Explaining Wellbeing and Inequality in Cameroon: A Regression-Based Decomposition

Marinus Arrey Arrey

Research Paper 378

Bringing Rigour and Evidence to Economic Policy Making in Africa

AFRICAN ECONOMIC RESEARCH CONSORTIUM

CONSORTIUM POUR LA RECHERCHE ÉCONOMIQUE EN AFRIQUE

Explaining Wellbeing and Inequality in Cameroon: A Regression-Based Decomposition

By

Marinus Arrey Arrey
Faculty of Economics and Management,
University of Yaoundé II, Cameroon
Yaoundé

AERC Research Paper 378
African Economic Research Consortium, Nairobi
January 2020

THIS RESEARCH STUDY was supported by a grant from the African Economic Research Consortium. The findings, opinions and recommendations are those of the author, however, 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

ISBN 978-9966-61-071-3

Contents

List of tables

List of figures

Abstract

This study sets out to estimate the determinants of household economic wellbeing and to evaluate the relative contributions of regressed-income sources in explaining measured inequality. In particular, a regression-based decomposition approach informed by the Shapley value, the instrumental variables econometric method, and the 2007 Cameroon household consumption survey, was used. This approach provides a flexible way to accommodate variables in a multivariate context. The results indicate that the household stock of education, age, credit, being bilingual, radio and electricity influence wellbeing positively, while rural, land and dependency had a negative impact on wellbeing. Results also show that rural, credit, bilingualism, education, age, dependency and land, in that order, are the main contributors to measured income inequality, meanwhile, the constant term, media and electricity are inequality reducing. These findings have policy implications for the ongoing drive to scale down both inequality and poverty in Cameroon.

Keywords: Regression-based decomposition, Inequality, Household Economic wellbeing, Cameroon.

1. Introduction

Background

Understanding wellbeing and its distribution remains topical in both high and low-income countries. The 1990s witnessed a renaissance in theoretical and empirical attention by economists, to the distribution of income and wealth (Atkinson and Bourguignon, 2000). This was motivated by the recognition that inequality was not only an outcome but also a determinant of growth (Yuko et al., 2006). As such, the measurement and analysis of wellbeing and inequality are crucial for cognitive purposes (to know what the situation is); for analytical purposes (to understand the factors determining this situation); for policy making purposes (to design interventions best adapted to the issues); and for monitoring and evaluation purposes (to assess the effectiveness of current policies); and to determine whether the situation is changing (Coudel et al., 2002).

A glimpse into the literature indicates that a substantial amount of work in economics and social sciences has investigated the relationship between income inequality and economic growth, and a variety of social phenomena (such as political conflict, education, health, and crime) with different standpoints. For example, the classical approach (Kaldor, 1956) argues that more inequality favours capital accumulation because the rich have a higher marginal propensity to save than the poor, thereby resulting in rapid economic growth. With the modern approaches, higher initial inequality of income leads to lower economic growth through different channels or paths (Thorbecke and Charumilind, 2002). To reconcile the two approaches, Galor (2000) also argues that for a country in an early stage of development, inequality would promote growth because physical capital is scarce at this stage and its accumulation requires savings. An increased share of the rich in the population would then result in higher savings and rapid growth. On the other hand, at a later stage of development, the increased availability of physical capital raises the return on investment in human capital. But, with credit market imperfections, the poor, who do not have the ability to provide collateral, may find their access to capital curtailed (Galor and Zeira, 1993; Agion and Bolton, 1997). The poor will therefore find it difficult to invest in human capital. Income inequality would then result in a poverty trap and lower growth. As such, to cue up from the various standpoints calls for empirical evidence.

Before the past two decades, wellbeing evaluation in developing countries was based solely on economic growth. As such, all too often, GDP is interpreted as a measure of welfare or wellbeing, which it is not and was never designed to be (Bergheim, 2006). GDP only measures the market value of final goods and services produced in a country. Looking at the GDP growth of a country is fundamental to understanding its development process, but it is never sufficient to sketch a reliable picture of the welfare situation at the household level.

For instance, Cameroon experienced economic recovery from 1996 subsequent to the 1994 devaluation of the CFA Franc, registering annual real GDP growth rates of 5.2% in 1996, 4.7% in 2001 and 3.5% in 2007 (National Institute of Statistics-NIS, 2004; 2008). However, at the household level, the incidence of poverty was 53.30% in 1996, 40.20% in 2001 and 39.95% in 2007, while the Gini index of inequality was 0.406 in 1996, 0.408 in 2001 and 0.390 in 2007. The indication is that the modest macroeconomic performance was not translated into comparable improvements in the living conditions of the average Cameroonian due to disparities in the distribution of the fruits of growth.

Faced with this situation, growing policy concerns have been generated to focus on inequality of outcomes, with inequality of opportunities of human capital and other socio-economic determinants as an important channel to mitigate income inequality and enhance wellbeing. It is apparent that an initial maldistribution of human capital inputs as well as associated endowments should make it much harder for the poor to participate in, and gain from the process of economic growth. Human capital inputs have been recognised as critical factors in achieving sustained growth in productivity in some African countries (Schultz, 2003).

Rousseau (1754) stated more than two centuries ago in his Discourse on the Origin of Inequality that, as individuals departed from the "primitive state" to conformed societies where private property predominated and individuals developed a specific role in those societies, the conditions were set for the generation of all sorts of inequalities among them. It was revealed in both ECAM1 and ECAM2 that among the characteristics of households, the level of education most highly discriminated between poor and non-poor households (NIS, 2008).

Inequality of outcomes can be considered a composite indicator comprising inequality of exogenous circumstances to which an individual may not be held responsible and inequality of endogenous effort to which an individual can largely be held responsible (Baye and Epo, 2013). Presumably, the root causes of the inequality in wellbeing are the prevailing inequality of opportunity and the different levels of efforts exerted by individuals. As such, a greater equity in the distribution of educational opportunities enables the poor to capture a larger share of the benefits of economic growth, and in turn contributes to higher growth rates. In contrast, large-scale exclusion from educational opportunities results in lower economic growth and persistent income inequality (Thorbecke and Charumilind, 2002). As far as this study is concerned, we are not going to make a distinction between the two types of inequality, but we will concentrate on inequality of outcomes by considering average stock of household education, which is made up of both the prevailing inequality of opportunity and the different levels of efforts exerted by individuals.

Household living conditions are typically reflected in wellbeing sources such as: access to education, health, social amenities such as electricity, telephone and portable water. If we proxy wellbeing by household consumption expenditure and explain it in terms of dimensions of living conditions using appropriate econometric tools, then we can comfortably use parameter estimates to explain measured inequality. This process is described in the literature as regression-based decomposition analysis. Understanding how much of total inequality is captured by regressed-wellbeing sources is important for targeting the roots of inequality in Cameroon. Such an analysis is thought of as more revealing than the traditional group and source decompositions of inequality that can easily be considered simple accounting procedures with no behavioural anchor. In this context, developing a broader understanding of inequality is necessary because it is increasingly recognised as important in policy discussion, for poverty reduction, for growth and behind most criminal activities. In addition, the relevance of income inequality to economic development efforts can be judged by the spread of researchers that have kept a close focus on it in the past decades.

Research issue

A glimpse into Cameroon's economic structure and performance does not meet the criteria for a socially responsible market democracy, and does not constitute a framework that allows citizens adequate freedom of choice. After some decades of structural adjustment and reforms without any major change in the welfare of the majority in Cameroon, the emphasis of economic policies has shifted to education, health and other socio-economic determinants to mitigate income inequality and enhance wellbeing. However, these aspects manifest disparity in access, as equal opportunities are not provided for individual ability to be expressed. According to the World Bank (2005), and Ferreira and Gignoux (2008), countries where inequality of opportunities are accentuated witness low economic growth rates given that inequalities discourage investment in human capital.

Faced with this situation, many researchers have been fine-tuning ways to mitigate inequality to improve on the welfare of Cameroonians, using household data for 1996, 2001 and 2007 (Baye and Fambon, 2002; NIS, 2002; Araar, 2006; Chameni and Wendji, 2012 ; Fambon, 2010; Fambon, 2011; 2014). These studies were interested in inequality of outcomes or its decomposition using different indicators. Fambon (2014) applies the method of quantile regression and total income inequality decomposition into population sub-groups to analyse the data of the third Cameroonian household survey (ECAM3) focusing on the within-group and between-group decomposition. His results show that there are considerable differences in the average consumption expenditure of households and in within-group inequality.

Interested in the determinants of income inequality using regression-based approach are Epo et al. (2011), and Baye and Epo (2011; 2013). Baye and Epo (2011) apply the regression-based inequality decomposition approach to explore the determinants of income inequality in Cameroon using synthetic variables for education and health constructed using the multiple correspondence analysis method to reflect the multidimensional character of health and education, and the 2007 Cameroon household consumption survey. This is equally the case in Baye and Epo (2013), where the synthetic variables for education and health are largely considered as endogenous effort-related variables to explain income inequality.

Despite all these studies, knowledge on how wellbeing and inequality are determined using socio-economic characteristics such as employment status, occupation, sector of activity, zone of residence and other characteristics thought vital for policy formulation are still underexplored in Cameroon. This is because most studies have attempted to capture within and between group effects of inequality; a mode of analysis increasingly being labelled as a mere accounting exercise.

The theme and approach pursued in this study presents a double interest: (1) a scientific interest; and (2) an economic policy interest. Scientifically, reviewing previous works on Cameroon, one could notice that very few studies applied the regression-based decomposition approach such as Tabi (2009) and Epo et al. (2011), but failed in their estimations to account for potential endogeneity. This renders their efforts limiting in terms of policy formulation. Equally, the novelty of this work is the use of household consumption per capita and stock of average household education, which is instrumented by distance to school and classroom density – an aspect that is not addressed by studies mentioned in the Cameroon literature. This study, however, seeks to fill this gap by using the instrumental variables econometric model to account for this problem.

In terms of policy, knowing the determinants of wellbeing as well as overall income inequality will make it possible for development efforts to be concentrated on reducing income inequality and enhancing the welfare of the less privileged. Knowledge of the sources of income inequality will therefore help in reducing poverty, because several studies have established that poverty is invariably related to multiple sources of inequality. This, no doubt, would assist policy makers in selecting the best options for ensuring rapid economic growth to meet Vision 2035 of the Cameroon government.

However, regression-based decompositions are increasingly being used in the Cameroon literature, in trying to illuminate on the subject of inequality. Unlike the traditional method and the use of OLS in regression-based approaches, which dominated the evaluation of inequality decomposition before now in the Cameroon literature, the use of the instrumental variables econometrics to capture the determinants of wellbeing and inequality in this study will permit us to generate more robust results. In this context, a key research question arises: which variables explain wellbeing and income inequality in Cameroon? Specifically:

- What are the determinants of household economic wellbeing?
- What are the relative contributions of regressed-income sources in explaining measured income inequality?

Objectives of the study

Main objective

The main objective is to empirically identify the variables that determine wellbeing and account for income inequality.

Specific objectives

- To evaluate the determinants of household economic wellbeing.
- To assess the relative contributions of regressed-income sources in explaining measured income inequality.
- To formulate policy implications of the findings.

These objectives are guided by two hypotheses:

- Other things being equal, human capital endowments are prominent in explaining household economic welfare.
- Regressed sources for education, credit and rural dwelling dominate in accounting for measured household inequality.

Background of the study

In Cameroon, analysis of indicators of education, health and other socio-economic determinants confirms the existence of rural-urban, regional and ethnic disparities in access to these social infrastructures. The level of attainment in the education sector depends both on the ability of people as well as the availability of educational facilities. Educational infrastructures are reflected in indicators such as number of recognised educational institutions, number of teachers, number of classrooms, etc per unit of the population. It is important to take into account the public provisioning of such facilities. The level of attainment depends not only on the overall level of provisioning but also on its quality of provisioning. For example, the quality of an educational institution might be assessed in terms of availability of basic amenities such as buildings, safe drinking water, proper toilet facilities, electricity, library, etc. Again, the quality of teaching in a school may depend on factors such as the pupil-teacher ratio (PTR), student-classroom ratio (SCR), proportion of professionally qualified teachers, etc. As such, an overly skewed distribution of education tends to have a negative impact on per capita income in most countries (Lopez, Vinod and Yan, 1998).

Table 1: Some key regional primary school education indicators in Cameroon (2008-2009)

Region	Number of pupils (public and private) schools	Number of schools	Number of class rooms	Average number of pupils for 50 seating places	Pupil/ teacher ratio 2006/ 2007
Adamawa	170,021	756	2,719	82	64
Centre	568,560	2,553	14,132	57	43
East	580,604	807	7,648	74	54
Far North	184,223	1,861	3,391	99	72
Littoral	346,974	1,540	9,964	45	36
North	334,591	1,022	4,336	74	79
North West	362,646	1,766	8,827	51	42
West	464,337	1,726	9,706	53	48
South	120,394	808	3,528	58	34
South West	218,312	1,017	5,553	43	41
Cameroun	3,350,662	13,856	69,804	59	48

Source: Statistical Yearbook MINEDUB_2008/2009, NIS (2008)

It is evident from Table 1 that there is much disparity in the distribution of social infrastructure in the different regions in Cameroon. Rural areas are still poverty-stricken, as roughly 90% of their schools have no access to electricity. About 31.8% of primary schools have no access to drinkable water, especially in rural areas (NIS, 2008).

2. Literature review

Since the early works of Roemer (Roemer, 1993; 1998 and Roemer et al., 2003), the concept of inequality has been extended to investigate the effects of opportunities caused by different socio-economic factors beyond individual control. Indicated in the literature, education and health affect the productivity of an individual and therefore earnings, and consequently household economic welfare (Mincer, 1958; Becker, 1964). A large body of empirical research has been conducted on the effect of education on income inequality. For example, Knight and Sabot (1983) observe that there are two effects of educational expansion on income inequality: the composition effect, raising the earnings of those who are more educated tends to increase income inequality, and the wage compression effect, which follows the expansion of the educated labour supply relative to demand, and which tends to decrease income inequality.

Using cross-section data from 59 countries, Park's (1996) econometric results show that a higher level of educational attainment in the labour force has an equalising effect on income distribution. The larger the dispersion of schooling among the labour force, the greater the income inequality. Using the dataset of Deininger and Squire (1996) and the "world income inequality dataset" (WIID), Checchi (2000) finds that average years of education have a strong negative effect on income inequality.

Bourguignon et al. (2007) decompose earnings inequality into a component due to unequal opportunities and a residual term using Brazilian data. Distinguishing between circumstance-based and effort-based variables, they associate inequality of opportunities with the inequality attributable to circumstances that lie beyond the control of the individual – father's and mother's education; father's occupation; race; and region of birth. They interpret their decomposition as establishing a lower bound on the contribution of opportunities to earnings inequality.

According to Gibbons et al. (2005), travel distances have a greater role to play in primary school choice because children of this age are not independent travellers. This means that geographical criteria are likely to be much more relevant in deciding which school to attend, so that the availability of schools can be more confidently inferred from geographical measures of accessibility. This implies that a more equalitarian distribution of education may constitute an efficient means of reducing irregularity of income distribution (Glomm and Ravikurmar, 1992).

Likewise, in the Millennium Development Goals, education is seen as a powerful instrument not only 'for reducing poverty and inequality, but also for improving health and social wellbeing, laying the basis for sustained economic growth, and being essential for building democratic societies and dynamic, globally competitive economies' (United Nations University-UNU, 2000).

Decomposition methodologies have witnessed a lot of evolution in the search for efficient decomposition methods in different strands of literature. This could be seen with the traditional inequality decomposition technique of Shorrocks (1980; 1984), the semi-parametric and non-parametric techniques of DiNardo et al. (1996) and Deaton (1997), to regression-based decomposition techniques of Oaxaca (1973) and Blinder (1973), Fields and Yoo (2000), Morduch and Sicular (2002), and to Wan (2004).

The traditional inequality decomposition gives exposure on the population sub-group decomposition and income source decomposition. While the population sub-group decomposition has been the leading approach to quantifying how the level of education, age cohorts and other household characteristics affect inequality, it then demands the breaking down of sample data into mutually exclusive groups according to one or more category variable (e.g. Urban vs Rural residents, Primary vs. secondary education etc), which will permit the calculation of the within and between inequality of the sub-samples. The inequality decomposition by income source requires an identity to express income as the sum of several components (Shorrocks, 1982). The application of these traditional methods of decomposition techniques by researchers has led to the exposure of its limitations.

The shortcoming of these traditional approaches lies first of all on the restriction of the inequality measure that can be used, which is limited only to the generalised class of entropy measures of inequality based on the population sub-group decomposition. Shorrocks and Wan (2005) point out that the Gini index cannot be used for this purpose unless income from different sources does not overlap at all. Secondly, the decomposition can only be carried out over discrete categories even though some factors such as education and age are more appropriately considered as continuous variables (Morduch and Sicular, 2002). Thirdly, the decomposition results obtained are usually contaminated by other factors, as these methodologies are unable to incorporate control variables (Wan, 2007).

However, the regression-based decomposition technique is welcomed as a solution to these problems. The reliable nature of this decomposition technique is appealing because it overcomes many of the limitations of standard decomposition by groups. This decomposition approach allows for a mix of explanatory variables, which might be discrete, continuous or proxies. For example, continuous variables are permissible, and it is possible to control for endogeneity. As such, the flexibility of this approach, particularly its ability to accommodate endogeneity of income determination and random errors, makes it rather more attractive. To date, work on regression-based decomposition techniques of inequality has been a gradual process, with each proposed approach having different properties and using different inequality indices.

These methods, such as that of Fields and Yoo (2000), and Morduch and Sicular (2002), were later extended by Wan (2004) to enlarge the flexibility and the accommodating characteristics of the regression-based decomposition approach.

The question of flexibility and the accommodating characteristics of regression-based decomposition lie on the restriction on the inequality measures to use and on the model specification. To get rid of these restrictions, Wan (2004) makes an attempt to marry the conventional regression models with the Shapley procedure of Shorrocks (1999). The regression model establishes a relationship between a target variable such as consumption or income and its determinants such as human capital, family characteristics and locality. This is, however, to produce reliable estimates suitable for decomposition. The Shapley procedure relies on the estimated function to attribute inequality in income or consumption to the various determinants.

Relative to earlier methods, the approach in Wan (2004) has a number of advantages. First, it is applicable under any inequality measure. Relative inequality measures such as the Gini and generalised entropy indices or absolute measures such as the Kolm index can all be used. Second, it can control for as many variables as data availability permits, rendering decomposition results more precise and reliable. Third, it does not require a pre-defined identity to express income as a sum of its components, although such an identity can be treated as a special regression model without the residual term. Finally, it imposes few constraints on the regression model. The model can be highly non-linear, can include interactive terms, and can be one of the equations in a simultaneous equation system (Wan, 2007).

All these approaches begin with an income-generating function. The results of the estimation of the income-generating function are then used to quantify the contribution of any number of factors to total inequality. The method suggested by Fields (2003) manipulates the equation so that it can be written in terms of covariance. The contribution of the independent variables to distributional changes is then expressed as a function of the size of the coefficient of the income equation and the magnitude of the change in the variable relative to the variation in income. In the Morduch and Sicular (2002) method, the resulting coefficients are regarded as estimates of the income flow attributed to household variables. This permits the application of decomposition by income sources or factor income to apportion inequality to any number of explanatory variables. This is in contrast to the method proposed by Bourguignon et al. (2001), which can be used to decompose differences in income distribution to just three broad components: price effects, participation effects, and population effects.

However, the Morduch and Sicular method has been criticised. Though the methodology requires the inclusion of an error term into the original income generating equation, it does not make any contribution towards overall inequality (Wan 2004: 352). According to Wan (2004), the value added of including this term in decomposition analysis is that it indicates the proportion of the contribution of sources which are not captured by the income-generating function when explaining inequality. In contrast, Field's decomposition methodology accounts for the contribution of the regression error to total inequality, but this tends to be large, leaving unexplained a major proportion of inequality. Further, Wan (2002) observes that neither method accounts for the contribution of the constant term to total inequality.

3. Methodology

The methodology used for this study consists of the use of instrumental variables econometric model to provide estimates that are then used in the decomposition exercise through the method inspired by Wan (2004). Wan's (2004) extension of the decomposition technique permits us to assess the contribution of specific factors, the constant and the residuals to total inequality, where the amount explained by each factor is dependent on the inequality measure used.

Regression model

In order to address the objectives of this study, it is necessary to estimate reliable parameters that will be suitable for the inequality decomposition procedure. To motivate the need for the method of instrumental variables, consider a linear population model presented as follows:

$$Y = Z_1\vartheta_y + \alpha H\mathrm{E} + \mu \qquad (1)$$

where Y is the outcome variable (economic wellbeing, surrogated by household consumption expenditure per capita), and HE the endogenous determinant of wellbeing (education) which might be correlated with μ. Z_1 is a vector of exogenous variables included in the model; ϑ_y is a vector of parameters comprising the constant term and those of the exogenous explanatory variables that correlate with the income generating function to be estimated. α is the parameter representing the coefficient of the potential endogenous variable (education) in the wellbeing function. μ represents the error term. $E(\mu) = 0$ and $Cov(Z_1, \mu) = 0$.

The method of instrumental variables (IV) provides a general solution to the problem of endogeneity (joint determination of response variable and regressors), omitted variable bias, and error in variable (measurement error in the regressors). Given that education is endogenous, the direct estimation of its coefficient – α to show its effect on the household economic wellbeing function – might likely produce bias and inconsistent estimates using ordinary least squares (OLS) method of estimation. As such, to generate consistent estimates that will express the effect of education on the household wellbeing function, potential instruments will be used. To situate ideas, it might be helpful to think of μ as containing omitted variables that are uncorrelated with all explanatory variables except HE. In other words, these instruments must satisfy two conditions; the first being that the instruments must not correlate with the error term in the household welfare function and, secondly, the instruments must be partially correlated with the endogenous variable (education). A precise statement of the reduced form equation of household demand for education requires the linear projection of on all the exogenous variables, including instruments.

$$H\mathrm{E} = \mathrm{Z}\vartheta_{he} + \mu_2 \tag{3}$$

whereby the definition of linear projection error, $\mathrm{E}(\mu_2) = 0$ and μ_2 is uncorrelated with Z, where HE is household education. Z represents a vector of exogenous variable which is made up of Z_1 covariates that belong to the outcome equation and Z_2 a vector of instrumental variables that affects the endogenous input (education), but have no direct influence on the household wellbeing function, Y. ϑ_{he} represents a vector of parameters of exogenous explanatory variables and those of instruments in the reduced form equation to be estimated, while μ_2 is the error term.

However, as long as the standard assumption is that there are no exact linear dependencies among the exogenous variables, we can consistently estimate the parameters in the reduced form equation by OLS, presented as follows:

$$\widehat{H}\mathrm{E} = \mathrm{Z}\hat{\vartheta}_{he} \tag{4}$$

Where $\widehat{H}\mathrm{E}$ is the estimated value of household demand for education (HE), in the wellbeing function, determined by $\mathrm{Z}\hat{\vartheta}_{he}$ in the reduced form equation, accounting for the two stage least squares (2SLS) estimator econometric method. As such, for each observation i, define the vector $\hat{\mathrm{X}}_i = (1, z_{i1}, \ldots z_{i10}, \widehat{H}\mathrm{E})$. Where i = 1,2,...N with N representing the total number of household visited, $Z_1 = (z_{i1}, z_{i2}, \ldots z_{i10})$, which captures the vector of exogenous variables included in the outcome equation. Substituting $\widehat{H}\mathrm{E}$ in equation (1) for the endogenous variable (education) gives the IV estimator.

Decomposition of income inequality

The next question we attempt to answer is how much income inequality is accounted for by each explanatory factor, and how much is unexplained, as gauged by the residual. Before now, the best-known and most general formula for decomposing inequality into its component sources is due to Shorrocks (1982), who established a measure of inequality which pertains to inequality indices that can be written as a weighted sum of income sources:

$$I(y) = \sum_{i=1}^{N} a_i(y) y_i \tag{5}$$

Where y_i is the income of household i, N is the number of households in the population, $y = (y_1, \ldots, y_N)$ and $a_i(y)$ is a weighting factor. $I(y)$ is the weighted sum of total household income corresponding to an inequality measure. Each y_i may be observed as the sum of component incomes y_i^{K} coming from K different sources or endowments, which could be written as; $y_i = \sum_{k=1}^{K} y_i^{\,k}$. As such, the proportional contribution of source k to overall inequality can be defined as:

$$S^{\mathrm{K}} = \frac{\sum_{i=1}^{N} a_i(y) y_i^{\mathrm{K}}}{I(y)} \tag{6}$$

However, Morduch and Sicular (2002) extended this decomposition rule (6) of Shorrock to a regression-based decomposition by determinants of household income as:

$$y = \mathrm{X}\beta + \varepsilon \tag{7}$$

Where X is a nxM matrix of explanatory variables (such as education, household size, residence, etc), β is an M-vector of coefficients, and ` is an n-vector of residuals. The vector β can be defined as the effects of the independent variables on income. Given the vector of consistently estimated parameter $\hat{\beta}$, income can be expressed as a sum of predicted income and predicted errors as in equation (8), and considered as the estimated income source flows of the various explanatory variables.

$$y = \mathrm{X}\hat{\beta} + \hat{\varepsilon} \tag{8}$$

y is considered a vector of economic wellbeing as measured by the log of household income expenditure per capita, which is to be accounted for by the vector X, a vector of estimated coefficient and the error term. A sample of observations can be used to estimate the model. Using equation (8), per capita income of household i can then be presented as follows:

$$y_i = \sum_{m=1}^{M} \hat{\beta}_m x_i^{\,m} + \hat{\varepsilon} \tag{9}$$

Where is the coefficient estimates and is the residual for household i. These estimated income flows can then be used directly to calculate decomposition components for all regression variables. Following the analogy from equation (6), the shares can take the form:

$$S^m = \hat{\beta}_m \left(\frac{\sum_{i=1}^{n} a_i\left(y\right) x_i^{\,m}}{I\left(y\right)} \right) \tag{10}$$

This decomposition rule can be applied to any inequality index that can be written as a weighted sum of incomes. Another attractive feature of the regression-based method is that it allows for the incorporation of relevant dummies in the expenditure regression to capture the effects of different geographical regions and other ordinal variables such as gender, while controlling for other factors that affect consumption expenditure.

However, Wan (2004) extended this method to account for the contribution of the intercept of the income regression to inequality. Wan (2002; 2004) remarked that most income inequality regression decomposition usually ignores or treats incorrectly the constant term, for example Fields and Yoo (2000). Meanwhile, Morduch and Sicular (2002) did not take up the issues of the constant term and error term of the regression model used. According to Wan (2002), ignoring the residual terms means throwing away useful information on non-included determinants of income or income distribution, which could distort decomposition results. A constant source of income is widely known to either lower the level of inequality when it is positive or raise the level of inequality when it is negative.

Adopting the potent extension of this procedure proposed by Wan (2004), free from the limitations of the traditional methods, the income inequality accounted for by each explanatory factor, constant term and how much of total inequality is gauged by the residual term can be assessed. Drawing from the simple, yet powerful procedure proposed by Wan (2002) for regression-based inequality decomposition, which is free from the pitfalls and limitations discussed above, the estimated income generating function can be presented as follows:

$$y = F\left(\mathrm{X}\right) + \varepsilon = \alpha + y^{*}\left(\mathrm{X}\right) + \varepsilon \tag{11}$$

where y is the income generating function or its transformation, and X is income determinants or their transformation. α is the constant term, ε is the error term, and $y^*(\mathrm{X})$ is the estimated income sources. $F(\mathrm{X})$ allows for any functional form; be it linear with a constant term or non-linear with the absence of this term.

Let the estimated part of equation (11) be written as shown below in equation 12, where $\hat{y}$ is the deterministic part of the estimated income generating function, $\hat{\alpha}$ is the estimated constant and $\hat{\varepsilon}$ is the estimated residual, while $\ddot{y}$ is a vector of estimated income sources excluding the constant and the estimated residual term.

$$y = \hat{y} + \hat{\varepsilon} = \hat{\alpha} + \ddot{y} + \hat{\varepsilon} \tag{12}$$

Following the general decomposition framework and using $I(\bullet)$ as an indicator for an inequality measure, the inequality from the deterministic part can be measured as:

$$I(\hat{y} / \alpha, x_1, x_2, ..., x_m) = \Gamma(\alpha, I) + \Gamma(x_1, I) + \Gamma(x_2, I) + ... + \Gamma(x_m, I) \tag{13}$$

The measured inequality from the predicted income values excluding the constant and the estimated residual term is given as follows:

$$I(\ddot{y} / x_1, x_2, ..., x_m) = \Gamma(x_1, I) + \Gamma(x_2, I) + ... + \Gamma(x_m, I) \tag{14}$$

The rationale behind the Shapley approach is based on the fact that the contribution of a single factor can be assessed as the difference between the overall income inequality and the inequality that would be observed should that factor be removed from the set of income determinants. As a consequence, the marginal impact of each factor $\Gamma(X_j, I)$, for $j = 1, 2, 3, ...m$ is calculated through the estimation of a sequence of regression models starting from the specification, which includes all the regressors, and then successively eliminating each of them. The overall marginal contribution of each variable is then obtained as the average of its marginal effects. Since the contribution of any factor depends on the order in which the factors appear in the elimination sequence, this average is calculated over all the possible elimination sequences.

The contribution $\Gamma(X_j, I)$ of the factor X_j to the explanation of the inequality measure is given by the following formula:

$$\Gamma(X_j, I) = \frac{1}{M!} \sum_{\pi \in \Pi_m} \left[I(\ddot{y} / \beta(\pi, X_j) U \{X_j\}) - I\left(\ddot{y} / \beta\left(\pi, X_j\right)\right) \right] \tag{15}$$

Where $I\left(\ddot{y}/X\right)$ is the inequality indicator of the vector of explanatory variables X Π_m, is the set of possible permutations of the m variables and $\beta\left(\Pi, X_j\right)$ represents the set of the variables preceding X_j in the given ordering Π.

The calculation of each factor contribution requires the estimation of 2^{m+2} income-generating models.

The total measured inequality can be broken down into the contribution of the constant term $I\left(\hat{\alpha}\right)$, the contribution of the estimated income sources $I\left(\ddot{y}\right)$ and the contribution of the predicted residual $I\left(\hat{\varepsilon}\right)$ as presented in equation 16 below:

$$I\left(y\right)=I\left(\hat{\alpha}\right)+I\left(\ddot{y}\right)+I\left(\hat{\varepsilon}\right) \tag{16}$$

Following Wan (2002), the contribution of the unexplained inequality $I\left(\hat{\varepsilon}\right)$ can be computed as the difference between the observed inequality $I\left(y\right)$ and the inequality from the deterministic part $I\left(\hat{y}\right)$ as follows:

$$I\left(\hat{\varepsilon}\right)=I\left(y\right)-I\left(\hat{y}\right) \tag{17}$$

The difference between $I\left(y\right)$ and $I\left(\hat{y}\right)$ is subtle and important. This is simply the case of the expected values of and ; they might be identical. The ranking between y and $\hat{y}$ differ and will be equivalent if and only if there is a good enough fit of the income-generating function. Looking at it from this perspective, the decomposition makes intuitive as well as theoretical sense. Decomposing equation (11) entails that the disturbance term is irrelevant and does not affect income inequality. This is not true because from the previous discussion, one should note that $I\left(y\right)\neq I\left(\hat{y}\right)$ unless $\varepsilon=0$.

One way to treat this residual term is to discard it altogether because the residuals are not explainable by the structural income generating function. If this is the case, one could focus on $\hat{y}$ and obtain further decomposition results. This, however, is not recommended. The residual term is to some extent viewed as representing factors or determinants other than those included in the regression model. Ignoring ε is certainly unwise as it does contain useful information. It therefore implies that once its contribution is identified, policy makers and others could be informed as to how much included factors can explain inequality.

Having identified the contribution of the residual term, the contribution of the constant term can be computed as the difference between the inequality from the deterministic part $I\left(\hat{y}\right)$ and the measured inequality from the predicted income values $I\left(\ddot{y}\right)$, excluding the constant and the predicted residual term as follows:

$$I\left(\hat{\alpha}\right)=I\left(\hat{y}\right)-I\left(\ddot{y}\right) \tag{18}$$

where all the contributions are simply attributed to the estimated income factors used in the decomposition process. To summarise, $I(y)$ can be decomposed into $I(\alpha)$, $I(\varepsilon)$ and $I(\ddot{y})$, which represent the estimated factor sources, as well as their percentage contributions, which add up to 100%.

Operationalising this decomposition, Arrar and Duclos (2009), following the uniform addition axiom of decomposition, the total income can be expressed in terms of regressed income sources and predicted residual as:

$$\begin{aligned} y &= y_0 + y_1 + y_2 + \ldots + y_m + \ldots + y_M + \hat{\varepsilon} \\ \hat{y} &= y_0 + y_1 + y_2 + \ldots + y_m + \ldots + y_M \\ \ddot{y} &= y_1 + y_2 + \ldots + y_m + \ldots + y_M \end{aligned} \tag{19}$$

Based on the semi-log linear functional form of the Shapley approach, the income-generating function takes the form:

$$\log(y) = y_0 + y_1 + y_2 + \ldots + y_m + \ldots + y_M + \hat{\varepsilon} \tag{20}$$

With the specification above,

$$\ddot{y} = Exp(y_0 + y_1 + y_2 + \ldots. + y_m + \ldots y_M + \hat{\varepsilon}) = Exp(y_0) * \prod_{m=1}^{M} Exp(y_m) * Exp(\hat{\varepsilon}) \tag{21}$$

From the above expression, it is clear that adding a constant leaves the measured inequality unchanged, since we are simply multiplying by a scalar. In this context, the constant component does not explain differences in inequality, and its concentration coefficient equals zero using the semi-log income-generating function.

4. Data and model identification strategy

Data

This study uses the 2007 Cameroon household consumption survey (ECAM 3), which is the most recent data collected by the National Institute of Statistics - NIS in Cameroon. The survey was carried out between May and July 2007. The targeted sample consisted of 12,000 households, of which 11,391 were effectively visited. The survey cuts across the national territory from north to south, west to east, covering the 10 regions of Cameroon. The whole territory was divided into urban, semi-urban, and rural strata, summing up to 32, the strata established for the survey. These comprises of 12 urban, 10 semi-urban and 10 rural strata. The two main cities of Douala and Yaoundé, stand as urban strata made up of 50,000 thousand inhabitants, with each of the 10 regions divided into semi-urban strata made up of at least 10,000 inhabitants and at most 50,000 inhabitants. The rural strata consisted of less than 10,000 inhabitants. The aim of the survey was to upgrade knowledge on poverty and welfare status in Cameroon by providing indicators that capture living standards of the local population and to compute poverty profiles that will provide a follow up of efforts made towards the implementation of the PRSP and the realisation of the MDG objectives (Epo et al., 2011).

This survey can be used to: (1) study all aspects of poverty, both at the national and regional levels (monetary poverty, household poverty, poverty in terms of potentials and subjective poverty), as well as establish correlations between these different types of poverty; (2) study the inter-temporal changes of poverty between 2001 and 2007, with the aim of evaluating the effects of macroeconomic policies of the last five years on household wellbeing; (3) evaluate the demand for education and identify its principal determinants; (4) evaluate internal tourism in Cameroon, and; (5) collect data on child labour in Cameroon (NIS, 2008).

Based on the above mentioned data, the following variables were selected for analysis: Log (total household expenditure per capita), as a surrogate for wellbeing and; Education, measured as average years of formal education of the household members. The instrumental variables for education are the distance to school captured at cluster level and classroom density. Other variables are: Age stock, measured as the average of the ages of the household members; Dependency, measured as the dependent population per household of ages less than 10 and more than 65 and; Credit access (have access =1 and 0 otherwise). Farmland ownership indicates households in which at least a member owns exploitable farmland and most farmland is inherited or owned communally in rural areas (own farmland = 1 and 0 otherwise); and Bilingual, are those who can read and write in French and English.

The growing focus on geographic inequality to find the nexus of location and living standards is based on zone of residence (rural =1 and 0 otherwise), electricity (electricity access =1 and 0 otherwise), and finally radio (own radio =1 and 0 otherwise). The variables credit, and radio, are captured at cluster level and expressed as cluster means, with the idea that a given household cannot influence a societal variable (community variable), being that a household without a radio can still listen to information from neighbouring households. Thus, considering the cluster mean in each primary sampling unit permits the reduction of potential endogeneity (Baye and Epo, 2011).

Model identification strategy

To address the identification strategy, it is useful to first consider the potential sources of endogeneity of education in the household wellbeing generating function. We can refer to three possible sources of endogeneity of education.

First, we measure household stock of education as the average years of schooling of the members, which captures both children and parental years of education. In this context, education is at least weakly endogenous. Besides, when parents decide as to whether or not their children will attend school, the choice is largely about income allocation. A particular household's demand for education depends not only upon their preferences, but also upon their financial situation. It may be to the economic advantage of the family that the child allocates his or her time to work either in the household or in the labour market. As such, the decision to send their children to school will influence current consumption as well as the production decisions of a household, thus the use of average household education, given that current consumption is not only determined by parental education.

Moreover, children of poorly educated parentage with some education can make their parents to be more innovative by helping to interpret farming and health-related instructions. Educated children are more likely to perform chores that contribute in valorising household production and consumption activities better than their non-educated counterparts.

Secondly, the liberalisation of education since the 1990s in Cameroon implies that the acquisition of more years of schooling is an ongoing process. This means that education attainment remains a choice even among workers. A worker can exploit the opportunity to acquire new skills that can make her more competitive in the labour market by being more eligible for promotion or increased job mobility in order to earn more and enhance household wellbeing. Thirdly, the full effect of education is not captured by years of education. Since years of education do not capture ability, it implies that ability is embedded in the error term of the structural equation. This source of endogeneity of education implies that the OLS estimates are biased, and the biasness does not disappear even if the sample size is increased indefinitely. In other words, the OLS estimates are biased and inconsistent.

To address this potential endogeneity problem of education, we appeal to the IV estimation approach that requires a credible identification strategy that respects the exclusion restrictions. In this context, the variables chosen as instruments should be uncorrelated with the error term in the wellbeing function (i.e. they should be exogenous), should be correlated with the endogenous variable in the wellbeing function (i.e. they should be relevant, or rather, their effects on the endogenous explanatory variable in the wellbeing function should be statistically significant), and should be excluded from the wellbeing function (Wooldridge, 2002; Murray, 2006; Mwabu, 2009; Brookhart, Rassen and Schneeweiss, 2010).

Following Brookhart, Rassen and Schneeweiss (2010), we propose distance to the nearest school captured at cluster level and classroom density captured at regional level to instrument for education. These instruments are expected to strongly affect education. They should be unrelated to other household characteristics, and should be related to wellbeing function only through their association with education.

The choice of these instruments is based on the assumption that distances to school are correlated to school attendance, which in turn leads to an increase in the stock of household education. Since many households in Cameroon are poor (NIS, 2008), they will hardly have the means of transport to cover long distances for schooling. Thus, the longer the distance to the nearest school, the higher the opportunity cost of an individual going to school and vice versa.

Secondly, the choice of these instruments overrules the argument in the literature against the use of school location and school characteristics as instruments, given that the choice of location of residence by parents makes the usual instrumental variables of school location and school characteristics endogenous, such that they could no longer be used as instrumental variables. This argument is overruled based on the idea that a given household cannot influence a societal variable (community variable). Better still, even though an individual household can choose to live near a school because of education for both children and parents, all the households in a region are unlikely to make this decision simultaneously every time.

Thirdly, since these instruments are more public policy determined, the shorter the distance to school the higher the accessibility. Also, high classroom density in a region provides an enabling environment for quality education, which leads to higher stock of education in households. Moreover, following the liberalisation of professional training, many professional trainers provide evening classes in primary school premises to permit workers to attend and improve on their skills and qualifications. In this context, proximity and density of classrooms captured at cluster and regional levels will likely fulfill the exclusion restrictions, and are thus relevant and valid for use in this study.

5. Empirical results

This section discusses the descriptive statistics, the regression results and the results of the regression-based inequality decomposition.

Descriptive statistics

Table 2 presents the summary of descriptive basic characteristics used in the study using the Cameroon household consumption survey of 2007. The survey indicates an average stock of education of about 5 years of study per household and an average age stock of 22 years, ranging from a minimum of 7 years to a maximum of 95 years. The statistics equally reveal that Cameroon is a country with a relatively large family size of over 6 persons per household, with 65% of the households interviewed residing in rural areas. Statistics indicate an average dependency rate of 34% per household, which implies that there is a high dependent population. Concerning access to credit, an average of 5.8% of households had access to credit, indicating that over 94.2% had no access to credit. On average, about 64% of households own a radio, while about 24% of household heads were bilingual – that is, can relate in both French and English. About 61% of the households had access to land – mainly communally in rural areas. Distance to the nearest public and private primary school from the homestead ranges from 0.5 kilometres to about 6 kilometres, with an average distance of about two kilometres. The number of classrooms per region ranges from 2,719 to 14,132 classrooms with an average number of classrooms of 7,770. These characteristics provide the context within which wellbeing as well as income inequality are discussed in this study.

Table 2: Descriptive statistics for regression analysis

Variables	Weight	Mean	Std. Dev.	Min	Max
Endogenous Variables					
Log of total expenditure per head	17,878,688	12.427	0.6914	11.1851	16.243
Education stock	17,878,688	5.159	3.551	0	20
Exogenous Variables					
Age stock	17,878,688	22.053	9.583	7	95
Household size	17,878,688	6.476	3.9869	1	43
Dependency	17,878,688	0.336	0.215	0	1

continued next page

Table 2 Continued

Variables	Weight	Mean	Std. Dev.	Min	Max
Exogenous Variables					
Access to credit (cluster mean)	17,878,688	0.058	0.090	0	0.647
Rural	17,878,688	0.647	0.478	0	1
Electricity	17,878,688	0.900	0.300	0	1
Land	17,878,688	0.608	0.488	0	1
Radio (cluster mean)	17,878,688	0.675	0.254	0	1
Bilingual	17,878,688	0.244	0.429	0	1
Instruments used for Education					
Classroom density	17,878,688	7770.11	3878.62	2719	14132
Distance to school (cluster mean)	17,878,688	2.167	1.337	0.5	5.941
Total number of observations	11391				

Source: Computed by author using CHCS III (2007) and STATA 10.

Regression results

Table 3 shows the regression results capturing the effect of the explanatory variables on wellbeing. It presents the OLS results (column 1), the reduced form (column 2), the two-stage least squares (column 3) and the two-stage using the residual addition approach (column 4). The selected variables give the estimated structural equations a global significance with an R-square of 0.50. It is equally observed that all of the effects are highly statistically significant. Comparing the models, the results indicate that the instrumental variable two-stage least square estimates have an edge over the ordinary least squares results because they are more robust after accounting for the potential endogeneity bias.

Table 3, column 3 presents the test on the relevance, strength and exogeneity of instruments. Following Shea (1997), the first-stage F-statistic and the partial R-square convey vital information as to the relevance and strength of instruments in the case of a single endogenous variable. The strength of the excluded instrument is also assessed by the Cragg–Donald F-statistic (Stock and Yogo, 2005) which stood at 1,210.65 greater than the Stock-Yogo weak ID test critical values: 10% maximal IV relative bias of 19.93. The (Durbin-Wu-Hausman Chi-square Statistic = 10.132, p-value=0.00) test of the endogenous regressor shows that stock of education is indeed endogenous, which indicates that the IV estimates are preferred to OLS estimates, which are not reliable for inference. Meanwhile, the magnitude of the coefficient of education stock and its t-statistics are lower than those of the OLS, an indication that the OLS estimates may be over-estimating the effect of education on household wellbeing.

A closer look at Table 3 reveals that the following variables: education, age, credit, bilingual, radio, and electricity are positively associated with household welfare. The importance and linkage of education to the development and growth of any society and welfare are well-known in most literature. This is evident in the Cameroon setting, where the level of education is considered a signal for high performance; be it in the public or private sector. Table 3 (column 3) reveals that education is directly proportional to wellbeing. Thus, increasing the household stock of education is consistent with improvements in household welfare.

Access to better education typically facilitates vertical social mobility by breaking barriers and enhancing the bargaining power in the labour market. This leads to exposure to high paying job opportunities and income earning capacity of household members, and hence higher wellbeing. This result corroborates that obtained by Baye and Epo (2011) for Cameroon, Awoyemi and Adekanye (2003), and Oyekale (2007) for Nigeria, and Maria and Jose (2008) for Cape Verde with an added value being the use of distance to school and classroom density as instruments for the stock of household education.

In today's global economy, the ability to communicate is key, and as more companies expand internationally, the ability to communicate in another language has become a significant advantage in the workforce. As such, in a country like Cameroon with two official languages (French and English), being bilingual exposes an individual to higher paying job opportunities (both in the public and private sector) and a higher income-earning capacity of a household, hence higher wellbeing.

Having access to credit means having more financial leverage to embark on relatively larger scale economic activities that may enhance profits and welfare in a household. These results are consistent with findings by Hao (2005) who, using Vietnamese data, found that credit contributes positively and significantly to the economic welfare of households. There appears to be a premium for bilingualism. Being bilingual widens the scope of an individual to access the labour market, thus reflecting an opportunity for more decent employment as well as the potential to generate more income. The age stock of the household has relatively modest positive impact on the level of expenditure.

Access to electricity in a household has both production and consumption value, which improves on household wellbeing. The consumption value can come as a result of the use of electricity to raise household utility through the use of electricity in a household for lightening, for water cooling, for ironing, for water heating, for air conditioner, for washing machines, and so on. The production value could be seen through the use of electricity to improve household income, for example, use of microwave ovens to dry farm products such as cocoa, coffee and others to improve on their quality for higher market value and other agricultural farming equipment to increase output. It equally permits the conservation of present consumption to future dates through the use of appliances such as fridges for conserving perishables for consumption or retailing.

That 65% of the households interviewed reside in rural areas is an indication that most of them depend on agriculture. Radio has a broad reach in Africa as well as in Cameroon than any other media. As such, the use of radio in a household has both consumption and production value. The consumption value could be seen in the direct use of radio for entertainment, whereas the production value could be seen through the use of radio to improve on household income. For example, African Farm Radio Research Initiative (AFRRI) created a series of programmes designed to educate farmers to make informed decisions about adopting modern agricultural techniques, and enable them to improve their agricultural practices and improve food security in Africa. Besides, a radio equally serves as a medium for informing household members on job opportunities. Thus, electricity and radio have production values that generate both market and non-market endowments that have implications for household welfare.

On the other hand, variables such as rural, dependency and land associate negatively with household welfare. The larger the number of dependants in a household, the higher the pressure on meager household income and consequently an overall deterioration in wellbeing. Residing in rural relative to urban areas reduces the welfare of households. This may perhaps reflect the inaccessibility of rural households to markets due to lack of roads and other social infrastructure compared to their urban counterparts, given that they embark mostly on agricultural activities. In terms of land, the negative correlation with wellbeing may be due to the observation that most households do not make meaningful gains from the piece of land they exploit due to lack of credits to indulge in large scale activities or due to lack of incentives to invest in communally-owned land.

Table 3: Determinants of household economic well-being-dependent variable is log of household expenditure per head

	(1)	(2)	(3)
Variables	OLS	Reduced form	One-step 2SLS
Education stock	0.0702*** (0.00199)		0.0565*** (0.00476)
Age stock	0.0131*** (0.000511)	0.00159 (0.00219)	0.0132*** (0.000513)
Dependency	-0.1000*** (0.0270)	-6.667*** (0.0994)	-0.196*** (0.0407)
Credit (cluster mean)	0.393*** (0.0513)	-0.581*** (0.223)	0.386*** (0.0514)
Rural	-0.324*** (0.0123)	-0.614*** (0.0566)	-0.346*** (0.0142)
Electricity	0.0741*** (0.0162)	0.354*** (0.0692)	0.0828*** (0.0164)
Land	-0.185*** (0.0115)	-0.427*** (0.0494)	-0.193*** (0.0118)
Radio (cluster mean)	0.0948*** (0.0194)	0.509*** (0.0847)	0.114*** (0.0204)
Bilingual	0.163*** (0.0128)	2.462*** (0.0497)	0.202*** (0.0178)

continued next page

Table 3 Continued

Variables	(1) OLS	(2) Reduced form	(3) One-step 2SLS
Classroom density		0.000240***	
		(5.94e-06)	
Distance (cluster mean)		-0.403***	
		(0.0188)	
Constant	11.94***	5.801***	12.03***
	(0.0282)	(0.115)	(0.0401)
Observations	11,391	11,391	11,391
R-squared	0.502	0.655	0.500
Fisher test [p-value]		2159.56; [0.000]	1147.94; [0.00]
Partial R-square for education		0.1754	
Test of excluded instruments: F-stat[p-value]		1210.65; [0.000]	
Test of joint significance of identifying variables/Cragg Donald weak identification test			
F-statistic [10% relative bias]			1210.65 (19.93)
Under-identification tests (Anderson Canon corr. LR statistics)			
Chi- Sq [p-value]			2197.385; [0.00]
Sargan statistics (Over-identification test of all instruments)			
Chi- Sq [p-value]			0.069; [0.7924]
Endogeneity test of endogenous regressors			
Chi- Sq [p-value]			10.132 ; [0.00]

Source: Computed by author using STATA 10. Standard errors in parentheses
Significant levels respectively. *** p<0.01, ** p<0.05, * p<0.1

Regression-based inequality decomposition results

Based on the regression results obtained for this study, the Shapley approach was used to decompose income inequality based on the linear income generating function and the semi-log income generating function to quantify the importance of the various income components in explaining measured income inequality in Cameroon. The results were computed using STATA 10 and the DASP 2.1 Software developed by Araar and Duclos (2009). The Gini index was used as the measure of income inequality with a total contribution of different income sources to overall income inequality of 0.4077.

According to Shorrocks (1999), the Shapley Value approach based on a set of axioms can be used to inform the computation of the weighted marginal contributions of an estimated income source in various coalitions of income sources, with the weighted contributions summing exactly to the considered inequality measure.

Table 4, column 2, shows the income shares, while columns 3 and 4 present the decomposition results based on the linear income generating function and the semi-log income generating function to quantify the importance of the various income components in explaining measured income inequality in Cameroon. In column 2 of Table 4, putting aside the constant term, the estimated income sources for education, ages, bilingual, electricity and credit registered positive income shares, while rural, land, radio and dependency registered negative income shares.

Based on Table 4 column 4 for the semi-log functional form, all the included consumption sources contributed positively while the constant term mark a zero contribution to measured income inequality. The contribution of the constant term to total inequality depends on two things: the functional form (linear, semi-log, and double-log) and the sign of the estimated constant term. Such a special source is factually known to lower (raise) total inequality if it is positive (negative) for linear income-generating function using the Shapley value decomposition. Based on the logarithmic functional forms using the Shapley value decomposition approach, the contribution of the entire included source to measured inequality is always positive while the contribution of the constant term to measured inequality is always zero. This, however, is based on the construction of proposed decomposition methodologies (Wan, 2002).

The 0 contribution from the constant term is acceptable because, with this particular income generating function (semi-log using the Shapley value), the constant term acts as a multiplicative scalar of total income for all income recipients (equation 21). Thus, the presence or absence of the constant term should not matter at all, and this is exactly what is given by the proposed procedure. In passing, it is noted that the Shapley value decomposition will always attribute 0 contributions to constant terms in regression models – a potential drawback of the semi-log functional form using Shapley value approach (Wan, 2002). Given the above potential drawbacks of the semi-log functional form using the Shapley value, the interpretation of the different sources, as well as their marginal contributions to measured inequality will be based on the linear functional form.

A glance at column 3 of Table 4 shows that both the absolute and relative contributions to inequality of rural, credit, bilingual and education, respectively, in that order are accounting for measured income inequality. Other sources that contributed marginally in explaining measured inequality are dependency, age and land. It is becoming clearer that the influence of location could be at the root of inequality in Cameroon. Our results show that rural residency accounts for up to about 40.4% of measured inequality. This underscores the need to seek ways and means of equalising economic and political opportunities between urban and rural areas.

There is no doubt that the estimated income source for credit accounts for up to about 22.3% of measured inequality because the descriptive statistics reveal that only about 6% of households had access to credit. Having access to credit means having more financial leverage to embark on relatively larger scale economic activities that may enhance profits and consequently household welfare. Since this access is not enjoyed by all the households, credit tends to be an important source of the gap in wellbeing between households endowed with this attribute and those that are not.

Being bilingual has some premium as it significantly explains about 19.5% of the measured inequality. The inequality augmenting effect of bilingualism could emanate from the observation that those who are bilingual have more access to the national institutions than their non-bilingual counterparts. Bilingual individuals have the opportunity to engage and be a part of two different and diverse communities without feeling excluded. As such, in a country such as Cameroon with two official languages (French and English) being bilingual, exposes an individual to more income-earning opportunities (both in the public and private sectors) than their non-bilingual counterparts. This is reflected in the gaps in wellbeing between those households endowed with this attribute and those that do not have it. Concerning our right hand-side variable of interest, it stands out clear that over 18% of the total inequality in consumption expenditure is explained by education. It is evident that a substantial share of measured inequality could be attributed to correlates of education. The longer the distance to the nearest school, the lower the affinity for schooling and consequently a low household education stock and vice versa. Also, the differences in classroom density in the regions may engender inequality in education, given that limited number of classrooms in a region may mean overcrowding and poor achievements in school or limited pupils' intake in schools. These disparities in access to education indicate the important role education plays in enhancing wellbeing and exacerbating inequality. This result is qualitatively consistent with the findings by Oyekale et al. (2007) and Baye and Epo (2011).

According to the classical human capital theory, more educated workers have a higher labour-force participation rate and work longer hours, are more prone to mobility in the labour markets to seek for better jobs, are more likely to enter the state-owned or monopoly sector to earn high wages, and have a higher probability of escaping from illness and maintaining a stable income level. As such, disparities in access to schooling infrastructure and education quality leads to differences in the ability to earn income and consequently disparities in consumption expenditure.

The average age of household members accounts for about 4% of the measured inequality in this study. This could be visible in the Cameroon employed workforce, which is dominated by an ageing adult population, with youths largely unemployed or under-employed. Farmland ownership accounts for about 2% of household income inequality. The inequality from farmland may emanate from disparities in farming techniques used by the different households, given that most of them do not have access to credit to carry out agriculture on a large scale, leaving most of their farmland unexploited. The estimated source for dependency accounts for measured inequality with a relative contribution of 2.5%. This could be possible given that the mean dependency ratio is about 33%. This finding could be self-informing that a route to improve wellbeing and inequality in Cameroon could be through effective and efficient family planning programmes. The residual term was inequality increasing. This informs policy makers on how much variables not captured by our income-generating function would account to measured inequality.

While the residual term and the above mentioned sources were inequality-increasing, the constant term, radio and electricity, were inequality-reducing. There is no doubt that the constant term is inequality-reducing, given that the more one has of a good, the less each additional unit of good tends to be worth to that person. Maximising utility tends to maximise equality as well. Consider a pauper and a millionaire: if the two were to gain an additional thousand dollars, it would probably make a significant difference to the quality of life of the pauper than that of the millionaire due to diminishing marginal utility. Thus, maximising utility tends to maximise equality.

Table 4: Decomposition of total inequality by estimated income sources using the Shapley Value Approach

Income Sources	Gini Index		
	Income Shares (1)	Linear Equation Absolute [Relative] Contributions (2)	Semi-log Equation Absolute [Relative] Contributions (2)
Education stock	0.526267	0.0743 [0.1823]	0.0611 [0.1498]
Age stock	0.362831	0.0163 [0.0400]	0.0252 [0.0618]
Dependency	-0.00437	0.0104 [0.0255]	0.0110 [0.0270]
Credit (cluster mean)	0.017760	0.0910 [0.2231]	0.0027 [0.0067]
Rural	-0.21913	0.1647 [0.4039]	0.0524 [0.1288]
Electricity	0.022717	-0.0016 [-0.004]	0.0027 [0.0066]
Land	-0.08586	0.0077 [0.0190]	0.0256 [0.0629]
Radio (cluster mean)	-0.02998	-0.0094 [-0.023]	0.0048 [0.0117]
Bilingual	0.051582	0.0794 [0.1947]	0.0226 [0.0555]
Residual	0.000000	0.0767 [0.1882]	0.1995 [0.4894]
Constant	0.358194	-0.1018 [-0.2496]	0.0000 [0.0000]
Total	0.000000	0.4077 [1.0000]	0.4077 [1.0000]

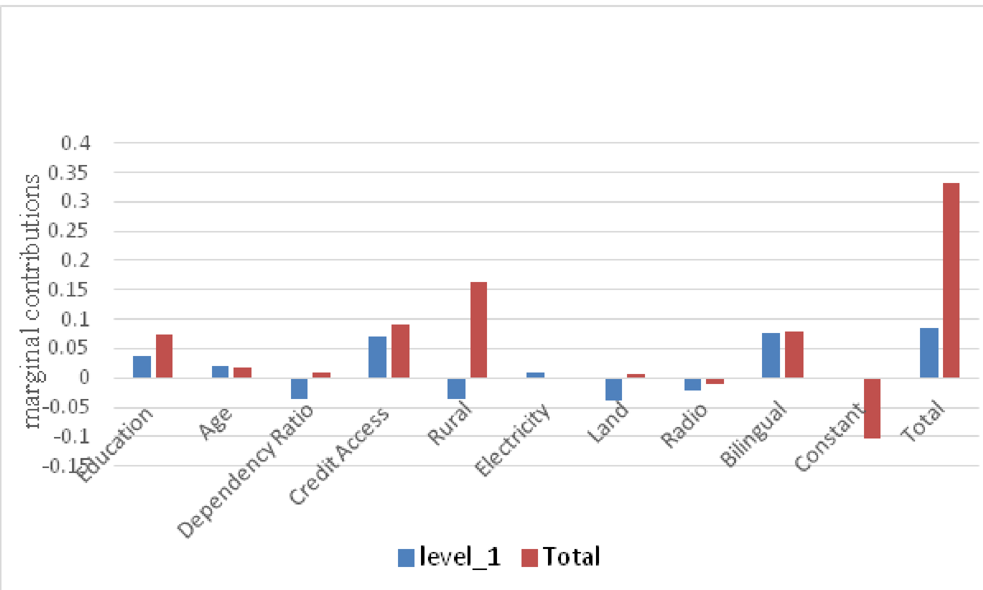
Source: Computed by authors using STATA 10 and the DASP 2.1 Software developed by Araar and Duclos (2009). Relative contributions in parentheses

Radio and electricity are inequality-reducing, given that from the descriptive statistics, an average of about 68% of households had access to radio and about 90% had access to electricity. This is an indication that the consumption and production value of radio and electricity enjoyed by most of the households is inequality-reducing. From the above results, and using the Gini index, the contribution of the observed factors and the constant term account for more than 80% of the observed inequality.

Appendix Table 1 shows the marginal contributions using the Shapley value for the linear functional form, presented in the Appendix Table 1. Meanwhile, Figure 1 presents the initial inequality of the different variables (level 1) and the overall inequality of each variable. This is, however, to give a quick view of the initial situation of the different variables before and after joining the various coalitions in the Shapley value set up. As indicated in the methodology, the Shapley value is an allocation method that assigns the gains of a coalition of sources among other sources as a function of what they contribute to the coalition. As the contribution of a source also depends on the order in which the source joins the coalition, the Shapley rule weights each possible coalition by its probability and assigns to every source the average of all marginal contributions that this source can make to all coalitions.

Level 1 of Appendix Table 1 gives the marginal contributions of each variable in the absence of other regressed-income sources. Attention to level one is perhaps more interesting because the other levels remain a black box, since we do not know members of the coalition after level one, being joint by a source . Sources that are inequality-increasing before being accompanied by other sources through level two are: education, age, credit, electricity and bilingual, while dependency, rural, land and radio are inequality-reducing at level one. For instance, of the weighted mean of marginal contributions of the variable, education, of about 0.0743 to measured income inequality of 0.4077, about 0.0387 is realised at level 1; that is, in the absence of other regressed-income sources. Accompanying with other sources at level 2, 3, 4 and 5, education remains inequality-increasing at a decreasing rate. From level 6 to 10, it becomes inequality-reducing though the total net effect remains positive. This is an indication that much work still needs to be done to equalise the effect of education on welfare.

Figure 1: Initial and overall marginal contributions to measured inequality



The variable, rural, recorded the highest positive contribution to measured inequality but had a reducing effect at level 1 and 2 and inequality-increasing in coalitions across level 3 to 10. This shows the magnitude of the gap that exists between rural and urban dwellers. The variable dependency is equally inequality-increasing through level 2 to 10 in all coalitions with a net positive effect of 0.0104 to overall inequality. This, therefore, could be an indication that there is a large dependent population in most Cameroon households. Land, on its part, is inequality-reducing only at levels 1, 2 and 3 and becomes inequality-increasing through level 4 to 10.

Bilingual and credit registered the highest increasing marginal contributions to inequality at level 1 and were inequality reducing when combined with other sources through level 3 to 10 for credit, with a positive net effect of 0.0909 to measured inequality, and at levels 3, 6, 7 and 8 for bilingual with a positive net effect of 0.0794. This is an indication that these sources still appeal for equalising policies to make them inequality-reducing. Age, on its part, had increasing positive marginal effects from level 1 and negative contributions when merged with other sources from level 6 to 10 with a net positive contribution to inequality. This implies that paying some attention on the labour market age-structure could unleash its inequality-reducing potential.

The variables that had a net reducing effect on observed inequality are electricity, radio and the constant term. Other things being equal, in the absence of other regressed-income sources, the contribution of the constant term at level 1 is zero, which is evident in real life. As the constant term joins other sources through level 2 to 10, it becomes inequality reducing with a heavy net effect of -0.1018. This is an indication that policies that favour income sources with characteristics of non-excludability and non-rivalry should be encouraged. For example, the polio campaign over the national territory is a recommendable policy for inequality reduction.

6. Conclusions and policy recommendations

This study sets out to estimate the determinants of household economic wellbeing and to evaluate the relative contributions of regressed-income sources in explaining measured inequality. In particular, a regression-based decomposition approach informed by the Shapley value, the instrumental variables econometric method, and the 2007 Cameroon household consumption survey, were used. The Gini index was used as a measure of inequality. Both the STATA 10 and DASP 2.1 packages were used to generate results. To fill the gap identified in the literature, we adopt an estimation strategy that controls for potential endogeneity of education. Results show that education, age, bilingualism, electricity, radio, and credit were positive and strongly significantly associated with wellbeing, whereas land, rural and dependency were significantly and negatively associated to wellbeing.

Results also show that rural, credit, bilingualism, education, age, dependency and land, in that order, were the main contributors to measured income inequality. Meanwhile, the constant term, media and electricity, were inequality-reducing. In particular, our variable of interest, education – a human capital endowment – contributed up to about 18.2% to measured income inequality. This should be viewed as a major role played by human capital endowment in the distribution of household income. The policy implication emanating from this finding is that educational expansion (both quantity and quality) could be achieved by shortening distances to schools and increasing classroom densities. Above all, to ensure adequate returns to investment in education, vocational training and skill development programmes could be integrated to help reduce the rate of dependency, which contributed to about 2.6 percentage points to measured inequality.

Secondly, policies that can increase the accessibility of rural areas could be interesting given that being in rural area contributed about 40.4% to measured inequality in the distribution of household income. This can take the form of the provision of quality infrastructural facilities such as good roads, electricity, and communication networks which have both consumption and production values to households. This will help to attract people to the rural areas and workers to render their services to the needy, as such improving on rural wellbeing. Improving access to credit facilities may also play a key role in enabling the poor to smooth their consumption expenditures, and to finance investments, which could improve productivity in agriculture and other economic activities, given that a greater population resides in the rural areas and depends mostly on agriculture. In such a context, initiatives such as creation of agricultural banks as announced by the Head of State during the agro-pastoral show of 2011 at Ebolowa could be inequality-reducing if they target small and medium-sized farmers.

Being bilingual contributed about 19.5% to measured inequality in the distribution of household income. As such, policies linked to bilingualism could be highly welcome, such as the teaching of both French and English in the national curriculum and the creation of language pilot centres in all the divisional headquarters of the country for workers and job seekers to improve on their language skills.

References

Agion, P. and Bolton, P. 1997. "A theory of trickle-down growth and development". Review of Economic Studies, Vol. 64: 151-172.

Araar, A. 2006. "On the decomposition of the Gini coefficient: An exact approach, with illustration using Cameroonian data". Cahiers de recherché, 0602, CIRPEE.

Araar, A. and Duclos, J. Y. 2009. Distributive Analysis Stata Package (DASP) 2.1 software. University of Laval, CIRPEE and the Poverty Economic and Policy Research Network.

Atkinson, A.B. and F. Bourguignon. 2000. "Introduction: Income distribution and economics". In A.B. Atkinson and F. Bourguignon (eds), Handbook of income distribution, Vol.1, North Holland: Amsterdam.

Awoyemi, T. T. and Adekanye, T. O.. 2003. "Gender analysis of economics efficiency in cassava-based farm holdings in South-Western Nigeria". Journal of Economic and Rural Development. Department of Agricultural Economics, University of Ibadan, pp.45-59.

Baye, F. M. and Fambon, S. 2002. "Decomposition of inequality in the distribution of living standards in Cameroon". African Journal of Economic Policy, Vol. 9, No: 51-75.

Baye M. F. and Epo, N. B. 2009. "Explaining inter-household gender inequality in Cameroon: An Oaxaca-Blinder Approach". Paper presented at the 65th annual congress of the International Institute of Public Finance (IIPF), in Cape Town, South Africa, 13-16 August 2009.

Baye M. F. and Epo, N.B. 2011. "Inequality decomposition by regressed-income sources in Cameroon". Paper prepared for the special IARIW-SSA conference on measuring national income, wealth, poverty, and inequality in African countries. Cape Town, South Africa, 28 September-1 October 2011.

Baye M. F. and Epo, N. B. 2013. "Impact of human capital endowments on inequality of outcomes in Cameroon". Review of Income and Wealth, 2013.

Becker, G. S. 1964. Human capital". (1st edn), New York: Columbia University Press.

Bergheim, S. 2006. "Measures of wellbeing". Deutsche Bank Research Frankfurt am Main Germany. Available on the internet at http://www.dbresearch.com.

Blinder, A. S. 1973. "Wage discrimination: Reduced form and structural estimates". Journal of Human Resources, Vol. 8, No. 4: 436- 455.

Bourguignon, M. Fournier and M. Gurgand. 2001. "Fast development with a stable income distribution: Taiwan, 1979-1994". Review of Income and Wealth, Vol. 47, No. 2: 139-63.

Bourguignon, F., G. Ferreira and M. Menendez. 2007. "Inequality of opportunity in Brazil". Review of Income and Wealth, 53: 585-618.

Brookhart, M. A., Rassen, J. A. and Schneeweiss, S. 2010. "Instrumental variable methods in comparative safety and effectiveness research". Pharmacoepidemiology and Drug Safety, Vol. 19, No. 6: 537-554.

Chameni Nembua, C. and Miamo Wendji, C. 2012. "Inequality of Cameroonian households: An analysis based on Shapley-shorrocks decomposition". International Journal of Economics and Finance, Vol. 4, No.6: 149-156.

Checchi, D. 2000. "Does educational achievement help to explain income inequality?" Working Paper No. 208. World Institute for Development Economics Research, United Nations University.

Coudel, A., J. S. Hentschel and Q. T. Wodon. 2002. "Poverty measurement and analysis". In J. Klugman (ed), A sourcebook for poverty reduction strategies, Washington, DC: World Bank.

Deaton, A. 1997. "The analysis of household surveys: A microeconometric approach to development policy". Washington DC: World Bank.

Deininger, K. and Squire, L. 1996. "A new data set measuring income inequality". World Bank Economic Review,10: 565-591.

DiNardo, J., N. M. Fortin and T. Lemieux. 1996. "Labour market institutions and the distribution of wages, 1973-1992: A semi-parametric approach". Econometrica, Vol. 64 No.5: 1001-44.

Epo, B., Baye, F. and T. Manga. 2011. "Explaining spatial and inter-temporal sources of poverty, inequality and gender disparities in Cameroon: A regression-based decomposition analysis". Poverty and Economic Policy (PEP) Research Network. Research Project No. 11321, June 2010, www.pep-net.org.

Fambon, S. 2010. "Inequality in the distribution of household expenditure in Cameroon: A Gini decomposition analysis based on the Shapley-value approach". Educational Research, Vol. 1, No.2: 021-031.

Fambon, S. 2011. "Inequality in the distribution of household expenditure in Cameroon." IARIW -Statistics South Africa conference on experiences and challenges in measuring national income, wealth, poverty, and inequality in African countries, 28 September-1 October 2011, Cape Town, South Africa.

Fambon, S. 2014. "Sources of income inequality: Empirical evidence from Cameroon". The annual bank conference on Africa, Paris, 23-24 June, 2014.

Ferreira Francisco H. and Jeremie Gognoux. 2008. "The measurement of inequality of opportunity: Theory and application to Latin America". Policy Research Working Paper 4659.

Fields, G. and Yoo G. 2000. "Falling labour income inequality in Korea's economic growth: Pattern and underlying causes". Review of Income and Wealth, Vol. 46, No. 2: 139–159.

Fields, G. S. 2003. "Accounting for income inequality and its change: A new method with application to distribution of earnings in the United States." Research in Labour Economics, 22, 1.38.

Galor, O. and Zeira, J. 1993. "Income distribution and macroeconomics". Review of Economics Studies, Vol. 60, N0.1: 35–52.

Galor, O. 2000. "Income distribution and the process of development". European Economic Review, 44: 706-712.

Gibbons, S., S. Machin and O. Silva (2005). "Competition, Choice and Primary School Performance". Mimeo CEP-LSE.

Glomm, G. and Ravikumar, B. 1992. "Public versus private investment in human capital: Endogenous growth and income inequality". Journal of Political Economy, Vol. 100, No. 4: 818-834.

Hao, Q. M. 2005. "Access to finance and poverty reduction: An application to rural Vietnam." University of Birmingham: 1-327.

International Monetary Fund. 2003. "Cameroon: Poverty Reduction Strategy Paper, August 2003". IMF Country Report No. 03/249: http://www.imf.org.

Kaldor, N. 1956. "Alternative theories of distribution". Review of Economic Studies, Vol. 23, No. 2: 83-100.

Knight, J. and Sabot R. 1983. "Educational expansion and the Kuznets effect". American Economic Review, Vol. 73, No. 2: 1132-1136.

Lopez, R., Vinod, T. and Yan, W. 1998. "Addressing the education puzzle: The distribution of education and economic reform". World Bank Policy Research Working Paper No. 2031.

Maria, F. and F. Jose. 2008. "The determinants of poverty in Cape Verde. An econometric approach". Bulletin of Statistic and Amp; Economics.

Mincer, J. 1958. "Investment in human capital and the personal income distribution". Journal of Political Economy, 66: 281-302.

Morduch, Jonathan and Sicular Terry. 2002. "Rethinking inequality decomposition, with evidence from rural China". The Economic Journal, Vol. 112: 93-106.

Murray, M. P. 2006. "Avoiding invalid instruments and coping with weak instruments," Journal of Economic Perspectives, 20: 111–32.

Mwabu, G. 2009. "The production of child health in Kenya: A structural model of birth weight". Journal of African Economies, 18: 212-60.

National Institute of Statistics - NIS. 2002. Preliminary results of ECAM I and ECAM II. Ministry of Economic Affairs, Programming and Regional Development.

National Institute of Statistics. 2004. Annuaire Statistique du Cameroun. Edition 2004. Tome III.

National Institute of Statistics. 2007. Enquête Camerounaise Auprès des Ménages (ECAM III).. National Institute of Statistics, 2007.

National Institute of Statistics. 2008. Tendance, Profil et Déterminant de la Pauvreté au Cameroun en 2007. Disponible sur www.statistic-Cameroun.org. May 2009.

Oaxaca, R. 1973. "Male-female wage differences in urban labour markets". International Economic Review, Vol. 14, No. 3: 693-709.

Oyekale, A.S., A.I. Adeoti and T.O. Oyekale. 2007. "Regression-based approach to income inequality decomposition in rural and urban Nigeria". International Journal of Economic Perspective, Vol. 1, No. 2: 45-54.

Park, K. H. 1996. "Educational expansion and educational inequality on income distribution". Economics of Education Review, Vol. 15, N0.3: 51-58.

Roemer, J. E. 1993. "A pragmatic theory of responsibility for the egalitarian planner". Philosophy and Public Affairs, 22: 146-66.

Roemer, J. E. 1998. "Equality of opportunity". Cambridge MA: Harvard University Press.

Roemer, J. E., R. Aaberge, U. Colombino, J. Fritzell, S. Jenkins and A. Lefranc (2003). "To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?" Journal of Public Economics, 87, 539-65.

Rousseau, J.J. 1754. "Discourse on the origin of inequality". Indianapolis, Cambridge.

Schultz, T. P. 2003. "Wage rentals for reproducible human capital: Evidence from Ghana and Ivory Coast," Economic Growth Center, Yale University, Discussion Paper 868.

Shea, J. 1997. "Instrument relevance in multivariate linear models: A simple measure". Review of Economics and Statistics, 49 (2): 348-52.

Shorrocks, A. F. 1980. "The class of additively decomposable inequality measures." Econometrica, Vol. 48: 613-625.

Shorrocks, A. F. 1982. "Inequality decomposition by factor components". Econometrica, 65: 177-207.

Shorrocks, A.F. 1984. "Inequality decomposition by population sub-groups". Econometrica, 52: 1396-1385.

Shorrocks, A.F. 1999. "Decomposition procedure for distribution analysis: A unified framework based on Shapely value". University of Essex and Institute for Fiscal studies, Colchester, England.

Shorrocks, A.F. and Wan G. 2005. "Spatial decomposition of inequality". Journal of Economic Geography, Vol. 5, No.1: 59-81.

Stock, J. and M. Yogo. 2005. "Testing for weak instruments in linear IV regression," in Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, Chapter 5: 80-108.

Tabi, A. J. 2009. "Components, sources and determinants of aggregate inequality trends in the distribution of household economic wellbeing in Cameroon, 1996-2001".Undefended and Unpublished PhD Thesis. University of Yaoundé II, Soa, Cameroon.

Thorbecke, E. and C. Charumilind. "Economic inequality and its socioeconomic impact." World Development Vol. 30, No. 9: 1477-1495.

UNU. 2000. "UNU/WIDER-UNDP World Income Inequality Database", v 1.0 ed. United Nations University/WIDER.

Wan, G.H. 2002. "Regression-based inequality decomposition: Pitfalls and a solution procedure". UNU-WIDER Discussion Paper No. 2002/101. Helsinki.

Wan, G.H. 2004. "Accounting for income inequality in rural China: A regression-based approach". Journal of Comparative Economics. Vol. 32, No.2: 348-63.

Wan, G.H. and Z.Y. Zhou. 2005. "Income inequality in rural China: Regression-based decomposition using household data". Review of Development Economics, Vol. 9: 107–120.

Wan, G.H. 2007. "Understanding regional poverty and inequality trends in China: Methodological issues and empirical findings". Review of Income and Wealth, 53(1), 1-10.

World Bank. 2005. World Development Report 2006: Equity and Development. Washington DC: World Bank and New York: Oxford University Press.

Wooldridge, J. M. 1997. "On two Stage Least Squares estimation of the average treatment effect in a random coefficient model". Economics Letters, Vol. 56: 129-133.

Wooldridge, J. M. 2002. Econometric analysis of cross section and panel data, Cambridge, MA: MIT Press.

Yuko, A., J.M. Kim. and A. Kimhi. 2006. "Determinants of income inequality among Korean farm households". Economic Research Centre, Discussion paper No. 161, School of Economics, NAGOYA University.

Appendix

Appendix Table 1: Marginal contributions of the regressed income sources using the Shapley approach for a linear income generating function

GINI Coefficient											
Source	level_1	level_2	level_3	level_4	level_5	level_6	level_7	level_8	level_9	level_10	Total
Education	0.0387	0.0370	0.0441	0.0436	0.0101	-0.035	-0.017	-0.025	-0.019	-0.002	0.0743
Age	0.0214	0.0224	0.0348	0.0418	0.0140	-0.030	-0.016	-0.028	-0.028	-0.016	0.0163
Dependency ratio	-0.036	0.0056	0.0049	0.0059	0.0255	0.0024	0.0012	0.0006	0.0003	0.0002	0.0104
Credit access (cluster mean)	0.0712	0.0021	-0.007	0.0165	0.0265	-0.009	-0.005	-0.002	-0.001	-0.001	0.0909
Rural	-0.035	-0.016	0.0198	0.0522	0.0465	0.0161	0.0355	0.0215	0.0134	0.0106	0.1647
Electricity	0.0099	-0.010	-0.013	0.0121	0.0236	-0.012	-0.007	-0.003	-0.001	-0.001	-0.0016
Land	-0.039	-0.049	-0.038	0.0176	0.0545	0.0237	0.018	0.0105	0.0058	0.0044	0.0077
Radio (cluster mean)	-0.021	-0.016	0.0095	0.0244	0.0062	-0.024	0.0061	0.0030	0.0011	0.0008	-0.0094
Bilingual	0.0756	0.0233	-0.013	0.0028	0.0174	-0.015	-0.010	-0.004	0.0008	0.0009	0.0794
Constant	0.0000	0.0036	-0.013	-0.005	-0.032	-0.041	-0.025	-0.035	-0.032	-0.018	-0.1018
Total											0.3310

Source: Computed by authors using STATA 10 and the DASP 2.1 Software developed by Araar and Duclos (2009)

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.

www.aercafrica.org

Learn More

www.facebook.com/aercafrica

twitter.com/aercafrica

www.instagram.com/aercafrica_official/

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