OVER-REPORTING OF HOUSEHOLD CONSUMPTION EXPENDITURE ESTIMATES IN THE FIFTH ROUND OF GHANA LIVING STANDARDS SURVEY

N.N.N.Nsowah-Nuamah[1], L. Munyakazi[b1] J. Dede Anum[2]
[b]*Department of Mathematics and Statistics, Kumasi Polytechnic, Ghana*
[2]*Ghana Statistical Service, Ghana.*

ABSTRACT

The effectiveness of government policies and programs relies heavily on measurements of poverty and inequality. Living Standards Surveys (LSS), such as the recent Ghana Living Standards Survey of 2005/2006 provides information on both *income* and *expenditure* collected from 8687 households survey across the country. The survey identifies six (6) different sources of income and expenditure including *income* from employment, household agriculture, non-farm self-employment, rent, remittances and other incomes whilst *expenditure* were on food, housing, remittances and other expenditures.

The objectives of the study are to compare estimates from the two indicators of poverty and to determine the magnitude of over-reported households estimates of the consumption expenditure. Results seem to indicate that the *total average* annual household income is GH¢1,288.74 *vs*. GH¢1,983.24 (Ghana Cedis) for expenditure for an average household size of four. With respect to income, regional comparisons show Greater Accra region to have the highest income (GH¢ 1,673.60). In general, the coastal regions have income higher than the national average except Volta Region (GH¢ 961.57). Households in Greater Accra have a higher *expenditure* (GH¢ 3,061.07) followed by households in Ashanti region (GH¢ 2,055.67). In all levels of comparison, the reported expenditure estimates are *significantly higher* (31%) than estimates from income (p-value < 0.001). Estimated Gini coefficients are also different in the two indicators. They are 0.57 and 0.40 for total expenditure and total income respectively.

Key words: Income, Consumption Expenditure, Living Standards Survey, Indicator, Poverty measures, Gini coefficient, Ghana Cedis.

[1] Associate Professor in Statistics and Head of the Department
[3] Professor in Statistics and Rector of the Kumasi Polytechnic

Introduction

The Ghana Statistical Service (GSS) and the World Bank agreed in December 1995 that the establishment of a permanent welfare monitoring capability was necessary in the national poverty reduction strategy. The Ghana Living Standard Survey (GLSS) survey provides information on patterns of household's consumption and expenditure and serve as the basis for the construction of the Consumer Price Index basket and to updating National Accounts.

The wealth of information contained in the survey has resulted in helping to measure poverty trends of the country from 1990 to 2006 as it is a nationally representative survey. It has also been used in other measurements of welfare. The major focus of the GLSS survey is to provide estimates with acceptable precision for a variety of indicators on various aspects of living conditions, including household consumption and expenditure, health, education, employment,etc...

Income and expenditure forms the basis for differences in the welfare of citizens in a country. It can be said that not every person or a household as a unit of analysis of this kind of study earns an income in the formal sense, but almost all households expend some of either in cash or through barter. This phenomenon prevails in Ghana as well as in most part of the world except where some households depend on social grant.

Individuals are reluctant to disclose their incomes even when compelled. An alternative means is the use of the consumption expenditure method which is not closely tied to short term fluctuations in income but is smoother and less variable (Deaton et al, 2002).

Methodology

Sources of Data

The firth round of the Living Condition Survey conducted by the Ghana Statistics office has a sample size of 17,372 observations obtained from 8687 households across the country.

The design and collection of collection followed the international standard of a random two-stage sampling where the Primary Sampling Units (PSU) are the Enumeration Areas and the Secondary Sampling Units (SSU) are the households within each PSU (Levy and Lemeshow, 1999). Both the PSU and the SSU were selected randomly to be representative of the 10 Regions of the country, and all the other variables of the table above. The selected independent variables are separated domains that require reliable estimates.

The GLSS data is collected over a 12-month period in order to ensure continuous recording of household consumption and expenditures and the changes occurring across the country.

Method of analysis

The unit of analysis is the household which was defined as a person or group of related or unrelated persons, who live together in the same housing unit, who acknowledge one adult male or female as the head of the household, who share the same housekeeping and cooking arrangements, and are considered as one unit. The different estimates of income and expenditure are explored with the characteristics of the head of household since all the estimates for each individual have been aggregated to the household. Inequality measures will then be used to find out the extent of the deviations and the sources causing the most variations. Finally the two paired sources of estimates from each household would be presented in the descriptive statistics and as well as analyzed to find a relationship between using error-in-variables method (measurement error method). The Gini coefficients were calculated using STATA. All the other analyses were performed using the SAS® (SAS, 2009).

Model specification

The variables of interest (expenditure and income) were jointly analyzed following a general linear model formulation (Snedecor and Cochran, 1989; Steel and Torrie, 1980; Milliken and Johnson, 2002):

$$y = X\beta + \varepsilon \qquad (1)$$

where y is the vector of values for expenditure and income, X is the design matrix that contains type of expenses (either expenditure or income), region of the country (10 in all), the Household (HH) age and gender, the socio-economic group, the economic zone, the level of education attained, and HH size; the β is the vector of solution and ε are the residuals. An economic component of the analysis is the Gini coefficient to evaluate regional inequalities.

Empirical Results

Descriptive statistics

The descriptive statistics are not reported here. The results of the overall analysis that contains both qualitative and quantitative variables as formulated in (1) (Milliken and Johnson, 2002). The general linear model goodness-of-fit is F of 148.29 (p-value <0.0001) indicating a good fit of the data. A Bonferroni's adjustment was applied to separate means of interest. The difference between *Expenditure* and *Income* is highly significant (p-value <0.001). An additional age increases expenses by 6.14 ± 1.04 Ghana □. On the other hand, an additional member of HH increases the expense by 185.09 ± 5.62 Ghana □. On a regional level, Greater Accra shows the highest adjusted Expenditure, very different from the other Regions (p-value<0.001). However the analysis showed that the pattern of Income has a different trend. Greater Accra is only different from Upper East and Upper West Regions respectively (the p-values are equal to 0.0065 and <.0001). On the average male headed HH earns 6.76% more than female headed HH.

An Approximate Relationship between Consumption Expenditure and Income

A close relationship between the two variables is expected. In order to describe the case when both axes are subject to error, the following equations are needed (Gillard and Illes, 2006, 2007):

$$Y_i = \eta_i + \varepsilon_I \qquad (2)$$
$$X_i = \xi_i + \delta_i$$

each illustrates the errors associated with the measurement methods (in our case, *Income* and *consumption expenditure*). In addition, one assumes a true relationship exists between the unobservable or latent variables η_i and ξ_i:

$$\eta_i = \beta_0 + \beta_1 \xi_i$$

and thus (3)

$$Y_i = \beta_0 + \beta_1 \xi_i + \varepsilon_I$$

There are no easy solutions to (3). The maximum likelihood solution of the slope is given in Kendal and Smith (1961) and requires some strong assumptions.

In this study we apply a simple regression of Income on Expenditure as an approximation of such relationship (Weisberg, 1985; Draper and Smith, 1989). We calculate the *equivalence limits* against the linear hypothesis

$$H_0: \ \beta_0 = 0 \text{ and } \beta_1 = 1.$$

It remains to observe whether the estimated intervals for the *slope* include 1, and the limits of the estimated *intercept* of the same equations include 0. Our data indicated that the expenditure is a dependable function of income:

$$Expenditure = 324.810 + 1.308 \mid Income$$

The tests indicated that the intercept is significant (p-value=0.0.0015) from zero and so is the slope (p-value<0.00001). This result seems to suggest that *expenditure* is a 1.31 ± 0.05 multiple of the *income* values with an upwards shift of $GH¢\,324.81$. The coefficient R^2 is 0.97. Thus the linear hypothesis ($\beta_0 = 0$ *and* $\beta_1 = 1$) is rejected.

Gini coefficients for Income and expenditure

The Gini coefficient is among the traditional measures of inequality among values of expenditure or income. The coefficients from Expenditure are lower that the coefficients from Income. At the national level, they are 0.57 and 0.40 for Income and Expenditure respectively.

References

1. Akib-Bin-Amin (2011). Causal Relationship between consumption Expenditure and Economic Growth in Bangladesh. World Journal of Social Sciences Vol 1. No 2 158-169
2. SAS Institute Guide (2009). Cary, NC: SAS Institute Inc version. 9.2
3. Draper and Smith, (1998). Applied Regression Analysis. Wiley Series in Probabilitiy and Statistics. 706 p.
4. Milliken A.G and Johnson D.E (2002). Analysis of Messy Data. Vol III. Analysis of Covariance. Chapman Hall/CRC 605 pp
5. Spearman C. 1904). The proof and measurement of association between two things. Am. J. Psychol. **15**:72-101
6. Levy P.S, S. Lemeshow (1999). Sampling of Populations. Methods and Applications 3rd Ed Wiley Series in probability Statistics 525pp
7. Steel, R.G, and T.H. Torrie (1980). Principles and Procedures of Statistics. A Biometric Approach. 2nd ed. 633pp.
8. Snedecor G. W and W.G. Cochran (1989). Statistical Methods, 8th Ed. 503 pp
9. Draper, N.R and Y Yang (1996). Generalization of the geometric mean functional relationship. *Computational Statistics and Data Analysis* **23** 355-372.
10. Gillard J.W. and T.C. Illes (2006). Variance Covariance matrices for Linear regression with Errors in both variables. School of Mathematics, Senghenydd Rd Cardiff University
11. Gillard J.W. and T.C. Illes (2007). Methods of fitting straight lines where both variables are subject to measurement error. *Current Clinical Pharmacology*, **4**, 164-171
12. Weisberg, S. (1985). *Applied Linear Regression* (2nd ed). New York: John Wiley
13. Deaton A Zaidi S (2002): Guidelines for constructing Aggregates for welfare Analysis. LSMS Working Paper No 135.
14. Kendall G.M (1961). *The Advanced Theory of Statistics.* Charles Grffin and Co