.006)***	(0.008)***
Experience (*E*)	0.049	0.044	0.040
	(0.005)***	(0.005)	(0.004)***
Experience Squared	-0.001	-0.001	-0.001
	(0.0001)***	(0.0001)***	(0.0001)***
Female Dummy	-0.727	-0.781	-0.720
	(0.087)***	(0.085)***	(0.097)***
Rural Dummy	-0.615	-0.598	-0.611
	(0.217)**	(0.239)***	(0.210)**
1998 Dummy	0.543	0.560	0.550
	(0.197)**	(0.243)***	(0.204)**
2003 Dummy	0.957	0.911	0.950
	(0.338)**	(0.314)***	(0.338)**
Constant	7.936	7.823	7.878
	(0.356)***	(0.330)***	(0.318)***
Regional Dummies	Yes	Yes	Yes
Observations	13033	13033	13033
R squared	0.37	0.37	0.37

***significant at 1% level, **significant at 5% level, *significant at 10% level

Recent Publications in the Series

nº	Year	Author(s)	Title
144	2011	Gabriel Mougani	An Analysis of the Impact of Financial Integration on Economic Activity and Macroeconomic Volatility in Africa within the Financial Globalization Context
143	2011	Thouraya Triki and Olfa Maalaoui Chun	Does Good Governance Create Value for International Acquirers in Africa: evidence from US acquisitions
142	2011	Thouraya Triki and Issa Faye	Africa's Quest for Development: Can Sovereign Wealth Funds help?
141	2011	Guy Blaise Nkamleu, Ignacio Tourino and James Edwin	Always Late: Measures and Determinants of Disbursement Delays at the African Development Bank
140	2011	Adeleke Salami, Marco Stampini, Abdul Kamara, Caroline Sullivan and Regassa Namara	Development Aid and Access to Water and Sanitation in sub-Saharan Africa
139	2011	Giovanni Caggiano and Pietro Calice	The Macroeconomic Impact Of Higher Capital Ratios On African Economies
138	2011	Cédric Achille Mbeng Mezui	Politique Economique Et Facteurs Institutionnels Dans Le Développement Des Marchés Obligataires Domestiques De La Zone CFA
137	2011	Thierry Kangoye	Does Aid Unpredictability Weaken Governance? New Evidence from Developing Countries
136	2011	John C. Anyanwu	Determinants Of Foreign Direct Investment Inflows To Africa, 1980-2007
135	2011	John C. Anyanwu	International Remittances and Income Inequality in Africa

African Development Bank

Angle de l'avenue du Ghana et des rues Pierre
de Coubertin et Hédi Nouira
BP 323 –1002 Tunis Belvédère (Tunisia)
Tel.: +216 71 333 511 – Fax: +216 71 351 933
E-mail: afdb@afdb.org – Internet: www.afdb.org