Poverty, Inequality and Inclusive Growth Dynamics: Evidence from Nigeria's Panel Household Surveys

Henry C. Edeh
and
Jane O. Ozor

Working Paper GPIR-CC-008

Bringing Rigour and Evidence to Economic Policy-making in Africa

AFRICAN ECONOMIC RESEARCH CONSORTIUM
CONSORTIUM POUR LA RECHERCHE ÉCONOMIQUE EN AFRIQUE

Poverty, Inequality and Inclusive Growth Dynamics: Evidence from Nigeria's Panel Household Surveys

By

Henry C. Edeh
Department of Economics, University of Nigeria

and

Jane O. Ozor
Department of Economics, University of Nigeria

AERC Working Paper GPIR-CC-008
African Economic Research Consortium, Nairobi
August 2024

THIS RESEARCH STUDY was supported by a grant from the African Economic Research Consortium. The findings, opinions and recommendations are, however, those of the author and do not necessarily reflect the views of the Consortium, its individual members or the AERC Secretariat.

Published by: The African Economic Research Consortium
P.O. Box 62882 – City Square
Nairobi 00200, Kenya

Contents

List of tables

Abstract

Using updated Nigeria's micro panel household surveys, we estimate the impact of human capital endowments on household economic well-being – controlling for exogenous circumstance-related factors over which households have little or no control. We found that education and health endowments have significant causal impact on the well-being of the households. More so, education has significant causal impact on the income of households below the bottom 40% (specifically the households at the bottom 25%). Inequalities at the national level are mainly determined by rural-urban and northern-southern inequalities. These observed income inequalities in rural-urban sectors and northern-southern geopolitical zones are mainly propelled by differences in education and health endowments of the households. However, the equalization of human capital endowments in terms of education and health is indeed growth-enhancing. We suggest, therefore, that policies capable of eliminating inequalities in access to schooling and health will enable households at the bottom of the distribution to enjoy better economic well-being.

Keywords: *Poverty, Inequality, Pro-poor growth, Micro panel household data, Nigeria.*

1. Introduction

Despite slipping into recession in 2016, Nigeria is still recognised as Africa's largest economy. The economy is well noted as number one in Africa not just because it is the sixth most populous country in the world (World Bank's World Development Indicators (WDI), 2021). It is the giant of Africa not only on account of its oil wealth. It is significant more because of the energy of its people, whose creativity, resilience and spirit of enterprise continue to assure the country of its progress even in the face of seemingly hopeless situations. It is due to the toil and hard work of ordinary Nigerians (the nation's greatest asset) that the country attained a Gross Domestic Product (GDP) rebased at $510 billion in 2014, exceeding that of South Africa to become the biggest African economy. However, roughly half of these hard-working Nigerians still tussle with poverty (Oxfam International, 2017). Economic growth in Nigeria is yet to be translated into significantly reduced poverty levels, countering the United Nation's (UN) Sustainable Development Goal 1 (SDG1) of ending poverty by 2030 (United Nations, 2017).

Poverty in the country has been growing in the context of an expanding economy where the benefits have been reaped by a small proportion of the population (Oxfam International, 2017). Despite strong average economic growth, 8.5% in the 2000s, the poverty reduction process remained slow (World Bank, 2016). Although the 2016 oil shock disruptions made a dent on growth, the country's average rate of growth in the recent decade remains at a modest level of 4.0%. The modest average growth notwithstanding, Nigeria remains one amongst African countries that have experienced a rising trend in the level of poverty. The poverty rate grew from 26% in 1980 to 46% in 2004 (Nigeria's National Bureau of Statistics (NBS) 2012; World Bank, 2013). Although a decline to 35% was observed in 2010 (World Bank, 2013), the poverty rate has currently risen from the 35% seen in 2010 to 41% in 2019 (World Bank, 2013; NBS, 2020).

Although income inequality as measured by the Gini index in Nigeria marginally declined to roughly 40.0 in 2016, from an average value of 41.0 in 2004-2013, the latest Commitment to Reducing Inequality (CRI) report by Development Finance International (DFI) and Oxfam International records that the country remains at the bottom of the CRI index, with a shamefully low social spending, increase in labour rights violations, and low tax collection (NBS, 2018; DFI and Oxfam, 2020). More perturbing is the spatial dimension of poverty and income inequality. Compared to the rural areas, urban areas in Nigeria experience a significantly lower poverty rate. While poverty rates in rural areas were 48.5 and 52.1% in 2013 and 2019, the rates for the urban areas were only 15.9 and 18.0% in the two periods, respectively. Poverty rates in the north are higher as compared to poverty rates in the south. On average, poverty rates in northern regions in 2013 and 2019 were 48.5 and 60.0%, while the rates in the southern regions were only 22.8 and 25.3% respectively (World Bank, 2016; NBS, 2020). Closing spatial well-being disparities has become a priority within the Nigeria government's development policies, and is at the centre of public debates. Empirical evidence on household human capital endowment drivers of spatial economic well-being is indeed necessary to aid actualization of development policies.

The relationship between human capital, poverty, inequality and growth has been the subject of increasing literature in both developed and developing countries (Olopade et al., 2019). Several studies have shown that human capital variables (education and health) play an important role in the process of achieving growth, reduced poverty and reduced inequality (Olopade et al., 2019; Shahpari and Davoudi, 2014). This happens through human capital formation, in terms of increased levels of education and health (Olopade et al., 2019; Bahar, 2022). Earlier works (Becker, 1975; Marin and Psacharopoulos, 1976; Schultz, 1961; Fields, 1980) suggest that human capital formation leads to improvement in economic benefits such as equality in the distribution of income and enhanced productivity. In line with this suggestion, the study of Shahpari and Davoudi (2014) found that increasing human capital (mean level of labour schooling) can make income distribution fairer i.e., leads to a better distribution of income in Iran. In a cross-country analysis, Bahar (2022) has revealed a significant and positive effect of human capital (gross enrolment rate in tertiary education) on income equality – suggesting that countries with a high stock of human capital experience a higher level of income equality. In Bangladesh, Mamun and Arfanuzzaman (2020) found that education has significant explanatory power on household income. Lee and Lee (2018) have shown that a more equal distribution of education contributes significantly to reducing income inequality across a broad range of countries. Su and Guo (2022) found in China that education and health have a negative effect on vulnerability to relative poverty. In the Eastern Cape province of South Africa, Moyo et al (2022) found that an increase in human capital (attainment of education) leads to a decline in poverty levels. However, human capital was found to be positively related to income inequality – an indication of unequal economic opportunities and inequality in the education system. In Cameroon, Baye and Epo (2015) found that education and health significantly improve economic well-being and reduce inequality.

Several studies on the impact of human capital (education and health) on growth (Epo and Baye, 2012; Baye and Epo, 2015; Ogundari and Awokuse, 2018), poverty (Moyo et al., 2022) and inequality (Baye and Epo, 2015; Moyo et al., 2022) exist for some African countries. However, a comprehensive micro-level panel empirical study on these issues in Nigeria has barely been done. The current study examined the impact of human capital (education and health) on poverty, inequality and growth in Nigeria. Further, less attention has been paid to these issues across regions and geopolitical zones of the country.

Previous studies for Nigeria published in rated journals have made some contributions in this area in several ways. For instance, the work of Omoruyi, and Omoyibo, (2011) focused on effect of household income sources on inequality and uncovered that household farm income source was by far the most significant contributor to household income inequality. The work of Adetayo, (2014) indicates that the likelihood of being poor were more with large households, non-educated farm household and households without access to credit and other non-farm income. The contribution of the work of (Ataguba et al., 2013) is the use of capability approach in understanding the missing dimensions of poverty in Nigeria. Olowa, et al. (2013) focused on remittances effect and uncovered that both domestic and foreign remittances reduced the level, depth and severity of poverty. In contributing to the measurement of levels of poverty and inequality across zones and regions, Jaiyeola and Bayat, (2019) found an overall deepening of poverty between 2010 and 2013. The main limitations of these studies in terms of our research questions lies in examination of impact of human capital on welfare, regional poverty, national inequality, zonal and regional inequality and pro-poor growth. Hence, the current study attempts to fill these identified gaps.

More so, the studies of (Ohwotemu, 2010; Usman et al., 2016) have reported income/expenditure inequality between the urban and rural areas. The work of (Usman et al., 2016) for instance, has reported high rural income inequality, and differences in urban-rural household income. Urban regions mean income contributions in the total household average income are roughly double the mean income shares of the rural region across 2010-2019 period. While the urban areas' contributions to average total income were largely - 66 and 64%, the rural regions' contributions were only 34 and 36% in 2010-2013 and 2015-2019 respectively (NBS Survey Reports, 2010-2019). It is, thus, essential to grasp fully, the sources and drivers of regional income inequality to structure and implement effective and socially acceptable development policies in the country.

Poverty and income inequality issues in the country have always been among the policy choruses sung by different administration since 1980s to-date followed by various policy strategies and programmes with little or no effective implementation to accomplish meaningful effect. Various policies to bring down poverty levels and income gap in Nigeria have over the years, ranged from those of the pre-SAP era (River Basin Development Authority, Agricultural Development Programme) to those of the SAP era (the Directorate for Food, Roads and Rural Infrastructures, National Directorate of Employment) and the Democratic era (National Economic Empowerment and Development Strategy, National Poverty Eradication Programme), among others (Oxfam, 2017). These policies notwithstanding, poverty and inequality remain worrisome issues in the country.

Previous empirical studies (Aigbokan, 2008; Ofem et al., 2010; Sowunmi et al., 2012 among others) on poverty in Nigeria have generally focused on incidence of poverty with methods, ranging from Factor analysis technique, Foster-Greer-Thorbecke, to Descriptive statistics, and Ordinary least square (OLS). It is crucial to note however, that these studies have not used a panel data, and have mainly employed static methods. Results from these previous studies may have fallen short of causal interpretation, since they failed to deal with the issues of endogeneity, unobserved heterogeneity as well as sample selection biases. Further, some of these studies have only used a very small sample data that was clearly unable to represent Nigeria – a country with the largest population in Africa.

Other previous related studies (Ijaiya et al., 2011; Olofin et al., 2015; Ogbeide and Agu, 2015) have only used aggregate macro-level time series datasets, which were clearly unable to rigorously speak to the issues of poverty and inequality at the micro household level. In Nigeria, studies are scarce that have examined the relative impact of specific household human capital endowment factors on income quintiles (income of the bottom 40%) and the drivers of poverty and income gap between the urban and rural areas, and the north and south regions, using regression-based linear and non-linear decomposition methods and nationally representative panel data rounds spanning a decade. In terms of income inequality, existing studies (Olanrewaju, and Timothy, 2005; Olaniyan and Awoyemi, 2005) have mainly used static and dated approaches. However, studies in Nigeria that have established the human capital endowment sources and drivers of regional income inequality using a long panel household dataset and a regression-based decomposition technique that allows for causal inference are indeed rare.

Further, the above-noted studies have made very limited attempts at addressing the issues of pro-poor growth, together with the corresponding household endowment drivers. Aggregate appraisal of pro-poor growth as in few previous attempts may conceal the diversity in patterns and sources of pro-poor growth. While growth is needed to achieve sustained poverty reduction, comprehending both the quality and sources of growth is indeed crucial and will enable the design of real growth and inclusiveness policies in the country. Unlike the previous few attempts, the current study employed the relative measure of pro-poor growth (Kakwani et al., 2010) that links growth rates in mean income and in income inequality to address growth inclusiveness and its contributing sources in Nigeria (Kakwani et al., 2010).

In the light of the above identified gaps, the current study intends to contribute to existing literature in the following useful ways. First, unlike previous studies in Nigeria, this study will use a rich micro panel data to account for endogeneity, unobserved heterogeneity and sample selection biases using the panel instrumental variable (IV) and sample selection approaches in examining the human capital factors that determine household well-being. Second, it will examine the drivers of income at different conditional quantiles using a relatively recent panel quantile regression (QR) method with an in-built instrumental variable technique that not only enabled us to deal with endogeneity and unobserved heterogeneity problems, but also identify key regressor(s) - education years that do not sufficiently vary over time. The advantage of using QR is that it allows for covariates to have marginal effects that vary with households' positions in the welfare distributions. The aim here is to identify the specific human capital endowments that drive income growth of the poor within the income distribution.

Third, this study aims to analyse the drivers of poverty across national and regional levels by employing the random-effect logistic regression model of poverty drivers.

Although the random-effect logistic model assumes no endogeneity problem in the model, it enabled the identification of the key covariates – education years that do not sufficiently vary across time. To further supplement our analyses of the drivers of poverty, the study examined the human capital correlates of urban-rural and northern-southern poverty differentials, for the first time in Nigeria using the Fairlie (2003), which is more appropriate for decomposing binary outcomes (poor/non-poor). The logit and probit models are non-linear and may be used to decompose the overall poverty gap. However, these models have no unique way of revealing the detailed decomposition. Hence the use of the Fairlie decomposition in this paper. Results from above analyses are to inform policies relating to management of identified variables, deepening the overall poverty and poverty differentials, and in redistribution of resources between the two regions and geopolitical zones.

Fourth, the study estimated a 'regression-based inequality decomposition' which is thought to be more revealing and crucial for targeting the root causes of inequality, relative to the 'traditional approach' involving the mere accounting and decomposition of inequality commonly used in previous studies. In employing the regression-based decomposition approach developed in Fiorio and Jenkins (2007), we used the panel instrumental variable approach to deal with the issues of potential endogeneity in endogenous variables – education and health.

Fifth, unlike previous papers, this study tried to decompose the urban-rural and northern-southern income differentials into household human capital endowment drivers. This was achieved using the recently developed panel Kitagawa-Oaxaca-Blinder decomposition technique (Kroger and Hartmann, 2021). To ensure that our decomposition results are to be interpreted in a causal manner, we invoked the instrumental variable approach to purge the key regressors of potential endogeneity and unobserved heterogeneity problems. This approach is in line with the panel Kitagawa-Oaxaca-Blinder approach of ensuring that the estimators of the regressors in the first-step panel linear regression are unbiased, before being utilized in the second-step decomposition analysis. Such analyses are crucial for policy purposes in terms of revealing the factors that cause inequality in income in the country.

Sixth, unlike previous works in Nigeria, the current study analysed inclusive growth and corresponding human capital drivers based on the relative pro-poor growth measure recently proposed by Kakwani et al (2010). Specifically, this was achieved in a counterfactual policy simulation framework involving the 'without and with elimination of inequality in specific sources of wellbeing', to see the impact of specific policy changes on pro-poor household income growth in the country. Finally, the study employed several rounds of the Nigeria General Household Surveys (NGHS) - Panel, ranging from 2010 to 2019, in achieving the above identified objectives. These panel datasets are current, rich and span over a decade.

2. Objectives

Objectives of the study

The broad aim of the study is to assess the impact of human capital on poverty, income inequality and pro-poor growth using available panel household surveys in Nigeria.

Hence the study intends to achieve specific objectives, namely:

1. To analyse the impact of human capital endowment on household well-being in Nigeria.

2. To examine the impact of human capital endowment on incomes of the bottom 40% of the population at the national, regional and geopolitical zone levels.

3. To analyse the impact of human capital on national, urban-rural and northern-southern poverty.

4. To examine the human capital drivers of total income inequality using a regression-based decomposition in Nigeria.

5. To ascertain the impact of human capital endowment on urban-rural and northern-southern gaps in inequality.

6. To appraise the impact of specific household human capital endowments on pro-poor income growth in Nigeria.

Other things being equal, the specific objectives of the study are governed by the perspective that human capital endowments are the drivers of poverty, income inequalities and pro-poor growth in Nigeria.

To address these specific objectives, the study will employ the methods detailed below:

3. Methodology

Panel framework of the household well-being-generating function

The framework is based on a panel household welfare-generating function, where household income is determined by both a set of exogenous variables and another set of endogenous variables. The exogenous variables are factors beyond the control of the household, whereas the endogenous variables are factors that the household can directly influence. In this study, the exogenous variables mainly include age, gender, location, and regions of households readily available in the survey data, whereas the main endogenous factors are the human capital and labour market-related variables such as education, health and employment respectively. Since these variables are chiefly endogenous, endogeneity problems could erupt from unobservable individual-specific characteristics such as ability, family background or genetic endowment of parents. Panel data techniques have the advantage of not only studying the dynamics of change, but explicitly accounting for unobservable individual characteristics. In this study, we account for endogeneity and unobserved heterogeneity biases using the instrumental variable (IV) technique in the panel data context (see procedure below). Further, to account for sample selection bias, the study utilized a panel two-step estimation method to generate a sample selection term (inverse mills ratio) from a first-step sample selection equation and then include this selection correction term as an additional regressor in the second-step panel well-being-generating model. We model the household well-being-generating function across both individual and time dimension below:

Analysing the impact of human capital on household well-being in Nigeria

In applying a panel approach to the above objective, omitted variables are modelled as individual fixed effects. This means that unobserved effects (e.g., ability) are time-invariant and individual-specific. This can be shown as:

$$lnIncome_{it} = X_{it}\beta + Z_i\delta + c_i + u_{it} \qquad t = 1, \dots, T; \; i = 1, \dots, N \tag{1}$$

Where the subscript *i* represents the cross-sectional unit and *t* represents the time period. $lnIncome_{it}$ refers to the log of well-being of individual *i* at time *t*. X_{it} is a vector for observed exogenous and endogenous characteristics that changes across individual *i* and time *t*. Z_i is a vector of exogenous (e.g., gender) that changes only across individual and endogenous variables, including education years that do not sufficiently change across time. An important assumption underlying the above equation is that c_i denotes unobserved individual heterogeneity, i.e., individual-variant variables that are hard to observe. u_{it} is an error term also measured across both individual and time, and for which the strict exogeneity condition is assumed to hold i.e., $E(u_{it}|X_{it}) = 0$. The c_i captures omitted variables such as family background or ability. Ability could be correlated with education, since people with relatively high intelligence will go for higher education years i.e., $E[c_i \mid Education_i] \neq 0$. This means that, if the c_i is not taken into account, δ estimate is likely to be biased.

There are several common estimation approaches when utilizing panel data. The first is the 'random effects' approach. This approach assumes that the time-invariant individual heterogeneity c_i is present, but this heterogeneity is uncorrelated with the regressors. This assumption means that the ability characteristics do not affect education years i.e., $E[c_i \mid Education_i] = 0$. However, the key concern is that the random effects approach does not eliminate the c_i. Hence, the random effect assumption is a difficult one, particularly in the presence of endogenous variables such as education years. If in the random effect model, the composite error term v_{it} consists of c_i and u_{it}, then the presence of c_i causes the serial correlation of the error terms over time. Hence, computed standard errors may be incorrect and estimators inefficient. In the form of the Generalized Least Squares (GLS), the random effect deals with these correlated error terms. As previously mentioned, however, the issue with the random effect approach is that the estimator δ maybe be biased since education is likely to be correlated with ability or family background. The second panel data estimation approach is the 'first-differencing'. This first difference approach involves subtracting the lag value of each variable from its current value over time. This is the difference process that leads to the elimination of the time-invariant unobserved individual effect. Thus, we have;

$$Income_{it} - Income_{i,t-1} = (X_{it} - X_{i,t-1})\beta + u_{it} - u_{i,t-1} \dashrightarrow \Delta Income_i = \Delta X_i\beta + \Delta u_i \tag{2}$$

As seen in equation 2 above, the c^i has been eliminated and the new error term is now uncorrelated with the years of education regressor, and hence the problem of endogeneity has been solved. The 'fixed effect' is a third panel data estimation approach. Like the 'first difference', the fixed effect approach involves subtracting the average value of a given variable from its absolute value over time. Again, this leads to the elimination of the time-invariant unobserved individual effect, and hence the endogeneity in the education regressor is severely dealt with. The is shown as:

$$Income_{it} - \overline{Income}_i = (X_{it} - \bar{X}_i)\beta + u_{it} - \bar{u}_i \dashrightarrow Inc\ddot{o}me_i = \ddot{X}_i\beta + \ddot{u}_i \quad (3)$$

However, there is a worrisome problem with the first-difference/fixed-effect approaches, especially in the presence of endogenous regressors i.e., years of education in the well-being function above. In these two approaches, all regressors are assumed to be correlated with the individual unobserved effect and are hence differenced. The key problem is that the time-invariant regressors are at the same time not clearly identified or eliminated. Specifically, education does not sufficiently vary across time since most household heads – the key unit of analyses in our survey data – have finished schooling. Hence, a time-differencing approach will expunge this key regressor, and its coefficient cannot be clearly identified. Since education is a key regressor of interest in the well-being analyses, a time-differencing method is hence not effective.

To deal with this kind of problem, Hausman and Taylor (HT) developed an instrumental variable (IV) approach that can be conveniently used with the panel data. The Hausman and Taylor IV, henceforth the HTIV, serves as an intermediate approach between the random effect and the fixed effect. While the random effect assumes all variables to be exogenous, the fixed effects assume all variables to be endogenous. Hence the HT approach takes some variables to be exogenous and others to be endogenous. Thereafter, valid instruments are utilised in a way that endogenous regressors are purged of endogeneity, unobserved heterogeneity problems and time-invariant regressors are correctly identified. Specifically, the exogenous, time-varying regressors are used as instruments for the endogenous, time-constant variables. In achieving Objective One of the current study, we used the HTIV approach to not only deal with the issues of endogeneity and unobserved individual heterogeneity, but also for the clear identification of the household endowment regressor(s) – years of education that does not sufficiently vary across time. More so, we included in our transformed panel model a sample correction term to account for sample selection bias in employment, as suggested in Wooldridge (1995). The selection correction term was generated from a sample selection probit equation of employment status. Since we have a panel data, we estimated a sample selection panel probit model for the employment variable and computed the Inverse Mills Ratio (IMR) to be an additional regressor in the panel structural model. Hence, instead of estimating the general panel data model in equation (1) above, we estimated an expanded version, in line with the HTIV with a sample selection correction term specified as:

$$lnIncome_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\delta_1 + Z_{2i}\delta_2 + c_i + \theta IMR_{it} + u_{it} \quad (4)$$

Where X_{1it} is a vector of exogenous, time-variant variables (e.g., age, area of residence, occupation, industry of occupation) uncorrelated with unobserved heterogeneity c_i. X_{2it} is a vector of endogenous, time-variant variables (illness reporting and employment) initially correlated with unobserved heterogeneity c_i. Z_{1i} is a vector of exogenous, time-invariant variables (gender, race) uncorrelated with unobserved heterogeneity. Z_{2i} is a vector of endogenous, time-invariant variables (completed years of schooling) initially correlated with unobserved heterogeneity. In summary, all $X_{1'S}$ and $Z_{1'S}$ are exogenous i.e., $Cov(c_i \mid X_{1it}\ ,\ Z_{1i}) = 0$ and all $X_{2'S}$ and $Z_{2'S}$ are endogenous i.e., $Cov(c_i \mid X_{2it}\ ,\ Z_{2i})\ \neq 0$. IMR_{it} is the inverse mills ratio representing the sample selection correction term. So, the idea is that solving the endogeneity, unobserved heterogeneity and identification problems will involve the use of exogenous variables X_{1it}, Z_{1i} as valid instruments to both cleanse and identify the endogenous variables X_{2it}, Z_{2i}. The choice of the instruments meets the two conditions for instrumentation (see instrument discussion in section 4 below). X_{1it} and Z_{1i} not only fulfil the instrument exogeneity requirements but correlate with endogenous variables of interest. IMR_{it}it corrects for sample selection bias of formal employment. β_1, β_2, δ_1, and θ are vectors of parameters to be estimated. All regressors are assumed to be orthogonal to an idiosyncratic error term, u_{it}. Another advantage of the HTIV is that it is estimated on the basis of the random effects panel data model, so that it uses a weighting procedure to deal with the serial correlation in the error term, u_{it}. Hence, the HTIV produces unbiased, consistent and efficient estimates. The associated factors of the panel household economic well-being provide the inputs in analysing the drivers of inequality, poverty, income of the bottom 40% and pro-poor growth in Nigeria.

To examine the impact of human capital endowments on income of the bottom 40%

To analyse the drivers of income by decile of the bottom 40% of the population, the study employed the panel quantile regression (QRPD) in the context of the IV technique. Quantile regression (QR), unlike the mean regression, allows for covariates to have marginal effects that vary with households' positions in living standard distributions. That is, it estimates various quantiles of the response variable. QR provides a richer characterisation of the data, allowing the consideration of the impact of a covariate on the entire distribution of the outcome variable and not merely its conditional mean (Baum, 2013). Several previous studies (Ivan, 2011; Nwakuya and Ijomah, 2020) have attempted to estimate quantile regression in the context of the fixed-effect and random-effect panel data. As mentioned previously, the random effect technique makes a strong assumption that the unobserved heterogeneity is uncorrelated with the regressors. The fixed-effect demeaning technique eliminates the unobserved heterogeneity, so that the problem of endogeneity suffered by endogenous regressors is solved. However, the fixed-effect technique leaves the time-invariant regressors in the model unidentified. To solve this problem, we invoke the IV procedure in equation (4) in the estimation of the quantile regression panel data. Recently, Powell (2015) has developed a panel quantile regression technique that not only takes account of the individual fixed effect but uses the IV approach of the HT nature to correctly identify the endogenous, time-invariant regressors of interest. The QRPD with non-additive fixed effect under the IV approach is stated as follows;

$$Q[Y_{it}/X_{1it}, X_{2it}, Z_{1i}, Z_{2i} \ \tau] = X_{1it}\beta_1(\tau) + X_{2it}\beta_2(\tau) + Z_{1i}\delta_1(\tau) + Z_{2i}\delta_2(\tau) \quad (5)$$

The panel quantile regression model with corrected unobserved individual specific heterogeneity describes the τ-specific conditional quantile of an outcome variable Y_{it} given a list of vectors of exogenous ($X_{1it} Z_{1i}$) and endogenous ($X_{2it} Z_{2i}$) covariates. Simply put, the panel model above specifies the conditional distribution of Y_it given the exogenous and endogenous covariates and a_i. The β coefficient gives the τ^{th} panel quantile slope differentials associated with the covariate matrix of the regressors. In line with the HT technique, the exogenous covariates will serve as instruments for the endogenous covariates. The set of the outcome and explanatory variables to be utilized are as earlier listed under Equation 4 above.

Analysing the impact of human capital on national, urban-rural and northern-southern poverty

Initially, we intended to use the Fairlie in decomposing the poverty differential. However, it is important to note at once that the Fairlie non-linear decomposition is mainly applied to the analysis of cross-sectional data. Notably, the Fairlie is not developed to be analysed in a panel data context. This presents a limitation since the current study has access to panel data. However, considering the superiority of the panel regression analysis for causal inference, we first present a panel logit analysis of the determinants of poverty at the national and regional levels of Nigeria. Then, to serve as support analysis, we are only able to use the Fairlie in analysing the drivers of regional poverty gap for each repeated cross-section of the panel data.

In analysing the panel logit of the human capital and labour market drivers of poverty, this study invokes the IV approach in Equation (4). As previously detailed, the random effect approach assumes that the unobserved heterogeneity is present in the model, but uncorrelated with the regressors. The first difference and fixed effects cures the unobserved heterogeneity problem, but at the same time eliminates the time-invariant regressors. Hence, to attempt to solve the endogeneity, unobserved heterogeneity problems, as well as identify the time-invariant regressors, we modify Equation (4) in terms of the panel logit random effect model of poverty drivers. Then we include in this model all the variables that served as instruments and also the sample selection term from the HTIV model above. This is shown as:

$$P(Poverty_{it} = 1\,|X_{it}) = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\delta_1 + Z_{2i}\delta_2 + \theta IMR_{it} \tag{6}$$

Where $Poverty_{it}$ is the poverty outcome variable measured across both unit *i* and time *t*. $P(Poverty_{it} = 1\,|X_{it}, Z_i)$ is the probability that the household *i* is in poverty at time *t* given a set covariates X_{it} and Z_i, X_{it} and Z_i vectors of regressors are as previously defined. IMR_{it} is the sample selection term that attempts to correct for sample selection bias of employment in the labour market. To support the analyses of Equation 6, we decompose poverty gap for each repeated cross-section data. This is because, as mentioned above, the Fairlie non-linear decomposition is not designed to be analysed in the panel data context. Hence, to see the human capital factors that contribute to the urban-rural poverty gap, Equation (7) below was estimated for each of the cross-section data. Following Fairlie (2003), the decomposition for a non-linear equation such as Y=F(Xβ̂) can be written as:

$$\hat{Y}^R - \hat{Y}^U = \left[\sum_{i=1}^{R} \frac{F\left(X_i^R \hat{\beta}^U\right)}{N^R} - \sum_{i=1}^{N^U} \frac{F\left(X_i^U \hat{\beta}^U\right)}{N^U}\right] + \left[\sum_{i=1}^{N^R} \frac{F\left(X_i^R \hat{\beta}^R\right)}{N^R} - \sum_{i=1}^{N^R} \frac{F\left(X_i^R \hat{\beta}^U\right)}{N^R}\right] \tag{7}$$

Where N_i is the sample size for group *i*. To compute the decomposition, define $\bar{Y}^i$ as the average probability of being poor for group *i* and *F* as the cumulative distribution function. In Equation 7, *R* represents rural/northern households and *U* represents urban/southern households. The first term in bracket represents the part of the sector/zone poverty gap that is due to group differences in distributions of *X* (*covariates*), and the second term represents the part due to differences in the group processes determining levels of *Y*. The second term captures the portion of the sector/zone poverty gap due to group differences in unobserved endowments.

Evaluating the human capital drivers of income inequality in Nigeria: Regression-based inequality decomposition

Unlike most previous inequality studies that have merely used the restricted 'traditional approach' involving the estimation of inequality and it decomposition by income sources and sub-groups, the current study employed a regression-based inequality decomposition following Fields (2003). This is aimed at finding out the household human capital factors that determine income inequality in the country. The Fields approach mainly involves the decomposition of predicted log of income, that is, running an OLS and using its predicted value in a regression-based decomposition. In this study, instead of running the normal OLS, which could yield biased estimates, we exploit the panel IV approach within the random-effect framework to estimate the predicted log of income to be utilized in the regression-based inequality decomposition. Specifically, in the first step, we follow the panel framework to estimate an income-generating function, but accounting for endogeneity problems using the IV approach. Since the HTIV approach is estimated within the random-effect framework, we are also able to account for the serial correlation problems in error terms. Hence, we invoke the HTIV approach specified in Equation 4 in the first step estimation. In a second step, we use the predicted value of log income from the first step estimation to run the regression-based inequality decomposition. Since the predicted value of log income came from the instrumented first step estimation, we believe that the final regression-based inequality decomposition is fairly unbiased.

From a general form of income-generating function of y income, x_j the j-th explanatory variable, b_j coefficient and ε error term, the Fields approach estimates the share of the log-variance of income that is attributable to the j-th explanatory variable;

$$S_{j,FIELDS} = \frac{\hat{b}_j . cov(x_j, lny)}{\delta^2(lny)} \tag{8}$$

Where $\hat{b}_j$ is the coefficient of the j-th explanatory factor estimated from an OLS regression, δ^2 (lny) is the variance of the dependent variable and $cov(x_j, lny)$ is the covariance between the j-th explanatory factor and the dependent variable. The S_j refers to whether the contribution of factor x_j is inequality-decreasing $(S_j < 0)$ or inequality-increasing $(S_j > 0)$. Thus, we can write 8 as;

$$\sum_{j=1}^{k} S_{j,FIELDS} = \frac{\sum_{j=1}^{k} \hat{b}_j \, . cov(x_j, lny)}{\delta^2(lny)} \tag{9}$$

To ascertain the impact of human capital endowment on regional and geopolitical zone income gap

To ascertain the factors contributing to income/expenditure gap in the rural-urban and northern-southern areas, this study utilized the linear panel Kitagawa-Oaxaca-Blinder (KOB) decomposition. Initially, the current study proposed to use the Oaxaca-Blinder decomposition. However, the Oaxaca-Blinder decomposition can only be used to estimate decomposition levels in the context of cross-sectional data.

Currently, the Kitagawa-Oaxaca-Blinder decomposition has been extended to the analyses of the panel data (Kroger and Hartmann, 2021). This has been introduced to fit research questions focusing on the decomposition of group-based differences in change over time. This extension takes an interventionist approach to model the effect of an intervention between time periods on the outcome of interest. The idea of the interventionist approach is that we take the initial differences in levels between the groups (e.g., urban-rural) at the base time point 2010 as given. Then, the question is how the difference between the urban and rural income have changed if either the change in endowments had been different? The panel KOB decomposition involves a two-step procedure. The first step involves the estimation of a linear panel regression model. In the second step, the saved estimated panel regression model is utilized in the levels and changes decomposition. It is crucial to note that the interpretation of results produced by decompositions of levels or change depend on the assumptions made in the initial estimated panel regression model. This means that for the results of the decomposition to be interpreted in a causal manner, the estimators of the regressors in the original regression would need to be unbiased.

As previously mentioned, panel data presents several techniques that employ some assumptions to ensure the estimation of unbiased coefficients. The first difference and the fixed effects methods of differencing and demeaning procedure removes unobserved heterogeneity, but at the same time eliminates the time-constant regressors. In achieving the above objective, this study employed the panel IV approach to not only deal with endogeneity problem, but consistently identify all regressors. As mentioned above, the panel KOB decomposition takes two important steps. In the first step we estimate a similar panel IV model using the random-effect approach. Hence, we deal with endogeneity and serial correlation problems, since this model is estimated in an error component framework. We then use this estimated model to produce unbiased decomposition results in the second step.

In the second step, the panel KOB decomposition follows a multi-level approach of varying time points nested within units. Differences between units and time points in a multi-level framework can both be seen as a form of difference-in-differences decomposition. Given two groups 'A and B' for which we have data for two time periods t and s, the change in the outcome difference between the two groups and between the two points in time is given by;

$$\Delta Y = \Delta Y^A - \Delta Y^B = \left(E(Y_t^A) - E(Y_t^B)\right) - \left(E(Y_s^A) - E(Y_s^B)\right) \quad (10)$$

Where ΔY denotes change in household income, A and B refer to rural and urban areas/north and south regions respectively. Subscripts t and s refer to time periods 2010 and 2019 respectively. The equation above describing change in income gaps is determined by three components comprising endowments, coefficients and interactions expressed below:

$$\Delta Y = \Delta E + \Delta I + \Delta C \quad (11)$$

$$\Delta E = E_A - E_B = [E(X_t^A) - E(X_s^A)]\beta_s^A - [E(X_t^B) - E(X_s^B)]\beta_s^B \quad (12)$$

$$\Delta C = C_A - C_B = E(X_s^A)\,(\beta_t^A - \beta_s^A) - E(X_s^B)\,(\beta_t^B - \beta_s^A) \quad (13)$$

$$\Delta I = I_A - I_B = [E(X_t^A) - E(X_s^A)]\,(\beta_t^A - \beta_s^A) - [E(X_t^B) - E(X_s^B)](\beta_t^B - \beta_s^B) \quad (14)$$

Where ΔE is the endowment effect obtained by subtracting the groups' compositional changes over time weighted by their initial coefficients at time s. ΔC is the coefficient effect, referring to the change in income gap due to change in coefficients over time between the groups, given the groups' initial differences in endowments at time *s*. ΔI is the interaction effect, denoting the interaction between the change in endowments and coefficients.

Analyses of growth inclusiveness and its human capital endowment sources

Pro-poor growth rate

Based on the relative pro-poor growth measure recently proposed by Kakwani et al (2010), the current study analyzed growth inclusiveness and its sources in Nigeria. The study used the first and the last wave of the Nigeria panel surveys to analyze pro-poor income growth rate over 2010-2019 period.

Given *x* to be the mean per capita income of a household, with its function as $f(x)$, the real mean of the population at time t is shown as:

$$\mu_t = \int_0^\infty x f(x)\, dx \quad (15)$$

Following from (15), the growth rate of mean population income between year 2010 and 2019 is expressed as:

$$\tau = \Delta \ln(\mu) \quad (16)$$

In line with Kakwani and pernia (2000), a pro-poor growth reduces inequality, whereas a pro-rich growth increases inequality. Comprehending growth involves examining its distributive pattern. This involves linking growth to income distribution. Hence the need to state a social welfare function, which gives a bigger weight to the utility of the poor, relative to that of the rich. If $u(x)$ is utility function increasing in *x*, then a money metric social welfare function can be stated as:

$$u(x^*) = \int_0^\infty u(x)w(x)f(x)dx \quad (17)$$

Where x^* is income that is equally distributed to every individual in the society. $w(x)$ is the weighting function capturing the deprivation of the poor, relative to that of the rich in the society. If the log utility function $u(x)$ is expressed as $ln(x)$, then a social welfare function can be derived as:

$$\ln(x^*) = \ln(\mu) - \ln(I) \tag{18}$$

Where I indicates inequality. If the first difference of (18) is derived, then we have:

$$\tau^* = \tau - \varphi \tag{19}$$

Where τ^* is the growth rate of money metric social welfare function or the pro-poor growth rate as in Kakwani et al (2010), τ is the growth rate of mean income, φ is the growth rate of inequality. If $\tau^* < \tau$, then there is a loss in the income growth of the poor due to increase in inequality. However, if $\tau^* > \tau$, then there is a gain in the income growth of the poor due to decrease in inequality. Growth is pro-poor (pro-rich) if there is a gain (loss) in growth rate of mean income.

Policy simulation of the impact of specific human capital endowments on pro-poor growth

In this study, the specific sources of household wellbeing to be analyzed are education and health. The idea is to investigate the impact of specific policy variable on pro-poor household income growth through counterfactual simulation[1]. This is achieved by calculating the differences in pro-poor growth without and with eliminating inequality in specific sources of well-being growth.

Factual and counterfactual distribution scenarios:

Factual scenario: $\Delta\tau^*(0)$ = pro-poor growth calculated with specific sources of wellbeing as observed.

Counterfactual scenario: $\Delta\tau^*(1)$ = pro-poor growth calculated with specific sources of wellbeing adjusted. That is, after inequality in the source of wellbeing is eliminated and its average value distributed to all households.

Actual Impact: $\Delta\theta = \Delta\tau^*(1) - \Delta\tau^*(0)$ = this refers to the differential impact of counterfactual and factual distributions on pro-poor income growth.

If $\Delta\theta > 0$, then the specific source of household wellbeing has a clear positive effect on pro-poor income growth.

4. Sample selection problem of employment and instruments for endogenous variables

The employment engagement in the formal sector involves non-random choices of individuals in the labour market. These non-random decisions depend on several observed and unobserved individual and household characteristics. If we fail to address the potential sample selection bias resulting from unobservable variables, the estimates from the well-being-generating function would be clearly biased and inconsistent. In line with Wooldridge's suggestion in handling sampling bias in panel data, the sample selection follows two steps. First, since we have panel data, we run a sample selection panel probit model of employment, and then generate a sample selection term (inverse mills ratio). Note that the instrumental variable for the sample selection indicator to be used in the employment selection equation is the non-self-cluster proportion of employment and education i.e., the proportion of the employed and educated in the neighbourhood/community of a household head. African households normally interact and copy each other, hence forming a social environment, which affects their decision-making. This means, if a non-self-cluster-level measure of an endogenous variable is attributed to a household belonging to a cluster, it determines the household's decision-making on that variable (Mwabu et al., forthcoming).

To identify and purge the endogenous years of education variable of endogeneity problems, we use the age and the non-self-cluster mean years of schooling variables. We assume that the non-self-cluster mean years of schooling variable is strongly correlated with education and also independent of unobserved variables. The age variable is time-variant, and clearly exogenous. Hence, it passes the instrument exogeneity test. Age is also correlated with education especially in Africa, where older people have higher completed years of schooling relative to the younger ones. For the endogenous health variable (measured with illness dummy), we will follow related literature (Mwabu et al., forthcoming) to use household land ownership as instrument to cleanse it of potential endogeneity and unobserved heterogeneity problems. The land size owned by the household is exogenous to its income-generating decisions. In line with related literature in Asia and Africa (Akin et al., 1986; Mwabu, 1986), since land property is a good proxy for wealth, it is believed to strongly correlate with health status.

5. Household panel datasets

Data for this study was drawn from the 2010/11, 2012/13, 2015/16 and 2018/19 rounds of the nationally representative Nigerian General Household Surveys (NGHS) - Panel spanning a decade or the 2010s in the Gregorian calendar. The NGHS – Panel follows the same household over time. This makes it a powerful tool for studying the dynamics of change in poverty and inequality in Nigeria. The NGHS - Panel is representative of the 36 states in Nigeria and Federal Capital Territory (FCT), Abuja, with each of its rounds targeting around 5,000 households (Nigerian National Bureau of Statistics (NBS), 2019). It is a sub-component of the NGHS (2006-2009), covering around 22,000 households. The surveys contain indicators for the monitoring and evaluation of socio-demographic characteristics, education and health, housing characteristics and household assets, consumption, food security and shocks, employment, income-generating activities, labour and time use, agriculture, housing, household enterprises, immunization programmes, child nutrition, information technology.

The surveys cover a wide range of socio-economic topics, which are collected through three different survey instruments (questionnaires). These include; the Household Questionnaire, the Agriculture Questionnaire and the Community Questionnaire (NBS, 2019). In each of the surveys, 10 households were selected per EAs, where the EAs/clusters are 500, so that the total targeted number of households is roughly 5,000. The samples are proportionally selected in the states, such that different states have different samples sizes. The surveys were produced by the Nigeria National Bureau of Statistics (NBS) in collaboration with Federal Ministry of Agriculture and Rural Development (FMA & RD), National Food Reserve Agency (NFRA) and the World Bank (WB) (NBS, 2019). The Nigerian General Household Surveys (NGHS) – Panel datasets mentioned above can be assessed from http://www.nigerianstat.gov.ng/nada/index.php/catalog

Panel sample attrition

The GHS-Panel is a panel survey, so that every possible effort was made to maintain as many households as possible in the sample. Households that had moved away from their previous locations were interviewed in a separate tracking phase. However, there is some level of attrition for some various reasons. Hence, it must be noted that the four waves of the Nigeria GHS panel does not yield a straightforward balanced panel data. Roughly, the same number of households interviewed in the first wave were followed in the second and third wave with an average minimal attrition of only 7% across wave one and two. The main reason for the sample attrition was due to security challenges that prevented revisit of some enumeration areas in Borno and

Yobe state in the North-East Zone. Other reasons were households refusing to be re-interviewed, no longer traceable, or members had died since Wave 1 (NBS, 2019).

However, in the fourth wave there was a partial refresh of the GHS-Panel sample. The refresh was conducted in order to maintain the integrity and representativeness of the sample. New refresh EAs were selected from the same sampling frame as the original GHS-Panel sample in the wave one period, resulting in a total refresh sample of approximately 3,600 households. In addition, a 'long panel' sample of 1,500 households, designed to be nationally representative, was included to enable continued longitudinal analyses for the sample going back to 2010. In the long-panel sample, about 1,425 households were successfully interviewed, with sample reduction resulting from security issues in the North West zone. Hence, the first, second and the third waves yield a balanced panel of up to 4,407 households with the minimal attrition accounted for, whereas the four waves put together yield an unbalanced panel. From waves one, two and three, we can generate an equivalent sample size of 4,407 households for each of the waves, so that the total number of households in the balanced panel sample for the three waves is 4,407 x 3. In wave four, all the 4,407 sample of same households were not re-interviewed. Hence, there is a long form panel of 1,590 interviewed households in wave one. Out of these 1,590 interviewed households forming a long form panel across wave one, two and three, 1,425 were re-interviewed in wave four. This means the long form panel attrition for the four waves is a minimal 10.4%. Hence, to form a balanced panel for the four waves, we calculate the sample, so that each wave has a total of 1,425 households. Hence, the total number of households in the balanced panel sample for the four waves yields 1425 x 4. It is this balanced panel that we used in the current study. This enabled us to track the updated dynamics of poverty, inequality and growth in Nigeria[2].

6. Findings

Sample selection model of employment

The non-random choices of securing a job in the formal sector depend on observed and unobserved individual and household characteristics. The estimates of the panel well-being-generating function would be biased and inconsistent if we fail to address the potential sample selection problem due to unobservable variables. As previously mentioned, the sample selection process is a panel probit model of employment since we already have a panel data. The instrument(s) used for the sample selection indicator are the non-self-cluster proportion of employment and education. Before interpreting the models for the primary objectives of the study, we first present the sample selection model in table 1 below.

Table 1: Panel (Random effect) probit estimates of employment

Variables	Coefficients, Col. 1	Marginal Effects, Col. 2
Gender (1 = female and 0 = otherwise)	-0.0134***	-0.0050***
	(0.0003)	(0.0001)
Sector (1 = rural and 0 = otherwise)	-0.0076***	-0.0028***
	(0.0003)	(0.0001)
Zone (1 = northern and 0 = otherwise)	-0.0114***	-0.0043***
	(0.0003)	(0.0001)
Married (1 = yes and 0 = otherwise)	-0.0460***	-0.0171***
	(0.0003)	(0.0001)
Age of household head	-0.0007***	-0.0003***
	(7.8306)	(2.9206)
Non-self-cluster employment	-1.5715***	-0.5851***
	(0.0021)	(0.0008)
Non-self-cluster years of education	0.0088***	0.0033***
	(0.0001)	(0.0000)
Land ownership cluster proportion	-0.0029***	-0.0011***
	(0.0001)	(0.0000)

continued next page

Table 1 Continued

Variables	Coefficients, Col. 1	Marginal Effects, Col. 2
Constant	1.3123***	
	(0.0018)	
Sigma_u	0.1549	
	(0.0005)	
Rho	0.0234	
	(0.0002)	
Wald chi2(8)	601060.44	
Prob > chi2	0.0000	
Number of groups	1425	1425
Number of observation	5694	5694

Note: [1]Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Column 2 of Table 1 above reports the marginal effects of the employment model. To ease interpretation in a non-linear setting, we choose the marginal effects over the coefficient estimates. According to the marginal estimates, the household head is 0.59% less likely to find work for a unit increase in the non-self-proportion of employment in a neighbourhood hosting the excluded household. This is not surprising, since a typical scenario in Nigeria is that the employment of members of a given community hosting the household does not directly guarantee that the household head will find work. However, a unit increase in years of schooling of neighbourhood hosting the household leads to 0.003% chances that the household head will find a job. This is expected since a bandwagon effect of schooling is common in Nigeria. Individuals are more propelled to strive to register for schooling when they see their peers consistently going to school. Hence, a rise in years of schooling of the community tends to raise that of the household head, which in turn leads to a rise in their chances of employment.

As in the pooled cross-section work of Mwabu et el (forthcoming), the age of the household head has a decreasing effect on employment. Relative to the male household heads, the female heads are 0.005 less likely to find a job. As expected, household heads living in rural and northern parts of the country are less likely to find a formal job, relative to their urban and southern counterparts. More so, an increase in the proportion of landed properties owned by the community does not guarantee that the house head will find formal employment.

Impact of human capital endowment and labour market participation on household well-being in Nigeria.

Table 2 below reports the causal impact of human capital endowment and labour market participation on household wellbeing using various panel data estimation techniques. In column 1 of Table 2, the exogenous (gender) and endogenous (education) variables that do not vary sufficiently across time were not well identified in the process of dealing with the unobserved individual heterogeneity in the panel fixed effect model. The panel random effect model in column 2 does not eliminate these variables, but also does not very well account for individual unobserved heterogeneity that is possibly correlated with the endogenous variable (education). Hence, we choose the panel HTIV model (exploiting both the fixed and random effect features), since it not only deals with the problem of endogeneity using the IV approach, but precisely and significantly identifies all right-hand-side variables, and at the same time yielding relatively smaller standard errors. In short, the panel HTIV model in column 3 yields clear causal estimates.

Table 2: Drivers of log of income per capita (measured with per capita household expenditure)

Variables	Fixed Effect, Col. 1	Random Effect, Col. 2	HTIV Effect, Col. 3
Gender (1 = female and 0 = otherwise)	0.0000	0.0521*	0.0484***
	(0.0000)	(0.0297)	(0.0003)
Sector (1 = rural and 0 = otherwise)	-0.0400	-0.1377***	-0.0733***
	(0.0296)	(0.0219)	(0.0001)
Zone (1 = northern and 0 = otherwise)	-0.0922**	-0.2316***	-0.1580***
	(0.0385)	(0.0219)	(0.0002)
Age	-0.0105**	-0.0070*	-0.0109***
	(0.0048)	(0.0036)	(0.0000)
Age squared	0.0001**	0.0001**	0.0001***
	(0.0000)	(0.0000)	(0.0000)
Land size ownership	-0.0970	0.0035	0.0012***
	(0.2367)	(0.0079)	(0.0001)
Education years	0.0000	0.0036	0.0922***
	(0.0000)	(0.0028)	(0.0009)
Health (1 = sick and 0 = otherwise)	-0.0277	-0.0217	-0.0256***
	(0.0277)	(0.0241)	(0.0002)
Formal employment (1 = yes and 0 otherwise)	-0.0163	0.0056	0.0065***
	(0.0254)	(0.0192)	(0.0002)
Year dummy (1 = 2019/12 and 0 = 2011/10)	0.1843***	0.1753***	0.1823***
	(0.0204)	(0.0205)	(0.0001)

Inverse mills ratio			0.5742***
			(0.0029)
Constant	12.6278***	11.8023***	10.5156***
	(1.7951)	(0.1197)	(0.0113)
F(8,1424)	11.5		
Prob > F	0.0000		
Wald chi2(10,11)		294.28	401546.3
Prob > chi2		0.0000	0.0000
Number of groups	1425	1425	1425
Number of observations	5699	5699	5699

Note: [1]Standard errors are in parentheses. [2]Significance level: *** p<0.01, ** p<0.05, * p<0.1
[3]In the HTIV model, we have used age and land size variables as instruments for education and health respectively.
[4]All regressions were estimated with a household panel spanning 4 waves (2010/11, 2012/13, 2015/16, 2018/19).
[5]Inverse mills ratio is the sample selection term controlling for bias in formal employment. This was generated from our sample selection panel probit model and then included in the panel HTIV regression.

The results in Table 2 (column 3) show that education is positively and significantly correlated with household economic well-being. An additional year of schooling leads to about 0.092 log-points increase in household economic well-being. Improving education would increase household welfare through enhanced occupation. This finding is not only in line with the panel data work of Biyase and Zwane (2017) for South Africa, but also consistent with results obtained in the pooled data works of Baye and Epo (2015) and Mwabu et al (forthcoming) for Cameroon and Kenya respectively. Other key effort-based variables included in the HTIV model in column 3 above are formal sector employment and health status. To identify the effect of health (measured with sickness dummy) on household wellbeing, we instrumented sickness with household land ownership. In line with the literature (Akin et al., 1986; Mwabu, 1986), land size owned by the household is exogenous to its income-generating decisions, and since land is a good proxy for wealth, it is believed to strongly correlate with health status.

Hence, the HTIV model in column 3 of Table 2 corrects for endogeneity and unobserved heterogeneity associated with the health variable (sickness reporting). Households reporting sickness suffered economic well-being decline in the order of about 0.023 log-points relative to those not reporting sickness within the panel periods[3]. This emanates from the fact that chronic ill health is likely to be associated with minimal or no work effort, limited savings/investment, and less motivation for better living standards. Column 3 of Table 2 shows that correcting for sample selectivity bias, securing formal sector employment is associated with 0.007 log-point increase in household economic well-being relative to informal sector workers. Similar to the result in Mwabu et al (forthcoming), this kind of finding, together with the positive education effect, implies that well-educated household heads are more likely, relative to their less educated counterparts, to secure a job in the formal sector.

The exogenous variables in the well-being-generating function are age, age squared, gender, location and regions of residence. Results in column 3 of Table 2 show that age has an increasing effect on household economic well-being. In specific terms, below 40 years of age of the head of the household, any additional year associates negatively with household economic well-being, above 40 years any additional year improves household welfare. This finding is consistent with the present-time experiences of households in Nigeria. Unstable jobs, underemployment or even unemployment usually characterise the early-life stage of most household heads. However, getting older tends to increase earnings through accumulated experience. Contrary to the general expectation of gender bias, though in line with some existing literature e.g., Mwabu et al (forthcoming), being a female household head raises wellbeing by 0.048 log points. This positive effect of having a female household head on per capita consumption is because the female household heads have higher consumption expenditures relative to the male household heads. We tabulated food and non-food purchases across male and female-headed households[4] and found that female household heads make relatively higher purchases. This is unsurprising since it is typical for Nigerian women to visit the market more often than men.

In terms of location and region of residence, the negative effects on well-being in Nigeria are expected. This is because the poor and the less educated people tend to concentrate more in the rural areas and northern regions of the country. Hence, it is not surprising that residing in the rural and northern regions of Nigeria reduces well-being by 0.073 and 0.158 log points. Land ownership of the household head associates positively with the household economic well-being. Specifically, increase in the ownership of more landed properties leads to about 0.001 log-point increase in household economic well-being. This finding, though contrary to some existing related studies (Mwabu et al, forthcoming), is expected. As previously stated, land property is a good proxy for wealth, so that increase in its ownership by a household head has the tendency to increase the well-being of the household. Looking at the year dummy, households enjoyed better welfare in later years on average i.e., 2019-12 relative to 2011-10 periods. Specifically, in the 2019-12 period on average, households enjoyed well-being increase in the order of 0.182 log points compared to the 2011-10 period[5].

Impact of human capital endowment on income of the bottom 40% at the national, regional and zonal levels in Nigeria.

Table 3 below reports quantiles causal impact of education and health variables on household well-being at not just the national level, but sectoral and zonal levels. As earlier noted, the advantage of the quantile estimates over the mean estimates reported in Table 2 above is that we see the causal impacts at various quantile points of the household income distribution. Here, we have employed the (Powel, 2016) panel quantile regression model that leverages the IV technique to not only account for unobserved individual heterogeneity, but cleanse endogenous variables of issues of endogeneity.

Table 3: Quantiles impact of human capital endowment on log of household income per capita

Covariates of interest	Quantiles				
	0.1	0.25	0.5	0.75	0.9
National Estimates					
Education	0.0013	0.0092***	0.0006	0.0009	0.0002
	(0.0016)	0.0012	(0.0012)	(0.0012)	(0.0015)
Health	0.0852	0.0925	-0.0278	-0.0203	-0.1421
	(0.0037)	(0.0038)	(0.0022)	(0.0019)	(0.0046)
Constant	10.7341***	11.0719***	11.6730***	12.1759***	12.6685***
	(0.0065)	(0.0049)	(0.0049)	(0.0053)	(0.0063)
Regional Estimates (Rural)					
Education	-0.0078***	-0.0033**	0.0148***	0.0083***	0.0141***
	(0.0019)	(0.0014)	(0.0014)	(0.0014)	(0.0017)
Health	0.0118	0.0832	-0.0108	0.0016	0.0241
	(0.0014)	(0.0037)	(0.0014)	(0.0005)	(0.0020)
Constant	10.7794***	11.1286***	11.4051***	11.9790***	12.3581***
	(0.0074)	(0.0058)	(0.0057)	(0.0061)	(0.0071)
Regional Estimates (Urban)					
Education	0.0152***	0.0196***	0.0129***	-0.0030	0.0034
	(0.0027)	0.0022	(0.0022)	(0.0023)	(0.0026)
Health	-0.0597	0.0913	0.0829	-0.0285	-0.0226
	(0.0031)	0.0038	0.0037	(0.0022)	(0.0020)
Constant	10.8213***	11.1541***	11.7257***	12.4420***	12.8220***
	(0.0115)	(0.0094)	(0.0090)	(0.0097)	(0.0112)
Zonal Estimates (Northern)					
Education	-0.0050**	-0.0004	0.0127***	0.0219***	0.0067***
	0.0022	(0.0015)	(0.0016)	(0.0016)	(0.0021)
Health	0.1169	-0.0640	-0.0877	0.0147	0.0352
	0.0043	(0.0032)	(0.0037)	(0.0016)	(0.0024)

Constant	10.6696***	11.0722***	11.3449***	11.7130***	12.3160***
	0.0087	(0.0063)	(0.0067)	(0.0066)	(0.0088)
Zonal Estimates (Southern)					
Education	-0.0259***	0.0083***	0.0034**	0.0023	0.0120***
	(0.0025)	(0.0017)	(0.0016)	(0.0018)	(0.0021)
Health	-0.0827	0.1009	-0.1318	-0.0463	-0.0218
	(0.0036)	(0.0040)	(0.0045)	(0.0028)	(0.0019)
Constant	11.2065***	11.2387***	11.8465***	12.3582***	12.6828***
	(0.0094)	(0.0071)	(0.0068)	(0.0074)	(0.0084)

Note: [1]Standard errors in parentheses. [2]Significance levels: *** p<0.01, ** p<0.05, * p<0.1
[3]Age and land size variables serve as instruments for education and health respectively.
[4]The health variable is measured with sickness dummy.

At the national level, we see positive relation between household well-being and years of education (table 3). This result is similar to the quantile estimates found in Cameroon and Vietnam (Fambon, 2017; Yamada, 2018). At the bottom 10% of the distribution (households below the bottom 40%), an additional year of schooling leads to about 0.001 income percentile point increase. However, this observed effect tends to diminish as we move towards the upper part of the distribution. This simply suggests that the causal effect of schooling is higher for poor households relative to the rich households at the national level. Moreover, we see more interesting dynamics for the quantiles impact of health (measured with sickness dummy) at the national level. For the 10 and 25% of the distribution, ill-health of the household head does not reduce the household well-being. This is of little surprise since the initial tabulation of quantiles of household income over the sickness dummy[6] reveals the concentration of relatively more healthy households at the 10^{th} and 20^{th} quantiles position. However, more sick household heads tend to concentrate towards the middle of the distribution, so that ill health status of the household head reduces the household well-being. Specifically, the national estimates in Table 3 above reveals that reporting ill-health at the 50, 75 and 90% of the distribution leads to 0.028, 0.020 and 0.142 percentile decline in household income respectively.

In terms of the rural estimates in the same Table 3, we see a negative relation between years of schooling and well-being at the 10 and 25% of the distribution. This is expected since the initial tabulation of quantiles of household income over years of education[7] shows the concentration of relatively less educated household heads (i.e., household heads with less number of years of schooling) at 10th and 20th quantile positions. However, more educated household heads tend to concentrate at the middle and the upper part of the distribution. Hence, an additional year of schooling of household heads at the 50, 75 and 90% of the distribution increases the well-being of the household by 0.015, 0.008, and 0.014 percentile points respectively. The ill-health estimates at the rural level looks similar to those at the national level, with a difference of a positive relation between ill-health and well-being at the upper part of distribution. We attribute this to the concentration of relatively low number of sick household heads at this upper part of the income distribution.

The next part of Table 3 above shows the urban estimates of quantiles causal impact of household human capital endowment on household well-being. Here, we also see interesting dynamics across the various quantile positions. At the bottom 10% of the distribution (households below the bottom 40%), an extra year of schooling leads to about 0.015 percentile point increase in household well-being. However, we see this effect tending to decrease as we move towards the upper part of the income distribution. Hence, as mentioned before, schooling tends to give more weight to the earnings of the poor relative to that of the wealthy. Ill-health reporting decreases the well-being of the household, with the effect relatively higher at the bottom 10% of the distribution. However, we a see a positive relation at the middle of the distribution, but this is attributable to the concentration of relatively less ill-health reporting in the 25th and 50th quantile positions of the distribution.

The next in table 3 above is the zonal (northern) estimates. To aid interpretation, we first tabulate the quantiles of household income over years of education[8]. This tabulation reveals the concentration of relatively less number of years of education at the bottom of the distribution. This complements the results and justifies the negative relation that we see between household well-being and years of schooling at the 10 and 25% of the distribution in northern Nigeria. In other words, household heads at the bottom of the distribution with relatively less number of years of education do not see an increase in their well-being level. The reverse seems to be the case when we look at the estimates for households towards the middle and the upper part of the distribution. Specifically, an extra year of schooling of household heads at the middle of the distribution significantly increases household income by 0.013 percentile points, though with this effect tending to diminish as we move towards the top of the distribution. In terms of health effect, we see fluctuating signs which as noted earlier is attributable to the varying concentration of ill-health reporting at different points of the distribution.

In the southern zone, results seem somewhat different. An additional year of schooling of household heads at the 25% of the distribution increases household well-being by 0.008 percentile point. Following the trend reported above, this effect tends to decrease as we move up the distribution. However, we see a little reverse as we get to the top of the distribution. Specifically, at the top of the distribution, households earn approximately 0.012 income percentile point increase for an extra year of education. The negative relation we see at the bottom 10% of the distribution confirms what is typical in the southern part of Nigeria – where some individuals from poor households finish schooling (usually in very low quality schools) but are still unable to find create reasonable value to raise their well-being. As expected, increase in ill-health decreases the household well-being, except in the lower part of the distribution (specifically 25%) where ill-health is relatively less reported.

Impact of human capital and labour market engagement on national, urban-rural and northern-southern poverty in Nigeria

Table 4 below reports various panel logit models of factors associated with poverty at the national level in Nigeria. Note that poverty is defined to take 1 when the household income is below a predefined poverty line.[9] The fixed effect model in column 1 does not clearly identify the key endogenous variable – years of schooling in the process of dealing with unobserved subject heterogeneity. Hence, we prefer the random effect model in column 2 since it not only identifies all right hand side variables but also includes various instruments correlated with the key endogenous variables in the model. The random effect model also includes the inverse mills ratio that attempts to control for sample selection bias in formal employment status of the household head. Hence we believe that the random effect model in column 2 provides a relatively more causal estimates of poverty drivers in Nigeria.

Table 4: Drivers of poverty – National estimates

Variables	Fixed Effect Logit, Col. 1	Random Effect Logit, Col. 2	Marginal Random Effect Logit, Col. 3
Gender (1 = female and 0 = otherwise)	0.0000	-0.0909***	-0.0181***
	(0.0000)	(0.0007)	(0.0001)
Sector (1 = rural and 0 = otherwise)	0.0460***	0 .3524***	0.0610***
	(0.0006)	(0.0005)	(0.0001)
Zone (1 = northern and 0 = otherwise)	0.1775***	0 .6476***	0.1286***
	(0.0007)	(0.0005)	(0.0001)
Age	0.0380***	0 .0345***	0.0068***
	(0.0001)	(0.0001)	(0.0000)
Age squared	-0.0004***	-0.0003***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
Land size ownership		-0.0176***	-0.0035***
		(0.0002)	(0.0000)
Education years	0.0000	-0.0010***	-0.0009***
	(0.0000)	(0.0001)	(0.0000)
Health (1= sick and 0 = otherwise)	0.0210***	0.0175***	0.0036***
	(0.0006)	(0.0006)	(0.0001)
Formal employment (1 = yes and 0 otherwise)	0.1097***	-0.0145***	-0.0029***
	(0.0005)	(0.0005)	(0.0001)
Non-self-cluster employment		-1.1426***	-0.2269***
		(0.0175)	(0.0035)

Non-self-cluster years of education		-0.0785***	-0.0156***
		(0.0002)	(0.0000)
Year dummy (1 = 2019/12 and 0 = 2011/10)	-0.5547		-0.1040
	(0.0005)		(0.0001)
Inverse mills ratio		-0.7482***	-0.1486***
		(0.0193)	(0.0038)
Constant		1.6391***	
		(0.0041)	
LR chi2	1640000		
Prob > chi2	0.0000		
Wald chi2		4880000	
Prob chi2		0.0000	
Number of groups	1425	1425	1425
Number of observations	5699	5699	5699

Note: [1]Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
[2]The constant term is omitted in column 3 because of reported marginal effects.

To facilitate interpretation in a non-linear setting, we choose the marginal effects in column 3 over the coefficient estimates in column 2 (Table 4). According to the marginal estimates, years of education of the household head correlate negatively with poverty – a finding that is consistent with that of Biyase and Zwane (2017) for South Africa. Specifically, an additional year of schooling leads to a 0.001% decline in likelihood that the household will be in poverty. This outcome strengthens the findings in column 3 of Table 2 above, alluding that an extra year of education of the household head raises the well-being of the household. Notably, the smaller size effect of education on income (0.092), relative to the effect on poverty (0.001), is attributable to differences in methods i.e., the HTIV model versus the random effect model (REM). The HTIV model uses an in-built IV technique to deal with endogeneity in education whereas the REM mainly assumes exogeneity. In terms of health status, household heads reporting ill-health are 0.004% more likely to be in poverty relative to their healthy counterparts. This confirms the result in column 3 of Table 2 above, suggesting that households reporting ill-health suffered economic well-being decline relative to those not reporting ill-health. Formally employed household heads are 0.003% less likely to be in poverty relative to their informally employed counterparts. This also corroborates our finding above, implying that formal sector employment is associated with some increase in household economic well-being. More so, this finding is consistent with the work of Biyase and Zwane (2017), showing that the probability of being poor for a household member whose head is employed is relatively lower in South Africa. Unexpectedly, though in line with our finding above, female-headed households are less likely to be in poverty than their male counterparts. More so, compared to the urban and southern areas, households residing in rural and northern areas are 0.061% and 0.129% more likely to be in poverty – consistent with the findings of Habyarimana, et al (2015) and Biyase and Zwane (2017) in Rwanda and South Africa respectively. This implies that rural and northern areas should continue to be a major target of poverty alleviation policies in not only Nigeria, but Africa in general.

As expected, land ownership decreases the probability of being in poverty. Specifically, a unit increase in ownership of landed properties decreases the chances of a household slipping below the poverty line by 0.004%. In accordance with the result in column 3 of Table 2 above, the chances of a household being in poverty is 0.104% less in 2019-12 on average, relative to their chances of impoverishment in the 2011-10 period. However, this does not directly imply higher levels of absolute poverty in 2011-10 compared to that of the 2019-12 period. A unit increase in the non-self-proportion of employment in a neighbourhood hosting the excluded household decreases the chances of the household falling below the poverty line by 0.227%. This result is somewhat expected, since a typical community doing well could set up some philanthropic structures that will enable an average household enjoy improved well-being. A rise in years of schooling of the community hosting the household reduces the chances of the household slipping into poverty by 0.016%. As previously noted, individuals are more propelled to strive to register for schooling when they see that their peers are consistently going to school (peer effect). Hence, improving their chances of a good living standard.

Table 5: Drivers of poverty – Sectoral Estimates

Variables	**Rural Estimates**		**Urban Estimates**	
	Random Effect Logit, Col. 1	**Marginal Random Effect Logit, Col. 2**	**Random Effect Logit, Col. 3**	**Marginal Random Effect Logit, Col. 4**
Gender (1 = female and 0 = otherwise)	-0.0928***	-0.0178***	-0.2989***	-0.0595***
	(0.0008)	(0.0002)	(0.0011)	(0.0002)
Sector (1 = rural and 0 = otherwise)				
Age	0.0238***	0.0046***	0.0334***	0.0066***
	(0.0001)	(0.0000)	(0.0002)	(0.0000)
Age squared	-0.0003***	-0.0001	-0.0004***	-0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Land size ownership	-0.0955***	-0.0184***	0.0419***	0.0083***
	(0.0002)	(0.0000)	(0.0003)	(0.0001)
Education years	-0.0119***	-0.0023***	0.0201***	0.0040***
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Health (1= sick and 0 = otherwise)	-0.0994***	0.0191***	0.1416***	0.0282***
	(0.0007)	(0.0001)	(0.0010)	(0.0002)
Formal employment (1 = yes and 0 otherwise)	-0.0414***	-0.0080***	-0.0093***	0.0078***
	(0.0006)	(0.0001)	(0.0008)	(0.0002)

Non-self-cluster employment	-9.1738***	-1.7642***	-6.1654***	-1.2261***
	(0.0209)	(0.0040)	(0.0293)	(0.0058)
Non-self-cluster years of education	-0.0168***	-0.0032***	-0.0768***	-0.0153**
	(0.0002)	(0.0000)	(0.0003)	(0.0001)
Year dummy (1 = 2019/12 and 0 = 2011/10)	-0.2822***	-0.0543***	-1.0585***	-0.2105
	(0.0006)	(0.0001)	(0.0010)	(0.0002)
Inverse mills ratio	8.2973***	1.5956***	5.6355***	1.1207***
	(0.0230)	(0.0044)	(0.0320)	(0.0063)
Constant	2.1667***		0.8809***	
	(0.0052)		(0.0064)	
Wald chi2	893669.1		1680000	
Prob chi2	0.0000		0.0000	
Number of groups	4001	4001	1,693	1,693
Number of observations	1379	1379	989	989

Note: [1]Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
[2]The constant term is somitted in column 2 and 4 because of reported marginal effects.

The primary results are the national estimates of drivers of poverty in Table 4 above. We have interpreted these results – focusing on all covariates. To complement interpretation of national estimates, we explain the drivers of poverty at the sectoral level – intentionally focusing on the key variables[10] of the study (marginal estimates in columns 2 and 4 of Table 5). Results at the rural level in column 2 show that education years correlates negatively with poverty. However, the reverse seems to be the case at the urban level – as shown in column 4. These results are somewhat expected. We have relatively more households reporting poverty at the rural level, so that any effort made in terms of an additional year of schooling makes a significant impact on their well-being. However, in the urban areas, most households are already living well, so that an additional year of schooling does not really add to their already high level of well-being. More so, this could be attributable to sample-size bias. Relative to the more representative national-level sample, the regional samples are smaller in size. For instance, we see that the national-level estimates, which are the main results, conform to expected signs. This is similar to what we see for formal employment across rural and urban areas. Formal employment reduces the likelihood that households in rural areas will fall into poverty. However, it is not so for households in the urban areas – suggesting that most households in the urban areas are already out of poverty. Ill-health reporting increases the likelihood of households in the rural areas moving into poverty by 0.019%. This is expected, since it is typical in rural areas that ill-health of household heads prevents them from working to sustain the family. More so, ill health reporting of household heads in urban areas increases the chances of households slipping into poverty by 0.028%.

Table 6: Drivers of poverty – Zonal Estimates

Variables	Northern Estimates		Southern Estimates	
	Random Effect Logit, Col. 1	Marginal Random Effect Logit, Col. 2	Random Effect Logit, Col. 3	Marginal Random Effect Logit, Col. 4
Gender (1 = female and 0 = otherwise)	-0.2675***	-0.0485***	0.0193***	0.0041***
	(0.0011)	(0.0002)	(0.0008)	(0.0002)
Sector (1 = rural and 0 = otherwise)				
Age	-0.0092***	-0.0017***	0.0527***	0.0111***
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Age squared	0.0000***	0.0000***	-0.0004	-0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Land size ownership	-0.0421***	-0.0076***	-0.0045***	-0.0010***
	(0.0003)	(0.0000)	(0.0002)	(0.0001)
Education years	-0.0204***	-0.0037***	0.0141***	0.0030***
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Health (1= sick and 0 = otherwise)	-0.0853***	-0.0155***	0.0941***	0.0199***
	(0.0009)	(0.0002)	(0.0007)	(0.0002)
Formal employment (1 = yes and 0 otherwise)	-0.1597***	-0.0290***	0.0367***	0.0078***
	(0.0007)	(0.0000)	(0.0006)	(0.0001)
Non-self-cluster employment	-13.387***	-2.4279***	0.3793***	0.0802***
	(0.0318)	(0.0057)	(0.0205)	(0.0043)
Non-self-cluster years of education	-0.0551***	-0.0010***	-0.0492***	-0.0104***
	(0.0003)	(0.0001)	(0.0002)	(0.0000)
Year dummy (1 = 2019/12 and 0 = 2011/10)	-0.0612	-0.0111***	-0.9241	-0.1953
	(0.0008)	(0.0001)	(0.0007)	(0.0001)
Inverse mills ratio	12.4978***	2.2607***	-2.4340***	-0.5144***
	(0.0351)	(0.0063)	(0.0227)	(0.0048)
Constant	3.4042***		0.8565***	
	(0.0064)		(0.0052)	
Wald chi2	755956.8		2320000	
Prob chi2	0.0000		0.0000	
Number of groups	1018	1018	1058	1058
Number of observations	2774	2774	2920	2920

Note: [1]Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
[2]The constant term is omitted in column 2 and 4 because of reported marginal effects.

As already mentioned, the primary results are the national estimates of drivers of poverty in Table 4 above. We have already interpreted them focusing on all covariates. Here, we interpret the drivers of poverty at the zonal level - intentionally focusing on the key variables of the study (marginal estimates in columns 2 and 4 of Table 6). Similar to the results at the rural level in column 2 of Table 5 above, results at the northern level in column 2 of Table 6 show that education years correlates negatively with poverty. However, the reverse seems to be the case in terms of signs at the southern level – as shown in column 4 of Table 6. The southern result is also similar to what we see at the urban level in column 4 of Table 5 above. These results are somewhat expected. We have relatively more households reporting poverty at the northern zone, so that any effort made in terms of additional year of education raises the well-being of the household. However, in the southern zone, most households are already living well, so that an additional year of education does not really add to their already high well-being level. This is similar to what we see for formal employment across northern and southern zones. Formal employment reduces the likelihood that households in the northern zones fall into poverty. However, it is not so for households in the southern zones – suggesting that most households in these zones are already out of poverty. Ill-health reporting does not increase the likelihood that households in the northern zones move into poverty. This is of little surprise since a relatively low number of households have reported ill-health in the northern parts of the country. However, ill-health reporting of household heads in the southern zones increases the chances of households slipping into poverty, specifically by 0.020%.

Impact of human capital endowment on urban-rural and northern-southern poverty gap in Nigeria

Table 4 above reports the national level logit model of drivers of poverty – among other covariates - controlling for the effect of sector and zone on poverty. We can see from this table that the coefficients of sector and zone are significant. This implies the existence of statistically significant rural-urban and northern-southern poverty gaps. In other words, the households in the rural and northern regions are significantly more in poverty than the households in the urban and southern regions. If significant rural-urban and northern-southern gaps exist, further decomposition of it into explanatory components would give more insight into what causes this gap in poverty. In order to achieve this, we use the methods of decomposing inequality in a binary variable into contributing factors – the fairlie non-linear decompositions presented in Table 7 below.

Table 7: Fairlie non-linear decompositions of sector and zone poverty gaps

Part one	Decomposition of rural-urban poverty differential				
Variable	**2010/11**	**2012/13**	**2014/15**	**2018/19**	**2010-2019**
Model					
Pr(Y!=0|G=0)	0.6475	0.6410	0.4462	0.2270	0.4649
Pr(Y!=0|G=1)	0.6584	0.6888	0.6991	0.4484	0.6299
Difference	-0.0109	-0.0479	-0.2529	-0.2214	-0.1650
Covariates contributions	0.0521	0.0276	-0.0737	-0.1632	-0.0520
Explained by					
Gender	0.0066***	-0.0001***	0.0026***	-0.0167***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Zone	-0.0101***	0.0182***	-0.0775***	-0.1303***	-0.0545***
	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Age	-0.0164***	0.0031***	0.0107***	0.0040***	-0.0050***
	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0000)
Age squared	0.0114***	0.0013***	-0.0121***	-0.0079***	0.0014***
	(0.0003)	(0.0001)	(0.0001)	(0.0002)	(0.0000)
Land size ownership	0.0264***	0.0015***	-0.0110***	0.0042***	0.0042***
	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Education years	0.0002***	-0.0016***	0.0010***	0.0084***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Health	-0.0002***	0.0004***	-0.0016***	0.0010***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Formal employment	-0.0001***	-0.0004***	-0.0001***	0.0002***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Non-self-cluster employment	-0.0071***	-0.0000	0.0196***	-0.0170***	0.0005***
	(0.0001)	(0.0001)	(0.0002)	(0.0000)	(0.0000)
Non-self-cluster education years	0.0066***	0.0010***	-0.0005***	-0.0084***	-0.0022***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Inverse mills ratio	0.0357***	0.0045***	-0.0025***	0.0012***	0.0027***
	(0.0001)	(0.0002)	(0.0000)	(0.0000)	(0.0000)

continued next page

Table 7 Continued

Part two	**Decomposition of northern-southern poverty differential**				
Variable	**2010/11**	**2012/13**	**2014/15**	**2018/19**	**2010-2019**
Model					
Pr(Y!=0|G=0)	0.6437	0.6667	0.4750	0.1977	0.4879
Pr(Y!=0|G=1)	0.6685	0.6837	0.7953	0.5604	0.6761
Difference	-0.0248	-0.0169	-0.3203	-0.3627	-0.1882
Covariates contributions	0.0330	0.0059	-0.0661	-0.0415	-0.0219
Explained by					
Gender	-0.0001***	0.0017***	0.0114***	0.0045***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Sector	-0.0034***	-0.0207***	-0.0499***	-0.0493***	-0.0323***
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0000)
Age	0.0257***	0.0219***	-0.0010***	0.0433***	0.0363***
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0001)
Age squared	-0.0213***	-0.0249***	-0.0126***	-0.0389***	-0.0414***
	(0.0001)	(0.0002)	(0.0002)	(0.0000)	(0.0001)
Land size ownership	0.0039***	0.0014***	-0.0126***	0.0065***	0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Education years	0.0009***	0.0010***	0.0002***	0.0012***	0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Health	-0.0019***	-0.0003***	-0.0039***	0.0019***	0.0011***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Formal employment	-0.0003***	0.0004***	-0.0001***	0.0012***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Non-self-cluster employment	0.0215***	0.0153***	0.0003***	0.0041***	-0.0015***
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Non-self-cluster education years	0.0015***	0.0014***	0.0013***	-0.0058***	0.0011***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Inverse mills ratio	0.0064***	.0096***	0.0001	-0.0019***	0.0164***
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0001)

Note: [1]Standard errors in parentheses.
[2]Significance Levels: *** p<0.01, ** p<0.05, * p<0.1
[3]G=0=Urban/Southern; G=1=Rural/Northern

In part one of Table 7 above, we see the sector poverty gap decomposition reported for each cross-section period of the data and for the panel data period (2010-2019) in the last column. Notably, the covariates (including the key covariates – education and health) are statistically significant. In the decompositions (in both part one and two of Table 7), a positive 'difference' implies that households in rural and northern regions are relatively less in poverty while a negative difference shows the opposite. It is clear from the decomposition results (in both part one and two of Table 7), that the 'difference term' shows negative values across all individual cross-section periods and panel sample period. We conclude, therefore, that poverty is more concentrated against households in the rural and northern regions of the country. We now look at the key contributing factors to these observed gaps in poverty. In part one of Table 7 reporting sector differentials, we see that the poverty gap rises with education years in 2010/11, 2015/16, 2018/19 and in the panel period 2010-2019. We also see that the poverty gap rises with health status in 2012/13, 2015/16 and in the panel periods. The poverty gaps only rise with employment in 2018/19, with an insignificant result in the panel period. Since the poverty gap rises with education and health in the panel periods 2010-2019, we conclude that the observed rural-urban poverty gap is mainly determined by differences in education and health status, among other covariates – such as age squared and land ownership. The results of northern-southern poverty gaps reported in part two of the same table are similar to results of rural-urban gaps. In part two of Table 7, we see that the poverty gaps rise with education years across all individual sample periods and the panel period. The poverty gap rises with health status in 2015/16 and in the panel period. More so, it only rises with employment in 2015/16. We conclude from these results that the observed northern-southern poverty gap is mainly determined by differences in education and health status among differences in other covariates – such as gender, age and land ownership.

Human capital endowment drivers of total income inequality using a regression-based decomposition in Nigeria

The table 8 below reports the regression-based decomposition of income inequality (ineqrbd in STATA) developed in Fiorio and Jenkins (2007) by using the decomposition rules/formulae of Fields (2003). The Fields decomposition rule involves obtaining the predicted values of income and using these predicted values to estimate the decomposition results. We have leveraged on this procedure in an attempt to account for endogeneity in the key endogenous variables (education and health). Hence, we first run an instrumental panel regression that accounts for endogeneity issues[11]. Then, we obtain the predicted values of income from the first step random effect regression and apply them to form the dependent variable in the final decomposition exercise. Since the random effect estimator is identical to an OLS estimator applied to conveniently transformed variables (Wooldridge, 2002), the results of the decomposition analysis following Fields approach have been obtained from an OLS regression of the transformed variables[12].

Table 8: Regression-based inequality decomposition results

Inequality contributions						
Variable	**Regression coefficient (Col.1)**	**100*s_f (Col. 2)**	**S_f (Col. 3)**	**100*m_f/m (Col. 4)**	**CV_f (Col. 5)**	**CV_f/ CV(total) (Col. 6)**
Gender	0.0544***	4.8121	0.0006	0.0883	2.0719	157.0799
Sector	-0.1437***	27.4157	0.0036	-0.8678	-0.6506	-49.3224
Zone	-0.2236***	63.0760	0.0083	-0.9361	-1.0261	-77.7926
Age	0.0002***	0.2884	0.0000	0.1075	0.2975	22.5575
Land size	0.0015***	0.2539	0.0000	0.1009	0.1942	14.7234
Education years	0.0037***	1.0099	0.0001	0.3636	0.3655	27.7125
Health	-0.0293***	0.0358	0.0000	-0.0464	-2.1069	-159.7397
Employment	0.0188***	0.3756	0.0000	0.1003	0.7810	59.2117
Constant	11.7645***					14.7234
Residual		4.8121	0.0006	-0.0000	-40000	-30000

Note: [1]The dependent variable is predicted log of income obtained from panel IV regression.
[2]Age and land size variables serve as instruments in the IV regression for education and health respectively.
[3]The health variable is measured with sickness dummy.
[4]Reference categories of dummies are already mentioned in the tables above.
[5]Standard errors of regression coefficients are intentionally not reported but available on request.
[6]Significance: *** p<0.01, ** p<0.05, * p<0.1.
[7]Results are based on STATA 14.0 ineqrbd developed by Fiorio and Jenkins (2007) following Fields (2003).
[8]Proportionate inequality contribution (s_f) = rho_f*sd(f)/sd(Total), where sd(f) = std.dev of factor. S_f = s_f*CV(Total), where CV is covariance. Mean of factor = m_f = mean(f).

Column 2 in Table 8 above shows the factors contributing to measured income inequality. The term s_f in column 2 refers to the main contribution to inequality in total of each factor (Fiorio and Jenkins, 2007). Hence, the interpretation of the regression-based inequality is focused on this term. We see from column 2 of Table 8 that the highest contributors to income inequality are sector and region of residence in 2010-2019[13]. Other sources that contributed marginally in explaining measured income inequality are gender, education and formal employment. Specifically, residing in the rural sector and northern zone relative to the urban sector and southern zone contributes to roughly 27% and 63% of inequality in total income respectively. These results are expected and not surprising following marked gaps in welfare between households in rural and urban areas, and northern and southern zones in Nigeria. In Nigeria, poverty level remains highest in rural and northern areas relative to the urban and southern areas (NBS, 2020). In Cameroon, the cross-sectional work of Arrey (2020) has revealed similar findings – rural residency is the largest contributor (40.4%) to inequality in Cameroon – suggesting the need to seek workable ways of equalising economic and political opportunities between urban and rural areas in Africa.

Impact of human capital endowment on rural-urban and northern-southern income gap in Nigeria

In the analysis in Table 8 above, we have identified sector (rural-urban) and zone (northern-southern) as the main contributors to income inequality in Nigeria. Since this is the case, it is important to examine the human capital factors that drive inequality within rural-urban sector and northern-southern zone. Hence, in Table 9 below, we present the human capital endowment drivers of income gap in rural-urban and northern-southern regions in Nigeria. We have used the recently developed panel Kitagawa-Oaxaca-Blinder decomposition (Kroger and Hartmann, 2021). This approach has been introduced to fit research questions focusing on the decomposition of group-based differences in change over time.

As previously mentioned, the panel Kitagawa-Oaxaca-Blinder decomposition involves a two-step procedure. The first step involves the estimation of a linear panel regression model. In the second step, the saved estimated panel regression model is utilized in the levels and changes decomposition. Hence, for the results of the decomposition to be interpreted in a causal manner, the estimators of the regressors in the original panel regression would need to be unbiased. To provide for this, we run an IV panel model in the first step regression – instrumenting for the key endogenous variables (education and health). In the second step, the estimated IV regression is employed in the decomposition. Following the standard procedure in (Kroger and Hartmann, 2021), the first step estimation is a regression of log of income on the time variable, the group variable of interest and the household endowment variable(s) of interest. In this study involving the analysis of the human capital endowment effect, we have focused on years of schooling and health as the household endowment variables. In the second step, we use the estimation from the first step, the endowment variables, the group variable of interest and the time (year) variable.

The 'decomposition of levels' results in column 1 and 2 in part one of Table 9 above reveal similar results, though differ somewhat in part two of the same table. The decomposition results in column 2 (both part one and two) of Table 9 show a relatively much lesser random effect. We choose to focus on this decomposition result since it attempts to deal with the issues of endogeneity using the IV approach in its first step panel regression. The first row showing 'outcome observed' for column 2 of Table 9 denotes the mean group differences in log household income, as estimated from the observed data across the panel data periods – 2010-2019. The rows in the decomposition section (part one) of the same column 2 show the results of the income gaps' decomposition into an endowments part, a coefficient part and an interaction part and the part due to the random effect. The rows in the decomposition % section (part two) display the same results, but now in percentage terms i.e., the relative contribution of the four decomposition effects to the overall income gap. We see in the 'outcome observed' row that the rural-urban income gap tends to rise over time from 0.066 in 2010/11 to 0.261 log incomes in 2018/19. Under part one of this table, showing absolute contribution – endowment contributes 0.003 log incomes to the income gap in 2010/11 and 0.008 log incomes in 2018/19. In terms of relative contribution – endowment contributes about 10.71% to the income gap in 2010/11 and roughly 10.82% in 2018/19. The coefficient part also contributed positively to the observed increasing rural-urban income gap. In absolute terms, the coefficient contributes 0.028 log incomes to the gap in 2010/2011 and 0.072 log incomes in 2018/19. However, the contribution of the interaction term is negative – highlighting the importance of disaggregated evaluation of the endowment and the coefficient parts of the decomposition.

Table 9: Decomposition of rural-urban income inequality/gap

	Decomposition of Levels, Col. 1			**Decomposition of Levels – IV Approach, Col. 2**		
	2010/2011	**2015/2016**	**2018/2019**	**2010/2011**	**2015/2016**	**2018/2019**
Outcome observed	0.066	0.220***	0.261***	0.066	0.220***	0.261***
	(0.045)	(0.040)	(0.034)	(0.042)	(0.040)	(0.030)
Decomposition	**Part One**					
Endowments	0.003	0.003	0.008	0.003	0.003	0.008
	(0.004)	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)
Coefficients	0.028	0.147***	0.072	0.028	0.147***	0.073*
	(0.038)	(0.033)	(0.039)	(0.034)	(0.033)	(0.037)
Interaction	-0.004	-0.004	-0.008	-0.004	-0.004	-0.008
	(0.006)	(0.006)	(0.006)	(0.004)	(0.006)	(0.007)
Random effect	0.038	0.074***	0.189***	0.000	0.000	0.000
	(0.020)	(0.021)	(0.022)	(0.000)	(0.000)	(0.000)
Total	0.066	0.220	0.261	0.028	0.146***	0.074*
	(0.045)	(0.040)	(0.034)	(0.034)	(0.034)	(0.037)
Decomposition %	**Part Two**					
Endowments	4.868	1.316	3.060	10.714	1.979	10.818
	(32.091)	(1.222)	(1.815)	(168.273)	(3.143)	(173.648)
Coefficients	43.044	66.649***	27.562	102.920	100.461***	99.365
	(341.466)	(9.796)	(18.220)	(162.070)	(4.504)	(156.348)
Interaction	-6.090	-1.624	-2.895	-14.522	-2.440	-10.183
	(26.419)	(2.945)	(2.430)	(320.884)	(6.178)	(310.132)
Random effect	58.177	33.659***	72.274***	0.000	0.000	0.000
	(326.496)	(9.605)	(18.178)	(0.000)	(0.000)	(0.000)
Total	100.000	100.000	100.000	100.000	100.000	100.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

continued next page

Table 9 Continued

	Decomposition of Change, Col. 1			Decomposition of Change – IV Approach, Col. 2		
	2010/2011	2015/2016	2018/2019	2010/2011	2015/2016	2018/2019
Outcome observed	0.000	0.154**	0.195***	0.000	0.154**	0.195***
	(0.000)	(0.058)	(0.053)	(0.000)	(0.057)	(0.046)
Decomposition	**Part One**					
Endowments	0.000	0.001	0.012	0.000	0.001	0.012
	(0.000)	(0.004)	(0.010)	(0.000)	(0.004)	(0.009)
Coefficients	0.000	0.117*	0.041	0.000	0.117**	0.042
	(0.000)	(0.049)	(0.052)	(0.000)	(0.045)	(0.057)
Interaction	0.000	0.001	-0.007	0.000	0.001	-0.007
	(0.000)	(0.006)	(0.018)	(0.000)	(0.006)	(0.020)
Random effect	0.000	0.036	0.150***	0.000	0.000	0.000
	(0.000)	(0.029)	(0.026)	(0.000)	(0.000)	(0.000)
Total	0.000	0.154**	0.195***	0.000	0.118**	0.046
	(0.000)	(0.058)	(0.053)	(0.000)	(0.045)	(0.051)
Decomposition %	**Part Two**					
Endowments	0.000	0.362	5.927	0.000	0.469	24.996
	(0.000)	(42.050)	(7.152)	(0.000)	(4.386)	(123.886)
Coefficients	0.000	76.198	20.769	0.000	99.103***	90.548
	(0.000)	(299.250)	(48.113)	(0.000)	(5.300)	(92.800)
Interaction	0.000	0.329	-3.747	0.000	0.428	-15.545
	(0.000)	(49.492)	(12.726)	(0.000)	(7.195)	(184.980)
Random effect	0.000	23.111	77.051	0.000	0.000	0.000
	(0.000)	(208.468)	(47.855)	(0.000)	(0.000)	(0.000)
Total	0.000	100.000	100.000	0.000	100.000	100.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note [1]Significance: *** p<0.01, ** p<0.05, * p<0.1.
[2]Age and land size variables serve as instruments in the IV regression for education and health respectively.
[3]By default the Kitagawa-Oaxaca-Blinder decomposition reports zeros for the reference year. Hence, 2010/2011 serving as reference year in the decomposition of change show zeros.

Further, we see a similar trend for the decomposition of change results in the lower part of Table 9 (column 2) above. We see the change in the rural-urban income gap in comparison to the reference year 2010/11 for the 2015/16 and 2018/19 periods. For the observed outcome, the gap increased between 2010/11 and 2018/19 by 0.195 log incomes. We now examine the effect of changing endowments and coefficients over time. In absolute terms, the changing endowment increased the gap by 0.012 log incomes and the changing coefficients contributed 0.042 log incomes to the increasing gap between 2010/11 and 2018/19.

Table 10: Decomposition of northern-southern income inequality/gap

	Decomposition of Levels, Col. 1			Decomposition of Levels – IV Approach, Col. 2		
	2010/2011	**2015/2016**	**2018/2019**	**2010/2011**	**2015/2016**	**2018/2019**
Outcome observed	0.134***	0.069	0.465***	0.134***	0.069	0.465***
	(0.036)	(0.039)	(0.029)	(0.039)	(0.040)	(0.032)
Decomposition	**Part One**					
Endowments	0.000	0.003	-0.001	0.000	0.003	-0.001
	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.007)
Coefficients	0.116***	0.037	0.366***	0.115***	0.036	0.367***
	(0.031)	(0.029)	(0.033)	(0.029)	(0.035)	(0.032)
Interaction	0.004	-0.002	0.002	0.004	-0.002	0.002
	(0.005)	(0.005)	(0.005)	(0.008)	(0.005)	(0.007)
Random effect	0.013	0.030	0.098***	0.000	0.000	0.000
	(0.017)	(0.020)	(0.023)	(0.000)	(0.000)	(0.000)
Total	0.134***	0.069	0.465***	0.120***	0.038	0.368***
	(0.036)	(0.039)	(0.029)	(0.029)	(0.035)	(0.032)
Decomposition %	**Part Two**					
Endowments	0.349	5.057	-0.109	0.361	9.089	-0.138
	(3.169)	(13.896)	(1.165)	(5.357)	(316.043)	(2.113)
Coefficients	86.800***	53.176	78.692***	96.144***	95.076	99.718***
	(10.743)	(418.812)	(5.013)	(4.705)	(77.769)	(0.783)
Interaction	3.113	-2.347	0.336	3.495	-4.164	0.419
	(3.872)	(23.562)	(1.110)	(7.359)	(242.386)	(2.163)
Random effect	9.737	44.114	21.081***	0.000	0.000	0.000
	(10.285)	(400.152)	(5.059)	(0.000)	(0.000)	(0.000)
Total	100.000	100.000	100.000	100.000	100.000	100.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Decomposition of Change, Col. 1			**Decomposition of Change – IV Approach, Col. 2**		
	2010/2011	**2015/2016**	**2018/2019**	**2010/2011**	**2015/2016**	**2018/2019**
Outcome observed	0.000	-0.065	0.332***	0.000	-0.065	0.332***
	(0.000)	(0.041)	(0.046)	(0.000)	(0.055)	(0.052)
Decomposition	**Part One**					
Endowments	0.000	-0.001	-0.004	0.000	-0.001	-0.004
	(0.000)	(0.003)	(0.007)	(0.000)	(0.004)	(0.008)
Coefficients	0.000	-0.081*	0.253***	0.000	-0.081*	0.253***
	(0.000)	(0.035)	(0.048)	(0.000)	(0.041)	(0.046)
Interaction	0.000	0.001	-0.002	0.000	0.001	0.014
	(0.000)	(0.004)	(0.012)	(0.000)	(0.004)	(0.011)
Random effect	0.000	0.017	0.085***	0.000	0.000	0.000

	(0.000)	(0.016)	(0.025)	(0.000)	(0.000)	(0.000)
Total	0.000	-0.065	0.332***	0.000	-0.082	0.248***
	(0.000)	(0.041)	(0.046)	(0.000)	(0.042)	(0.047)
Decomposition %	**Part Two**					
Endowments	0.000	1.693	-1.200	0.000	1.334	-1.627
	(0.000)	(15.228)	(2.230)	(0.000)	(9.239)	(4.015)
Coefficients	0.000	126.036	76.137***	0.000	99.335***	102.362***
	(0.000)	(168.161)	(9.359)	(0.000)	(4.765)	(6.596)
Interaction	0.000	-0.849	-0.583	0.000	-0.669	-0.735
	(0.000)	(9.553)	(4.004)	(0.000)	(7.574)	(6.863)
Random effect	0.000	-26.880	25.646**	0.000	0.000	0.000
	(0.000)	(164.716)	(9.144)	(0.000)	(0.000)	(0.000)
Total	0.000	100.000	100.000	0.000	100.000	100.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1Significance: *** p<0.01, ** p<0.05, * p<0.1.
2Age and land size variables serve as instruments in the IV regression for education and health respectively.
3 By default the Kitagawa-Oaxaca-Blinder decomposition reports zeros for the reference year. Hence, 2010/2011 serving as reference year in the decomposition of change show zeros.

The 'decomposition of levels' in column 1 and 2 in part one of Table 10 above also reveals similar results, though differs somewhat in part two. The decomposition results in column 2 (both part one and two) of Table 10 also show a relatively lesser random effect. We focus on these decomposition results since they attempts to deal with the endogeneity problems using the IV approach. The first row showing 'outcome observed' denotes the mean northern-southern differences in log household income, as estimated from the observed data across the panel data periods – 2010-2019. Similar to the increasing trend that we see for the 'outcome observed' for the rural-urban income gap in Table 9, the southern-northern income gap tends to rise over time from 0.134 in 2010/11 to 0.465 log incomes in 2018/19. In terms of the absolute role of endowment, coefficient and interaction, the endowment part, on average contributed a marginal 0.001 log income to the income gap in 2010/11-2018/19. The coefficient contribution to the observed increasing northern-southern income gap is relatively higher, rising from 0.115 log incomes in 2010/11 to 0.367 log incomes in 2018/19. The interaction part contributes 0.004 log incomes to the gap in 2010/2011 and 0.002 log incomes in 2018/19.

We observe a similar trend for the decomposition of change results in the lower part of the Table 10 (column 2) above. We see the change in the northern-southern income gap in comparison to the reference year 2010/11 for the 2015/16 and 2018/19 periods. For the observed outcome, the gap increased between 2010/11 and 2018/19 by 0.332 log incomes. We now look at the effect of changing endowments and coefficients over time. In absolute terms, the changing endowment tends to decrease the gap by some marginal 0.004 log incomes. However, the changing coefficients and interaction increased the gap by 0.253 and 0.014 log incomes respectively.

Impact of specific household human capital endowments on pro-poor income growth in Nigeria.

In Table 11 below, we simulate counterfactual scenarios to assess the impacts of the various well-being sources (specifically focusing on education and health) on growth pro-poorness in the 2010-2019 period. This follows several specific steps. First, the factual scenario is computed by explaining the well-being equation in (4) above, obtaining the predicted income values and using them to calculate actual and pro-poor growth rates in panels A and B using the Kakwani et al. 2010 method. Second, the counterfactual scenarios are calculated by explaining the same equation (4), but this time without the influence of the specific source of well-being under consideration, i.e., when inequality of this source is expunged by assigning its average value to all households. The predicted well-being values are then obtained and utilized in computing actual and pro-poor growth rates recorded in other panels, using the same Kakwani et al (2010) approach. Third, the impact of each income source is computed by taking the difference of actual and pro-poor growth with and without the respective sources of well-being across the 2010-2019 period.

As in Panels A and B of Table 11, we found no relative pro-poor growth for total incomes per capita over the period 2010-2019. Though observed growth in household income tends to expand – marginally over the studied period, poor households witnessed a loss attributable to some increase in inequality. In Nigeria, while inequality was 0.36 in 2004, on average it remained roughly 0.38 in 2013-2019 (NBS, 2020). The marginal expansion in well-being over the period is not surprising, following the 2014-15 oil price shock that plunged the economy's growth to negative levels in 2016. Hence, the aftermath of the oil shock partly explains the poor performance of incomes of poor households within the 2010-2019 period. Although a decline in poverty to 35% was observed in 2010 from 46% in 2004 (NBS 2012; World Bank 2013), it rose to 41 per cent more recently, in 2019 (World Bank 2013; NBS 2020). Therefore, the observation in Table 11 above, that total household income per capita was anti-poor in 2010-2019, manifests that losses in growth were driven by an increase in the number of households below the poverty line and inequality levels that have barely decreased in the country.

When we erase the effect of inequality associated with education-based source of well-being for all households, pro-poor growth increased over the studied periods. This can be clearly seen when we compare the factual scenario (panel B) and the counterfactual scenario (panel D) for equalizing education. Specifically, when we account for the difference between the factual and counterfactual growth and pro-poor growth values, actual growth for 2010 and 2019 witnesses an increase of 0.02 and 0.00 percentage points, respectively. Moreover, pro-poor growth witnesses an increase of 0.22 and 0.22 percentage points, respectively. Overtime, while average growth decreased by 0.02, pro-poor growth increased by 0.00 percentage points. This led to a net effect of a gain in growth attributable to a decline in inequality (panels C and D of Table 11). This indicates that poor households are likely to experience an expansion in their well-being if government puts in place policies that eliminate inequality in access to schooling for households at the bottom of the distribution. Hence, encouraging pro-poor policies in terms of equal opportunities in the country.

Table 11: Pro-poor growth in Nigeria: Relative pro-poor growth method of Kakwani et al (2010)

Period	2010/ 2011	2018/ 2019	Counter factual Vs. factual outcomes 2010 (I)	Counter factual Vs. factual outcomes 2019 (II)	Changes over time (II)-(I)
Factual total expenditure per capita					
A) Actual growth rate	11.67	11.68	n/a	n/a	n/a
B) Pro-poor growth rate	11.35	11.34	n/a	n/a	n/a
Outcome [gain (+)/loss (-)]	Loss in growth [-0.02] due to some increase in inequality.				
Counterfactual total expenditure per capita (when we equalized education)					
C) Simulated actual growth rate	11.69	11.68	0.02		
(line C- line A)	0.00				
(line C- line A)	-0.02				
D) Simulated pro-poor growth rate	11.57	11.56	0.22		
(line D- line B)	0.22				
(line D- line B)	0.00				
Outcome [gain (+)/loss (-)]	Equalizing education generates gain in growth [0.02] due to decline in inequality.				
Counterfactual total expenditure per capita (when we equalized health)					
E) Simulated actual growth rate	13.18	13.19	0.01		
(line E- line A)	0.00				
(line E- line A)	-0.01				
F) Simulated pro-poor growth rate	10.68	10.66	0.03		
(line F- line B)	0.03				
(line F- line B)	0.00				
Outcome [gain (+)/loss (-)]	Equalizing health generates a marginal gain in growth [0.01] due to decline in inequality.				

Note: 1Source: Computed by authors.
2n/a imply 'not applicable'.

We also see a net gain in growth, though relatively less, when we expunge the effect of inequality in health-based source of well-being by equalizing health outcomes (measured with sickness dummy) for all households. Although pro-poor growth values tend to decline when comparing the factual scenario (panel B) and the counterfactual scenario (panel F) for equalizing health, some increases were recorded for actual growth across these scenarios (panel A, relative to E). This implies that poor households are likely to experience an expansion in their per capita incomes if government puts in place policies that erase inequality in health for households located at the bottom of the distribution. Over time, while average growth decreased by 0.01, pro-poor growth increased by 0.00 percentage points – the net effect is a gain in growth attributable to an increase in overall income (a decrease in inequality) (Panels E and F of table 11). This outcome uncovers the importance of enhancing access for poorer households to health-based opportunities in the country.

7. Summary of findings and policy implications

Leveraging an up-to-date Nigeria's micro panel household survey dataset, this study has linked a panel income generating function (with effort and exogenous circumstance-related covariates) to welfare outcomes – poverty, inequality and pro-poor growth. Specifically, the paper (a) analyses the impact of human capital endowment and labour market engagement on household well-being; (b) examines the impact of human capital endowment on the income quintiles (including income of the bottom 40%); (c) assesses the impact of human capital and labour market engagement on national, urban-rural and northern-southern poverty; (d) establishes the human capital drivers of total income inequality using a regression-based decomposition in Nigeria; (e) ascertains the impact of human capital endowment on rural-urban and northern-southern income gap; (f) appraises the impact of specific household human capital endowments on pro-poor income growth in Nigeria.

In line with these specific objectives, the current study yields several crucial findings. Human capital variables – education and health – significantly affect household economic well-being in Nigeria. Specifically, an additional year of schooling significantly increases household economic well-being. In terms of health, sickness reporting of the household head significantly reduces household economic well-being. Securing formal sector job significantly increases the economic well-being of the household. Further, we uncovered that education years has significant causal impact on the income of households below the bottom 40% (specifically the households at the bottom 25%). Consolidating these findings, we show that poverty is significantly reduced for an additional year of schooling. Engaging in formal sector employment significantly reduces poverty. However, poverty is significantly increased with ill-health reporting of the household heads. These findings tend to be consistent across rural and northern parts of the country. The significant poverty bias we see in rural-urban and northern-southern regions is determined by significant differences in schooling years and health status of household heads.

The study further reveals that inequalities at the national level are mainly driven by rural-urban and northern-southern inequalities. Using a panel data regression-based decomposition approach, we uncovered that the observed income inequalities in rural-urban sectors and northern-southern geopolitical zones are mainly propelled by differences in education and health endowments of the households. The factual and counterfactual simulation analyses demonstrate that equalization of human capital endowments in terms of education and health is indeed growth-enhancing. It is suggested, in line with the above findings, that policies capable of eliminating inequalities in access to education and health will enable households at the bottom of the distribution to enjoy better and sustainable economic well-being.

Notes

1. A counterfactual effect predicts what would happen (in this case to pro-poor growth) if changes in the distribution of education and health were to be implemented.

2. Note that we have used the household identification number (hhid) and the individual identification number (indiv) to track the same household and individual across the four waves of the panel survey data.

3. Note that the reverse is the case in terms of signs when we re-categorize the sickness dummy to good-health dummy. In this case, the health variable takes 1 for good-health and 0 otherwise.

4. The tabulation is available upon request.

5. Note that our model does not accept factor variables. Hence, we have re-categorized the year variable, initially (1 = 2010/11, 2 = 2012/13, 3 = 2015/16 and 4 = 2018/19) to form a dummy (0 = earlier periods, 2010/11 and 1 = later periods, 2012/19).

6. Detailed initial tabulation of quantiles of household income over the sickness dummy is available upon request.

7. Detailed initial tabulation of quantiles of household income over years of education is available upon request.

8. Detailed tabulation is available upon request.

9. In Nigeria, the official poverty line is based on the 'cost of basic needs approach' referred to as the monetary value of food and non-food expenditures needed for an individual to achieve a basic level of welfare (NBS, 2020). This total monetary value for Nigeria equals 137,430 Naira (NLSS, 2018/2019). This value has been utilised in calculating poverty in this study.

10. This is to avoid repetition since results of non-key variables appear similar across national and regional levels.

11. Note that our results are overestimated when we fail to account for endogeneity following these procedures. Results of regression-based decomposition not accounting for endogeniety issues are available upon request.

12. The new variables are obtained by removing from the original ones a fraction λ of their average over time, where λ is calculated as a function of the variances of both the error term and the individual effect term.

13. Note that we see similar results even when we ignore the panel structure of the data and run the decomposition analysis for each cross-section component of the household data.

References

Adetayo, A. O. (2014). Analysis of farm households poverty status in Ogun states, Nigeria. *Asian Economic and Financial Review*, 4(3), 325−340.

Aigbokhan, B. E. (2008). Growth, Inequality and Poverty in Nigeria. Discussion Paper No. 3 of the United Nations Economic Commission for Africa (UNECA).

Akin, John S., Charles Griffin, David Guilkey and Barry Popkin (1986). The Demand for Primary Health Care Services in the Bicol Region of the Philippines. *Economic Development and Cultural Change*, 34(4): 755−782.

Ataguba, J. E-O., Ichoku, E. H., and William M. Fonta, M. W. (2013). Multidimensional poverty assessment: applying the capability approach. *International Journal of Social Economics,* 40(4), 331−354.

Arrey, M. A. (2020). Explaining Wellbeing and Inequality in Cameroon: A Regression-Based Decomposition. AERC Research Paper 378 African Economic Research Consortium, Nairobi.

Bahar, T. (2022). The effect of human capital on income equality: cross-sectional analysis. *Sinop Üniversitesi Sosyal Bilimler Dergisi*, 6(1), 183−199.

Baum, C. F. (2013). Quantile Regression. Applied Econometrics: EC 823. Boston College, Spring.

Baye, F.M. and Epo, B. N. (2015). Impact of Human Capital Endowments on Inequality of outcomes in Cameroon. *Review of Income and Wealth*, 61(1): 93−118.

Biyase, M. and Zwane, T. (2017). An Empirical Analysis of the Determinants of poverty and household welfare in South Africa. Munich Personal RePEc Achive (MPRA) Paper No 77085. Online at https://mpra.ub.uni-muenchen.de/77085/

Development Finance International (DFI) and Oxfam, (2020). Fighting Inequality in the Time of Covid 19: The Commitment to reducing Inequality Index, 2020.

Epo, B. N., and Baye, M. F. (2012). Determinants of Well-being and Poverty Changes in Cameroon: 2001–2007. *African Development Review*, 24(1), 18–33.

Fairlie, R. W. (2003). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. Working Papers 873, Economic Growth Center, Yale University.

Fambon, S. (2017). The Determinants of Inequality and Income Gap between Urban and Rural Areas in Cameroon: Evidence from the ECAM3 Household Survey. *Advances in Economics and Business* 5(7): 394−410.

Fiorio, C. V., and Jenkins, S. P. (2007). ineqrbd: Regression-based inequality decomposition, following Fields (2003), Motivation Model ineqrbd Example: wage inequality. Available at https://www.stata.com/meeting/13uk/fiorio_ineqrbd_UKSUG07.pdf

Habyarimana, F. Zewotir, T. and Ramroop, S. (2015). Determinants of poverty of households: semi parametric analysis of Demographic and Health survey data from Rwanda. *Journal of Economics and Behavioral Studies,* 7(3), 4755.

Ijaiya, G. T., Ijaiya, M. A., Bello, R. A., and Ajayi, M. A. (2011). Economic growth and poverty reduction in Nigeria. *International Journal of Business and Social Science*, 2(15), 147−154.

Ivan A. C. (2011). A simple approach to quantile regression for panel data. *Econometrics Journal,* 14, 368–386.

Jaiyeola, A. O., and Bayat, A. (2019). Assessment of Trends in Income Poverty in Nigeria from 2010–2013: An Analysis Based on the Nigeria General Household Survey. *Journal of Poverty,* DOI: 10.1080/10875549.2019.1668900

Kakwani, N., and Pernia, E. (2000). What is pro-poor growth? *Asian Development Review,* 16(1), 1–22.

Kakwani, N., Neri, M. C., and Son, H. H. (2010). Linkages between Pro-Poor Growth, Social Programs and Labor Market: The Recent Brazilian Experience. *World Development*, 38 (6), 881–894.

Kroger, H., and Hartmann, J. (2021). Extending the Kitagawa-Oaxaca-Blinder decomposition approach to panel data. The Stata Journal: Promoting Communications on Statistics and Stata, 21(2), 360−410.

Lee, J-W., and Lee, H. (2018). Human Capital and Income Inequality. ADBI Working Paper Series, 810.

Mamun, A., and Arfanuzzaman, M. (2020). The effects of human capital and social factors on the household income of Bangladesh: an econometric analysis. *Journal of Economic Development,* 45(3), 29−49.

Moyo, C., Mishi, S., and Ncwadi, R. (2022). Human capital development, poverty and income inequality in the Eastern Cape Province. Development Studies Research, 9:1, 36−47.

Mwabu, G. (1986). Health care Decisions at the Household Level: Results of a Rural Health Survey in Kenya. *Social Science & Medicine*, 22(3): 315−319.

Mwabu, G., Bayea, F. M., Epoa, B. N., Etyangb, M., Gachanjab, P., Metseyema, C., and Njugunab, A. (forthcoming). Poverty, Inequality and Growth in Sub-Saharan Africa: Survey Evidence from Cameroon and Kenya. An Unpublished African Economic Consortium Collaborative Research Framework Paper.

National Bureau of Statistics (NBS) (2012). Nigeria Poverty Profile 2010. Available at https://reliefweb.int/report/nigeria/nigeria-poverty-profile-2010-report

Nwakuya, M. T., and Ijomah, M. A. (2020). Panel quantile regression with penalized fixed effects and correlated random effects. *European Journal of Statistics and Probability*, 8,(1), 74−81.

NBS, (2018). Snapshot of Inequality in Nigeria. Available at http://www.nigerianstat.gov.ng/download/698

NBS, (2019). Basic Information Document. Nigeria General Household Survey – Panel, 2018−2019.

NBS, (2020). 2019 Poverty and Inequality in Nigeria: Executive Summary.

Oaxaca, R., (1973). Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review*, 14, 693–709.

Ofem, B., Akpan, U., and Umoren, V. (2010). Analysis of urban poverty and its implications on development in Uyo urban, Akwa Ibom State. *Global Journal of Social Sciences, 9*(1), 7-19

Ogbeide, E. N. S., and Agu, D. O. (2015). Poverty and Income Inequality in Nigeria: Any Causality? *Asian Economic and Financial Review,* 5(3), 439-452

Ogundari, K., and Awokuse, T. (2018). Human capital contribution to economic growth in Sub-Saharan Africa: Does health status matter more than education? *Economic Analysis and Policy,* 58, 131–140

Ohwotemu, O. L. (2010). A distributional analysis of income in Nigeria. An M.Sc thesis submitted to the department of economics, faculty of arts, university of Nigeria Nsukka .

Olaniyan, O., and Awoyemi, T. T. (2005). Inequality in the Distribution of Household Expenditure in Rural Nigeria: A Decomposition Analysis. Draft Final Research Report Submitted to the AERC, Nairobi for the Second Phase Collaborative Poverty Research Project.

Olanrewaju, O., and Timothy, T. A. (2005). Inequality in the distribution of household expenditure in rural Nigeria: a decomposition analysis. Draft Final Research Report Submitted to the AERC, Nairobi for the Second Phase Collaborative Poverty Research Project .

Olofin, O. P., Adejumo, A. V., and Sanusi, K. A. (2015). Determinants of Poverty Level in Nigeria. *Journal of Sustainable Development*, 8(1), 235–241.

Olopade, B. C., Okodua, H., Oladosun, M., and Asaleye, A. J. (2019). Human capital and poverty reduction in OPEC member-countries. *Heliyon*, 5, e02279.

Olowa, W. O., Awoyemi, T. T., Shittu, M. A., and Olowa, O. A. (2013). Effects of Remittances on Poverty among Rural Households in Nigeria. *European Journal of Sustainable Development*, 2(4), 263–284.

Omoruyi, O., and Omoyibo, K. U. (2014). Welfare Inequality in Nigeria. *Mediterranean Journal of Social Sciences*, 5(7), 579–588.

Oxfam International, (2017). Inequality in Nigeria. Exploring the Drivers. Retrieved from https://www.oxfam.org/sites/www.oxfam.org/files/file_attachments/cr-inequality-in-Nigeria 170517-en.pdf

Powell, D. (2015). Quantile Regression with Nonadditive Fixed Effects. RAND Labor and Population Working Paper.

Roemer, J. E. (2002). Equality of Opportunity: A Progress Report, *Social Choice and Welfare*, 19, (2), 455–471.

Shahpari, G., and Davoudi, P. (2014). Studying Effects of Human Capital on Income Inequality in Iran. *Procedia - Social and Behavioral Sciences* 109, 1386–1389.

Sowunmi, F. A., Akinyosoye, V. O., Okoruwa, V. O., and Omonona, B. T. (2012). The Landscape of Poverty in Nigeria: A Spatial Analysis Using Senatorial Districts- level Data. *American Journal of Economics,* 2(5), 61–74.

Su, J., and Guo, S. (2022). Human Capital and Rural Households' Vulnerability to Relative Poverty: Evidence from China. *Discrete Dynamics in Nature and Society*, 11.

United Nations, (2017). Sustainable Development Goals Report, 2017. Available at https://unstats.un. org/sdgs/files/report/2017/thesustainabledevelopmentgoalsreport2017.pdf

Usman, Z. A., Ritter, K., and Vinogradov, S. A. (2016). The Dynamics of Income Inequality in Rural Areas of Nigeria*. Asian Journal of Agricultural Extension, Economics & Sociology*, 11(4), 1–13.

World Bank, (2013). Nigeria Economic Report. World Bank, Nigeria Economic Report No.1.

World Bank (2016). Poverty Reduction in Nigeria in the Last Decade. Available at http://documents.worldbank.org/curated/en/103491483646246005/pdf/ACS19141-REVISED-PUBLIC-Pov-assessment-final.pdf

World Bank's WDI, (2021). World Bank national accounts data, and OECD National Accounts data. https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.CD&country.

Wooldridge, J. (1995). Selection Correction for Panel Data Models under Conditional Mean Independence Assumptions. *Journal of Econometrics*, 68: 115–132.

Wooldridge J. M. (2002). *Econometric Analysis of cross section and panel data*. The MIT Press, Cambridge, Massachusetts, London, England.

Yamada, T. (2018). Dynamics of Spatial Inequality and Poverty: Evidence from Two Decades of Surveys in Vietnam, 1993-2014, *Economics Bulletin*, 38(1), 404–418.

Mission

To strengthen local capacity for conducting independent, rigorous inquiry into the problems facing the management of economies in sub-Saharan Africa.

The mission rests on two basic premises: that development is more likely to occur where there is sustained sound management of the economy, and that such management is more likely to happen where there is an active, well-informed group of locally based professional economists to conduct policy-relevant research.

Bringing Rigour and Evidence to Economic Policy Making in Africa

- Improve quality.
- Ensure sustainability.
- Expand influence.

www.aercafrica.org

Learn More

www.facebook.com/aercafrica

www.instagram.com/aercafrica_official/

twitter.com/aercafrica

www.linkedin.com/school/aercafrica/

Contact Us
African Economic Research Consortium
Consortium pour la Recherche Economique en Afrique
Middle East Bank Towers,
3rd Floor, Jakaya Kikwete Road
Nairobi 00200, Kenya
Tel: +254 (0) 20 273 4150
communications@aercafrica.org