

# Does Education Pay In South Africa? Estimating Returns To Education Using Two Stage Least Squares Approach

Mduduzi Biyase, University of Johannesburg, South Africa

Talent Zwane, University of Johannesburg, South Africa

## ABSTRACT

*This paper investigates, using the first three waves of the National Income Dynamic dataset, the link between education and wages. Specifically it estimates the potential impact of the educational levels on wages in South Africa over the period 2008 – 2012. A two-stage least squares (2SLS) method is applied to account for endogeneity bias. More specifically, we use a lagged education as an instrumental variable in a two-stage least squares framework. Our results show that the proposed instruments is relevant and that there is an unambiguously positive effect on the wages of an individual from participation in education.*

**Keywords:** Endogeneity Bias; Earnings; Mincerian Model; Two-Stage Least Squares; Fixed Effects

## INTRODUCTION

The relationship between education and earnings have received a great deal of attention from both policy makers and economists in South Africa. It is believed that education can help in the way of alleviating poverty and unemployment, which is astonishingly high by international standards. According to Statistics South Africa (2014) the poverty rate in South Africa stands at 45.5%, while unemployment rate is in the region of about 25% (Finn and Leibbrandt, 2013). There has been a growing thread of literature on returns to education in South Africa. A number of studies have focused on racial and gender discrimination (Mwabu and Schultz, 1996; Moll, 1998; Ntuli, 2007; Burger and Jafta, 2006; Burger, 2011; Burger and van der Berg, 2011; Lam et al., 2011). While others have focused on returns to matric<sup>1</sup> and postsecondary education (Keswell and Poswell, 2004; Branson et al., 2013).

Many of the above studies (Mwabu and Schultz, 1996; Keswell and Poswell, 2004; Ntuli, 2007; Burger and Jafta, 2006) have relied on cross-sectional data rather than panel at a national level. The only study that has used panel only at a provincial level is by Lam et al., 2011. The lack of long-running national panel data in South Africa makes it difficult to track individuals over their lifetimes (Branson et al., 2013).

To fill a gap in the literature, this study employs the National Income Dynamics Study (NIDS) data set. The panel nature of the NIDS dataset makes it possible to go beyond a static approach to the question of returns to education in South Africa. The panel nature of the data also allows us to overcome common estimation issue of reverse causality.

The paper proceeds as follows. First, we review the literature on returns to education. Next we discuss the methodological issues that need to be considered when studying returns to education and describe the dataset used in this paper. The second last section describes the empirical approach and outlines the empirical results. The last section concludes with a discussion of the results.

---

<sup>1</sup> In South Africa, matriculation (or matric) is a term used to refer to the final year of high school and the qualification received on graduating from high school

## **Literature Review**

The causal nexus between the level of education attainment and wage earnings is well established in the empirical literature. The general understanding is that a subsequent increase in years of schooling would exert a simultaneous rise in the wage rate (Mocan, 2014).

There is a growing body of literature that investigates the link between education and earnings in country specific settings as well as cross-country settings. The former studies include Keswell and Poswell (2004); Aromolan (2006, 2004); Okuwa (2004); Kimenyi et al. (2006); Nielsen and Westergard-Nielsen (2001); Schultz (1999); Psacharopolous and Patrinos (2004); Bennell (1996); Angrist and Krueger (1991); Ashenfelter and Krueger (1994); Plug (2001); Aakvik et al.(2003); Mocan (2014); Oreopoulos (2006) and Garcia-Mainar et al (2005). While the latter studies include Trostel et al (2001), Salehi-Isfahani et al. (2009); Bennell (1996); Psacharopolous (1994); Harmon et al. (2001); De la Fuente (2003) and others.

Although a wealth of information is contained particularly in cross-country studies, these studies are likely to suffer from the well-known drawbacks. First, some of the cross country studies tend to present aggregate results without fully allowing for country specific effects that may affect the impact of education. Secondly, while there is a strong possibility that at least some of the explanatory variables could be endogenous, few studies attempt to control for this.

Typically, empirical work in this field has focused on establishing relationships between human capital indicators, such as education and training - on the one hand – and outcome measures such as wages. A variety of earnings measures have are used in the literature - annual, weekly, hourly - almost always in logarithmic form. The popularity of the log transformation reflects several factors. For one, the distribution of log earnings (especially log hourly wages) is surprisingly close to a normal distribution.

There is overwhelming evidence indicating that education does result in an increase in wages in many countries, using both individual country and cross-country analysis. In an individual country setting, Angrist and Krueger (1991), used quarter year of birth (i.e. Q2 1998) as an instrument for education. Their results revealed that on average, one additional year of education increased wages of individuals by 7.5 percent. Using the change in the legal school-leaving age in the United Kingdom, Oreopoulos (2006) compared his estimates of the returns to education to the local average treatment effect obtained from United States of America and Canada. The author found larger returns of 10-14 percent per year. The author argued that his results were similar to the finding obtained by previous scholars.

In her paper, Duflo (2000) investigated a reform that built 61,000 primary schools across Indonesia during the 1970s. Her results suggest that there is an increase of between 6.8 to 10.6 percent in wage per year of schooling. Likewise, a compulsory education reform in China was considered by Fang et al. (2012). The authors used reform as a potential instrument and reported that overall returns to education were approximately 20 percent of an additional year of schooling in contemporary China. The authors concluded that their results were almost consistent with the finding obtained in industrialised nations.

Using twins as an instrumental variable, Ashenfelter and Rouse (1998) found that an additional year of education presented a 9 percent increase in wages for these twins. In a similar study, Bonjour et al. (2003) studied a set of identical twins from the United Kingdom. Their results revealed that the returns to education for women was 7.7 percent per year of schooling. However, these studies were not very clear as to why one twin achieved more education than the other, while coming from the same genetics and upbringing.

Cross-country studies confirm the positive relationship between wages and education. For example, Salehi-Isfahani et al. (2009) examined the private returns to schooling in Egypt, Iran, and Turkey and found that returns to education increase in years of schooling in all three countries. More specifically, Private returns to tertiary education compared to upper secondary and vocational education were found to be high, confirming that university education is highly attractive in these countries. Their results were in line with other studies that found low returns to vocational training compared to general upper secondary.

In their Kuepie et al. (2006) investigated the effects of education on urban labour market participation and earnings and found that although education does not always reduce unemployment, it does increase individual earnings in some cities and open the door to get into the most profitable niches, which are found in the formal private and public sectors.

As noted in the introduction, there is a growing body of literature that investigates returns to education in South Africa. For example, Branson et al. (2013) investigated the changes in the employment and earnings in South Africa. More specifically their study focused on the changes in the distribution of education across birth cohorts in relation to changes in employment probability as well as distribution of earnings and earnings premium to complete secondary and tertiary education. The results revealed that young cohort faced worse labour market conditions relative to their predecessors.

A study by Burger and Van der Berg (2011) investigated the wage differentials between various race groups in South Africa, using a simulation model that employed matric results and education attainment to generate estimates of education quality. Their results suggest that wage gap can be attributed to differences in education quality.

### **Methodological Issues**

Investigating the relationship between education and earnings is not straightforward because several socioeconomic factors are simultaneously correlated with education and earnings. If these factors are not accounted for in the analysis, we cannot conclude that the estimated association between education and earnings is causal. There is, therefore a need to address the problems cited above, i.e. the “endogeneity” to be able to say something about the causal effect of education on earnings.

Some authors have estimated the link between education and earnings using the OLS method. However, Ordinary Least-Squares (OLS) estimation of the effect of education on earnings are likely to be biased. The reason for biased OLS estimates is twofold: It ignores the fact that people with more education tend to have higher earnings due to their greater ability or better family background. If there is some correlation between education and the error, then an OLS regression will produce a biased estimate of the return to years of education. Secondly it is possible that years of schooling and years of experience are highly correlated. To truly be able to make a causal claim, we use of an instrumental variable that is correlated with the endogenous variable (schooling) but uncorrelated with unobserved ability.

Many researchers (see Card, 2001; Heckman et al, 2006; Kerr and Quinn, 2010; Leyaro et al, 2010; Rankin Sandefur and Teal, 2010) have used the Two-Stage Least Squares (2SLS) techniques to overcome the bias associated with the OLS method.

In other words, one needs to find a set of relevant instruments which is correlated with the schooling variable but uncorrelated with the wage or error term. Finding a relevant instrument which is correlated with the schooling variable but uncorrelated with the wage or error term is extremely challenging. Perhaps unsurprisingly, some researchers end up using invalid instruments. In their study, Carneiro and Heckman (2002) observed that most empirical studies on return to education in the literature used invalid instruments that tend to have a high correlation with the omitted personal unobserved abilities. Given the challenges of finding a valid instrument, our study follows the work of De Gregorio et al. (2002) who used the lagged value of education as an instrument in their study.

### **Data**

To investigate the impact of education on earnings we exploit the available panel data in the first three waves of the National Income Dynamics (NIDS) data set. The NIDS is a longitudinal survey of private South African households designed as a biannual survey. The survey (which is conducted by the Southern Africa Labour and Development Research Unit (SALDRU) began in 2008 with a large nationally representative sample of over 28,000 individuals in 7,300 households across the country. After that those who were present in the initial sample as well as their spouses and children, are re-interviewed. The survey continues to be repeated with the same household

members every two years. The NIDS collects demographic information, including age, sex, marital status, living situation, a location of residence, a highest level of completed schooling and so on. Beyond these demographic questions, the survey primarily asks questions about behaviors and outcomes related to work.

Details on the variables used in the in this study are provided in Table 1 below. They are based on individual characteristics, household, and regional level characteristics. The literature on the micro level regarding education attainment was used as a guide in selecting the relevant variables (See for example, Bonjour et al, 2003; Angrist and Krueger, 1999). The following table describes the variables used in the study.

**Table 1.** Variable Description

<b>Explanatory variables</b>	<b>Description</b>
HH_age	age of household head (in years).
HH_married	marital status of the household head (married = 1 and not mar = 0)
HH_Female	gender of the household head (male = 1 and female = 0)
Coloured	Race of the household head (Coloured = 1 and 0 otherwise)
Indian	Race of the household head (Indian = 1 and 0 otherwise)
White	Race of the household head (White = 1 and 0 otherwise)
Farms	household in farms (1/0)
Urban	household in Urban areas (1/0)
Education	education of the household head (years of schooling).
Primary	Individual has primary education (1/0)
Secondary	Individual has secondary education (1/0)
Matric	Individual has matric education (1/0)
Tertiary	Individual has tertiary education (1/0)
Wages	Wages
Western Cape	individual resides in the Western Cape (1/0)
Eastern Cape	individual resides in the Eastern Cape (1/0)
Northern Cape	individual resides in the Northern Cape (1/0)
Free State	individual resides in the Free State (1/0)
KwaZulu-Natal	individual resides in KwaZulu-Natal (1/0)
North West Province	individual resides in the North West Province (1/0)
Gauteng	individual resides in Gauteng (1/0)
Mpumalanga	individual resides in Mpumalanga (1/0)

**Empirical Strategy**

Our task is to examine the relationship between education and earnings in South Africa. Rather like many studies in this field we use the earnings function proposed by Mincer (1974). The Mincerian wage regression has been mostly used in the literature on the returns to education (Card, 2001, and Belzil and Hansen, 2002a).

As already noted, using Pooled OLS in the presence of endogeneity and heterogeneity problems will result in biased and inconsistent estimates. The estimates might be biased due to the presence of an unobservable factor that is correlated with both schooling and wages. For example, well-educated people may have high earnings as a result of their greater ability or better family background. This study uses panel data, which permits us to control for unobservable individual characteristics. More specifically, heterogeneity problem was accounted for by using the Fixed effect that control for time-invariant unobserved characteristics. While Potential endogeneity between wages and educational attainment was tackled by pursuing the Two-Stage Least Squares approach. Thus, in this study we use two different models to empirically investigate the relationship between education and earnings in South Africa. First, we estimate the baseline fixed effects model to control the heterogeneity expressed as follows:

$$\ln W_{it} = \alpha_0 + \alpha_1 \text{education}_{it} + \alpha_2 \ln X_{it} + \delta_i + \pi_{it} \tag{1}$$

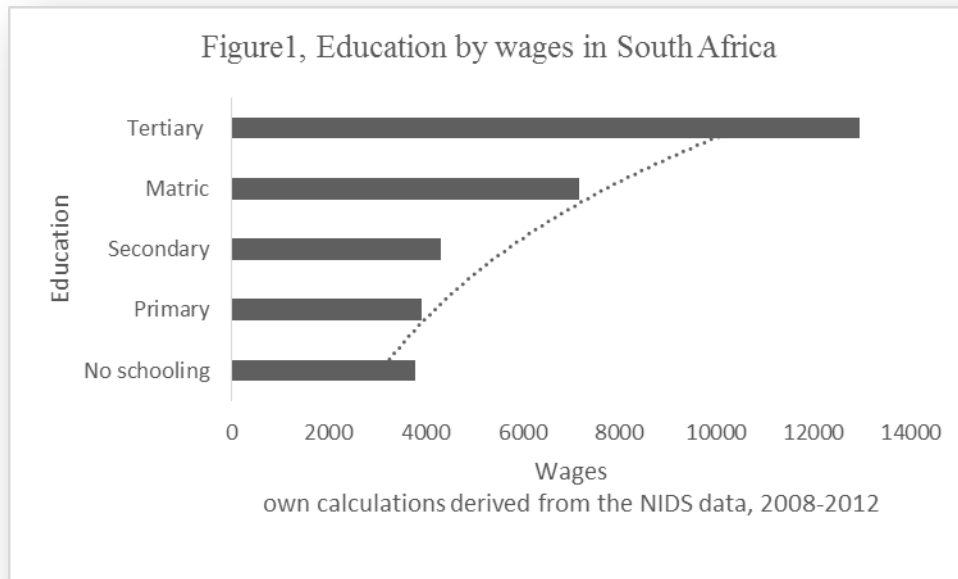
Second, the Two-stage Least Squares regression specification to control for endogeneity is as follows:

$$\ln W_{it} = \alpha_0 + \beta_1 \text{education}_{t-1} + \alpha_2 \ln X_{it} + \pi_{it} \tag{2}$$

Where  $LnW_{it}$  represents earnings of individual  $i$  at time  $t$ . The error term  $\pi_{it}$  includes both preference shocks and measurement error and are distributed identically and independently.  $X_{it}$  are individual characteristics including the age of the household head, education of the household head family size, etc. The model (1) also includes  $\delta_i$  individual specific time- invariant unobservables to control the heterogeneity and account for endogeneity (as noted in the previous section) by using the lagged value of education as an instrument, consistent with the work of De Gregorio et al. (2002).

**RESULTS**

Before we proceed to presenting empirical results, we provide a descriptive analysis showing that higher levels of educated are associated with more earnings.



Analysis of the NIDS data (figure 1) shows that educational attainment is strongly linked to average earnings, particularly for those with tertiary education. Specifically, an average person who had never attended school earned around R 3778 monthly while graduates of tertiary education earned R 12951. Employed adults who finished secondary school earned an average of nearly R 4311 while those who finished matric earned almost R 7175. In a nutshell, the NIDS data provide reassurance that spending time and money furthering ones studies is likely to pay off particularly in monetary value.

We report the empirical results in Table 2 for both the fixed effect and the Two-Stage Least Squares models. Column 2 of Table 2 contains the baseline regression results using equations (1), while column 4 of Table 2 accounts for the potential endogeneity of education and estimate Two-Stage Least Squares regression using equation (2). The baseline results are consistent with the human capital literature and previous estimates showing that an increase in household head’s educational level is directly associated with an increase in wages. More specifically, estimate suggests that an additional year of schooling increases the wage by 2.19%. The results for the control variables are as expected. For example, an age of the household head, gender of the household head, whether the head of household is married are positive and statistically significantly different from zero.

Column 4 of Table 2 reports the Two-Stage Least Squares results that corrects for potential endogeneity of the education variable. There are at least two crucial features govern the selection of an instrument for education: the instrument must be correlated with education and its effect on earnings must operate solely through its effect on

education and should not be directly affected by earnings. The instrument that is used here is the lagged value of education, which has been used in the previous studies.

There are some noticeable differences between the baseline estimates and the Two Stage Least Squire estimates: the latter estimate suggests the return is about 47%. The dramatic difference in the estimated coefficients is not surprising because the years of schooling is an endogenously related to wages. 2SLS estimator is likely to be able to provide less prone to mis-specification than FE estimator (Belzil, 2007; Keane, 2010). The control variable coefficients are quite similar to the baseline, the age of the household head, gender of the household head, whether the head of household is married still presents positive and significant estimates on wages. To summarize, the results across all models indicate a statistically significant impact of education on wages.

**Table 2.** Fixed effects and 2SLS estimates of education effects of on earnings

	Fixed Effect	Standard Errors	Two Stage Least Square	Standard Errors
Education yrs	.02195254***	0.0016477	.47282461***	0.0136854
Age	.03495984***	0.0041419	.0630058***	0.0048038
Hhsize	-.0082395***	0.0008158	.00962604***	0.0010537
Married	.47203634***	0.0243221	.33518136***	0.0286962
Gender Female	-.10726456***	0.0231173	-.18736302***	0.0269859
Coloured	.21746865***	0.039144	.189339***	0.0476376
Indians	1.0240231***	0.0928617	.5514029***	0.1133535
Whites	1.278171***	0.0486402	.69607918***	0.0680134
Urban	.4428578***	0.0311204	-.23493865***	0.0467089
Farms	0.0079166	0.0385741	0.06188911	0.0412988
Eastern Cape	-.13338767*	0.0587933	-0.07572971	0.0701118
Northern Cape	-.31599121***	0.0560299	-.31790759***	0.0651498
Free State	-.12831683*	0.0611709	-0.02717157	0.0722862
KwaZulu-Natal	-.21890974***	0.0618937	-.20898968**	0.0721912
North West	-.17533047***	0.0486082	-0.02431921	0.0561187
Gauteng	.16249545**	0.0596197	.17049937*	0.0687215
Mpumalanga	.12826965*	0.054104	.12924527*	0.0636305
Limpopo	.17379438**	0.0570057	.28861502***	0.067158

**Notes:** Standard Errors in parentheses. \* Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

## CONCLUSION

In this paper, we examined the labour market returns to education in South Africa, using the first three waves of the National Income Dynamic dataset. Two estimation methods were used in an attempt to overcome limitations faced in conventional earnings function analyses: fixed effect and two stage least squares. Accounting for endogeneity problem, we find that an additional year of schooling increases labour market earnings by 47%. The results for the control variables were as expected. For example, an age of the household head, gender of the household head, whether the head of household is married were positive and statistically significantly different from zero. Thus our results are broadly in line with many studies in this field, which suggest that private investment in higher education is a worthwhile investment.

## REFERENCES

- Aakvik A, Salvanes K. Vaage K (2003). Measuring heterogeneity in the returns to education in Norway using educational reforms. IZA Discussion Paper 815, Bonn
- Angrist, J. D. Krueger, A.B. (1991). Does compulsory school attendance affect schooling and earnings? The Quarterly Journal of Economics, 106(4), pp. 979-1014
- Aromolaran, A. B. (2004). Wage returns to schooling in Nigeria. African Development Review, 16(3): 433-455.
- Aromolaran, A. B. (2006). Estimates of mincerian returns to schooling in Nigeria. Oxford Development Studies, Vol. 34(2): 265-292
- Ashenfelter, O. Rouse, C. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. The Quarterly Journal of Economics, 113(1), pp. 253-284



- Ashenfelter, O. Krueger, A. (1994) Estimates of the economic return to schooling for a New Sample of Twins. *American Economic Review*, 84(5), pp. 1157-1173
- Belzil, C. (2007). The return to schooling in structural dynamic models: A Survey. *European Economic Review*, 51(5), pp. 1059–1105
- Belzil, C. Hansen, J. (2002). Unobserved ability and the return to schooling, *Econometrica*, 70(5), pp.2075-2091
- Bennel, P. (1996) Rates of return to education: Does the conventional pattern prevail in Sub-Saharan Africa? *World Development*, 24 (1), pp. 183-99
- Bonjour, D. Lynn, F. Cherkas, J. E. Haskel, D. Hawkes. D. Spector, T. (2003). Returns to Education: Evidence from UK Twins. *American Economic Review*, 93(5), pp. 1799-1812
- Branson, N. Ardington, C. Lam, D. Leibbrandt, M. (2013). Changes in education, employment and earnings in South Africa – A cohort analysis. SALDRU Working Paper 105, University of Cape Town
- Burger, R. (2011). Estimating the shape of the South African schooling-earnings profile, (Doctoral dissertation). Oriol College, Oxford University
- Burger, R Jafta, R. (2006). Returns to Race: Labour market discrimination in Post-Apartheid South Africa, University of Stellenbosch Working Paper
- Burger, C. Van de Berg, S. (2011). Modelling cognitive skills, ability and school quality to explain labour market earnings differentials, Stellenbosch Economic Working Papers: 08/11
- Card, D. (1999). The causal effect of education on earnings. *Handbook of Labour Economics*, Edited by Ashenfelter, O. and Card D. Volume 3 (Part 1): 1801-1863
- Carneiro, P. Heckman, J. (2002). The evidence on credit constraints in postsecondary schooling. *Economic Journal* 112(482), pp. 705-734.
- De Gregorio, J. Lee, W. (2002). Education and income inequality: new evidence from cross-country data. *Review of Income and Wealth* 48(3), pp. 395-416
- De la Fuente, A. (2003). Human capital in a global and knowledge-based economy. Assessment at the EU Country Level. Report for the European Commission. DG for Employment and Social Affairs
- Duflo, E. (2000). Schooling and labour market consequences of school construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91(4), pp. 795-813
- Fang, H. Eggleston, K.N. Rizzo, J. Rozelle, S. Zeckhauser, J. (2012). The Returns to education in China: Evidence from the 1986 Compulsory Education Law. NBER Working Paper 18189
- Finn, A. Leibbrandt, M. (2013). The dynamics of poverty in the first three waves of NIDS. Cape Town: SALDRU, University of Cape Town. SALDRU Working Paper No. 119/NIDS Discussion Paper 2013/1
- García-Mainar, I. Monteungu-Gómez, V. M. (2005). Education returns of wage earners and self-employed workers: Portugal vs. Spain. *Economics of Education Review*, 24(2), pp. 161–170
- Harmon C. Oosterbeek, H. Walker, I. (2003) ‘The Returns to education – A Review of evidence, issues and deficiencies in the literature’ *Journal of Economic Surveys*, 17(2), pp. 115-156
- Heckman J. Stirxud, J. Urzua, F. (2006). ‘The effect of cognitive and non-cognitive factors in behavioural and labour outcomes’, *Journal of Labour Economics*, 24(3), pp. 411-482
- Harmon, C. Walker, I. Westergaard-Nielsen, N. (2001), Education and earnings in Europe. A Cross Country Analysis of the Returns to Education, Cheltenham, UK and Northampton, MA, USA: Edward Elgar
- Keane, M. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics*, 156(1), pp. 3-20
- Kerr, A. S. Quinn, S. (2010). Returns to education in Tanzania: Exploiting a natural experiment. Paper presented in the Centre for the Study of African Economies Conference on Economic Development in Africa, 21-23 March 2010, St Catherine's College, Oxford, UK.
- Kimenyi, M. Mwabu, G. Manda, K. (2006). Human capital externalities and private returns to education in Kenya, *Eastern Economic Journal*, Vol. 32(3), pp. 493- 513
- Keswell, M. Poswell, L. (2004). Returns to education in South Africa: a retrospective sensitivity of the available evidence. *South African Journal of Economics*. 72(4), pp. 834-860
- Lam, D. Ardington, C. Leibbrandt, M. (2011). Schooling as a lottery: Racial differences in school advancement in Urban South Africa, *Journal of Development Economics*, 95(2): 121-136
- Martins P. Pereira, T. (2004). ‘Does education reduce wage inequality? Quantile regression evidence from 16 countries’, *Labour Economics*, 11 pp.355-371
- Mincer, J. (1974). Schooling, experience and earnings. National Bureau of Economics, New York: Columbia University Press

- Mocan, L. (2014). The impact of education on wages: Analysis of an Education Reform in Turkey. Wharton School of Business. Working Paper109, April 2014
- Moll, G. (1998). Primary schooling, cognitive skills, and wage in South Africa. *Economica*, 65, pp. 263-284
- Mwabu G. Schultz, T. (1996). Education returns across quantiles of the wage functions: alternative explanations for returns to education by race in South Africa. *The American Economic Review*, 86(2), pp. 335-339
- Ntuli, M. (2007). Exploring gender wage discrimination in South Africa, 1995- 2004: A Quantile Regression Approach. IPC Working Paper Series: 56
- Okuwa, O. B (2004). Private returns to higher education in Nigeria. AERC Research paper No. 139, Nairobi Kenya
- Oreopoulos, P. (2006). Estimating average and local treatment effects of education when compulsory schooling laws really matter. *The American Economic Review*. 96(1), pp. 152-175
- Plug, E. (2001). ‘Season of birth, schooling and earnings’, *Journal of Economic Psychology*, 22, pp. 641-660
- Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development* 22(9), pp. 1325-1343
- Psacharopoulos, G. Patrinos, H. (2004). Returns to investment in education: A further update. *Education Economics* 12(2), pp. 111-134
- Rankin, N. Sandefur, J. Teal, F. (2010). ‘Learning and earnings in Africa: where are the returns to education high? Working Paper. Oxford: Centre for the study of African Economies
- Salehi-Isfahani, D. Tunali, I. Assaad, R. (2009). A comparative study of returns to education of urban men in Egypt, Iran, and Turkey. *Middle East Development Journal*. 1(2), pp. 145- 187
- Schultz, T. (1999): Health and schooling investments in Africa. *Journal of Economic Perspectives*, 13(3), pp. 67-88
- Statistics South Africa. (2014). Poverty trends in South Africa: an examination of absolute poverty between 2006 and 2011. Statistics South Africa. Report No. 03-10-06
- Trostel, P. Walker, I. Woolley, P. (2001). Estimates of the economic return to schooling for 28 countries. *Labour Economics*, 9(1), pp. 1-16