Research Journal of Mathematics and Statistics 6(2): 16-22, 2014
DOI:10.19026/rjms.6.5808
ISSN: 2042-2024, e-ISSN: 2040-7505

Submitted: December 04, 2013 Accepted: December 26, 2013 Published: May 25, 2014

Research Article

Catalogistico Discriminant Analysis: A Methodology for Analyzing Catastrophic Spending on Health in Statistically Under-developed Countries

Felix O. Mettle, Abeku A. Asare-Kumi, Isaac K. Baidoo and Ezekiel N.N. Nortey
Department of Statistics, University of Ghana, P.O. Box LG 115, Legon

Abstract: This study proposes a methodology for analysis of catastrophic spending on health in statistically under-developed countries. A binary logistic regression model, based on data from households with reported non-zero expenditure on health, is proposed for the estimation of the likelihood of spending on health for all households irrespective of whether they spent on health or not within the reference period for the survey. “Univariate” discriminant functions, also based on data from households who spent on health within the reference period of the survey, were proposed for discriminating households that made catastrophic expenditure on health from those who did not. An application of this methodology to the data from the Ghana living Standards survey (round V) indicates that the binary logistic regression model estimates correctly at least 78% of household’s likelihood of spending on health while correctly discriminating the households as having a catastrophic expenditure.

Keywords: Binary logistic regression, catastrophic health expenditure, discriminant analysis, survey reference period

INTRODUCTION

The amount of money spent on health in economically under-developed countries place families at the point of insolvency. The lack of adequate insurance and other risk management facilities to help hedge against financial ruin has been addressed by many authors (Lave *et al*., 1998; Blankenau *et al*., 2009).

Examining catastrophic health expenditure to evaluate health systems in various countries dates back to Berki (1986). The inception of Berki’s work has led to much literature over the years on the catastrophic spending on health in various countries with various demographics and several proposed methodologies. According to Berki (1986), catastrophic expenditure is one which constitutes a large part of a household budget and hence affects the household’s ability to maintain its “normal” standard of living. Wyszewianski (1986) also defines catastrophic health expenditure as a proportion of a household’s expenditure. This means that if expenditure on health care constitutes a large proportion of the household’s budget, then the household may be considered as having spent catastrophically on health. The question that ought to be asked is how large is large enough?

Russell (1996) defined catastrophic health expenditure in relation to a household’s capacity to pay and hence focused on the opportunity cost of healthcare expenditure to that household concerned.

Wagstaff and van Doorslaer (2003) measured the incidence and extent of catastrophic health expenditure in Vietnam using proportion of the total expenditure in a household. Wagstaff and van Doorslaer (2003) in the manner of Berki (1986) defined catastrophic health expenditure as that which exceeds some fixed proportion of a household’s total expenditure. There is no hard and fast rule as to how to set this threshold proportion of household’s total expenditure on health. Wagstaff and van Doorslaer (2003) set this threshold at 10% of total household budget.

Xu *et al*. (2003) explored catastrophic health expenditure for multiple countries using regression analysis and like Russell (1996) defined catastrophic health expenditure in relation to a household’s capacity to pay. However he defined the capacity to pay as the household income after accounting for median level of food consumption in society. Xu *et al*. (2003) then in the manner of Wagstaff and van Doorslaer (2003) arbitrarily classified a household to have spent catastrophically on health if the household’s financial contributions towards health exceeded 40% of the family’s remaining income after it is able to meet its subsistence needs.

It is noteworthy that both Wagstaff and van Doorslaer (2003) and Xu *et al*. (2003) consider catastrophic health expenditure as involuntary and make the assumption that it deteriorates a household’s welfare (Wagstaff, 2008). Also both methods have had wide applications across countries of varying backgrounds and economic statuses.

Xu *et al*. (2003) also observed that significant differences existed between the proportions of households facing catastrophic household spending on

Corresponding Author: Abeku A. Asare-Kumi, Department of Statistics, University of Ghana, P.O. Box LG 115, Legon

health. They claim that household's in developed countries were protected from catastrophic expenditure on health due to their advanced social institutions such as social insurance or tax-funded health systems. This finding turns our focus to the under-developed countries where it is necessary that catastrophic expenditure be well assessed and remedial measures taken to address the issue.

Kim and Yang (2011) using South Korea as a case study, documented amongst others that the burden of health care cost and the effects on household economies depended on the countries health system and the ability of the individuals to pay and that the degree of household income loss was dependent on the employment status and the income earned by the sick member of the household.

More significant were the limitations of their results which stated that the data used in their analysis had no information on the health status of the ill family member and ill individuals who did not utilize health care due to the high cost. The lack of sufficient and relevant data in under-developed countries has always limited the statistical approach to analyze and make sense of the health data gathered from these countries. The duplicity of roles performed by various state agencies and storage of data are examples of well documented challenges to research in under-developed countries. Xu *et al*. (2006) explained that their research in Uganda on catastrophic health expenditures failed to capture those spending on alternative and traditional medicine which is a significant component of health expenditure in developing countries. It has also been observed that the living standard surveys and socio-economic surveys of many countries fail to capture in detail household expenditure on health especially that from alternative medicine.

Data collected in the 61st round of socio-economic survey conducted by the National Sample Survey Organization (NSSO) during July, 2004 to June, 2005 in India captured expenditure on necessary consumption items that affect health but with preference period of 30 days (Pal, 2010). This means that for a household that did not have to spend on health in the last 30 days, health care expenditure is zero. It becomes impossible to determine whether or not such a household which has similar characteristics with other households (with reported health expenditure in the same reference period) spend catastrophically on health. Pal (2010) sites the non-inclusion of transportation cost associated with medical expenses. Lara and Gómez (2011) in their discussion on their findings on factors that affect catastrophic spending in Bogota, Columbia cautioned that the level of catastrophic spending seen in their study was based only on the group of households that incurred some health expenditure during the year of the General Social Security Health System (GSSHS) and that unequal access to the system and related out-of-pocket expenses may actually be greater if they took into consideration that, there is still a group of the population without health expenditure simply due to geographical barriers or to a lack of economic capacity, which hinders them from incurring any health expenditure even if they wish to do so. Several discussions to findings in the research of health care expenditure have also documented similar lapses as a limitation to their results.

On methodology that has been used in analyzing health expenditure, Filmer and Pritchett (2001), constructed a wealth index using principal components analysis to identify which factors contributed to catastrophic spending on health. The World Health Organization (WHO) has proposed different methodologies for estimating financial protection; the only distinguishing factor in these methodologies is how payment capacity and catastrophic health spending are measured.

Xu *et al*. (2006) presented the current methodology proposed by the WHO, which affirms that "health spending is catastrophic when a household's out-of-pocket health payments are equal to or greater than 40% of the household's payment capacity or non-subsistence expenditure". "Prior studies considered different thresholds as reference points for establishing catastrophic spending; they vary from 10 to 50% depending on the reference country's level of development, the methodology used to measure catastrophic spending, the method employed to calculate payment capacity and the definition of subsistence expenditure" (Lara and Gómez, 2011).

Xu *et al*. (2006) compared a restricted regression analysis model to that of an unrestricted regression model and tested the difference in model coefficients of the separate equations for two levels of prosperity (poor and non-poor) within the data compared to the pooled equations coefficients using the log-likelihood Chow test. They also adopted Multinomial regression models for certain aspects of their analysis that took into consideration where health services were sought by the sick person. Daneshkohan *et al*. (2011) adopted the use of WHO's methodology to analyze catastrophic household expenditure on health in Iran. Nguyen *et al*. (2013) in discussing catastrophic spending on injuries in Vietnam adopted a prospective cohort study where a modified Poisson approach was used to predict catastrophic spending on injuries.

Governments, especially those of statistically under-developed countries have tried over the years with the use of these research findings (with significant gradual improvements in the past decade) to reduce the catastrophic spending by its citizens on health. Despite governments' efforts, catastrophic health care spending is not rare in under-developing countries.

Analysis on the data used to determine catastrophic health expenditure is only based on households with reported expenditure in the surveys reference period and only on the data collected by such agencies that collect the data. This study proposes a methodology capable of classifying all households with regards to catastrophic expenditure on health irrespective of whether the households reported on health expenditure for the specified reference period or not.

Although a lot of data is sought on health in living standards surveys, the reference period for which respondents ought to answer questions on health expenditure is small. The short reference period often yields a large number of zeros since only a small proportion of the population would have had expenditure on health within the specified reference period. This makes it difficult to effectively identify the total number of households that have had catastrophic health expenditure. To go around this predicament, we propose this methodology.

METHODOLOGY

To solve the problem of inadequate response on health expenditure which is widely evident in most living standard surveys the world over, we propose a methodology that is in two (2) parts. The first part seeks to solve the problem of inadequate/lack of a response on health expenditure which is widely evident in most living standard surveys the world over. The second seeks to discriminate households with regards to them having catastrophic spending on health.

Estimation of a household's likelihood of spending on health: Suppose V_i is the ith household expenditure on health for q weeks and U_i is the ith household's non-food expenditure in a year (nominal).

Then on average, the ith household spends $w_i = \frac{U_{i*}q}{52}\,(i = 1, 2, \ldots, n)$ on non-food expenditure in q weeks; where n is the total number of households who had reported (non-zero) expenditure in the last q-weeks at the time of study:

Next, let $Z_i = \frac{V_i}{w_i}, i = 1,2, \ldots, n$

Then in the manner of Xu *et al*. (2003) a household is said to have spent catastrophically on health if:

$$Z_i > c, where\ 0 < c < 1$$

For the ith household, we define a dichotomous variable y_i such that:

$$y_i = \begin{cases} 1, if\ Z_i > c \\ 0, otherwise \end{cases}$$

For a dichotomous outcome variable Alan (2012) explains the use of binary logistic regression in determining the likelihood of belonging to any of the two categories based on certain statistically significant explanatory variables.

Now suppose the binomial logistic regression model:

$$logit(\theta_i) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon,$$

$$\theta_i = P(y_i = 1 | X_1, \ldots, X_k) \quad (1)$$

is significant and correctly classifies at least 75% of households who have spent catastrophically on health, then an estimate of model can be written as:

$$logit(\widehat{\theta_i}) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k \quad (2)$$

where, $\hat{\beta}_i, i = 0, 1, 2, \ldots \ldots . . k$ are estimates of the parameters β_i and $\widehat{\theta_i}$ is the estimate of the likelihood of a household having a catastrophic spending on health. Given the explanatory variables $X_1, X_2, \ldots, X_k$, θ_i can be estimated as follows:

$$\widehat{\theta_i} = \frac{exp\{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k\}}{1 + exp\{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k\}} \quad (3)$$

Model (3) is then used to estimate the likelihood of a household's spending on health for all households observed to have spent on health in the past q weeks. These estimates are then used to develop a discriminant model (function) as discussed in the following sequel.

Discriminant models (functions): Suppose $f_1(\theta)$ and $f_2(\theta)$ are the probability density functions associated with the random variable θ for the populations π_1 (households that spent catastrophically on health) and π_2(households that have not spent catastrophically on health), respectively. A household whose likelihood θ of spending catastrophically on health must be assigned to either π_1 or π_2.

Let Ω be the sample space of θ and R_1 and $R_2 = \Omega - R_1$ form a partition of Ω. If R_1 is the set of all values of θ for which a household is classified as π_1 and R_2 is the set of values of θ for which a household is classified as π_2, then the probability of misclassifying a population π_1 as π_2 is given by; $P(2|1) = P(\theta_{ij} \in R_2 | \pi_1) = \int_{R_2} f_1(\theta)\, d\theta$ and the probability of misclassifying a population π_2 as π_1 is; $P(1|2) = P(\theta_{ij} \in R_1 | \pi_2) = \int_{R_1} f_2(\theta)\, d\theta$.

According to Johnson and Wichern (2007) a reasonable classification rule should have an expected cost of misclassification (ECM) as small as possible; with:

$$ECM = c(2|1)\, P(2|1)\, P(1) + c(1|2)\, P(1|2)\, P \quad (4)$$

where, $c(i|j)$ is the cost of misclassifying a population π_j as π_i, $(i = 1, 2)$ and $P(i), (i = 1, 2)$ is the prior probability of π_i and $P(1) + p(2) = 1$.

The regions R_1 and R_2 that minimize the ECM, according to Johnson and Wichern (2007) are defined by the values θ for which the following holds:

$$R_1 : \frac{f_1(\theta)}{f_2(\theta)} \geq \frac{c(1|2)P(2)}{c(2|1)P(1)}$$
$$R_2 : \frac{f_1(\theta)}{f_2(\theta)} < \frac{c(1|2)P(2)}{c(2|1)P(1)} \quad (5)$$

If π_i $(i = 1,2)$ has a normal distribution with mean μ_i and variance σ_i^2, then the density ratio based on θ is given by:

$$\frac{f_1(\theta)}{f_2(\theta)} = \frac{\frac{1}{\sqrt{2\pi}\sigma_1} * e^{-\frac{1}{2}(\frac{\theta-\mu_1}{\sigma_1})^2}}{\frac{1}{\sqrt{2\pi}\sigma_2} * e^{-\frac{1}{2}(\frac{\theta-\mu_2}{\sigma_2})^2}} = \left[\frac{\sigma_2}{\sigma_1}\right]^{\frac{1}{2}} e^{-\frac{1}{2}\left[\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}\right)\theta^2 - 2\left(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_2}{\sigma_2^2}\right)\theta + \left(\frac{\mu_1^2}{\sigma_1^2}-\frac{\mu_2^2}{\sigma_2^2}\right)\right]} \quad (6)$$

Rearranging and taking the natural logarithm of both sides, the first inequality in (3), by trivial algebra becomes:

$$\frac{1}{2}(\sigma_1^2 - \sigma_2^2)\theta^2 + (\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\theta + (\mu_1^2\sigma_2^2 - \mu_2^2\sigma_1^2) \geq \sigma_1^2\sigma_2^2 ln\left[\left(\frac{\sigma_1}{\sigma_2}\right)^{\frac{1}{2}} \frac{c(1|2)p(2)}{c(2|1)p(1)}\right] \quad (7)$$

However, if $\sigma_1 = \sigma_2 = \sigma$ Eq. (6) becomes:

$$\frac{f_1(\theta)}{f_2(\theta)} = \left[\frac{\sigma_2}{\sigma_1}\right]^{\frac{1}{2}} e^{\frac{(\mu_1-\mu_2)\theta}{\sigma^2} - \frac{1}{2\sigma^2}(\mu_1^2-\mu_2^2)} \quad (8)$$

Again re-arranging and taking the natural logarithm of both sides, the first inequality of (5) becomes:

$$(\mu_1 - \mu_2)\theta - \frac{1}{2}(\mu_1^2 - \mu_2^2) \geq \sigma^2 ln \left[\frac{C(1|2)p(2)}{C(2|1)p(1)}\right] \quad (9)$$

Now, labeling the left hand side of (7) and (9) as quadratic and linear discriminant functions $l_{(1)}$ and $l_{(2)}$ and the corresponding right hand sides as the critical values $c_{(1)}$ and $c_{(2)}$ respectively, the sample estimate of the discriminant functions and their critical values are given by:

$$\hat{l}_{(1)} = \frac{1}{2}(s_1^2 - s_2^2)\theta^2 + (\bar{\theta}_1 s_2^2 - \bar{\theta}_2 s_1^2)\theta + (\bar{\theta}_1^2 s_2^2 - \bar{\theta}_2^2 s_1^2) \text{ with } \hat{c}_{(1)} = s_1^2 s_2^2 ln\left[\left(\frac{s_1}{s_2}\right)^{\frac{1}{2}} \frac{c(1|2)p(2)}{c(2|1)p(1)}\right]$$

and $\hat{l}_{(2)} = (\bar{\theta}_1 - \bar{\theta}_2)\theta - \frac{1}{2}(\bar{\theta}_1^2 - \bar{\theta}_2^2)$ with $\hat{c}_{(2)} = s^2 ln\left[\frac{c(1|2)p(2)}{c(2|1)p(1)}\right]$

where, $\bar{\theta}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\theta_{ij}$ and $s_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(\theta_{ij} - \bar{\theta}_i)^2$ are based on samples of size n_i from population $\pi_i (i = 1,2)$; and $s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ is the pooled sample variance.

By the minimum ECM rule, a household with θ likelihood of spending on health is classified as having spent catastrophically on health if:

$$\hat{l}_{(2)}(\theta_{ij}) = \frac{1}{2}(s_1^2 - s_2^2)\theta^2{}_{ij} + (\bar{\theta}_1 s_2^2 - \bar{\theta}_2 s_1^2)\theta_{ij} + (\bar{\theta}_1^2 s_2^2 - \bar{\theta}_2^2 s_1^2) \geq \hat{c}_{(2)} \text{ For } \sigma_1 \neq \sigma_2$$

$$\text{OR } \hat{l}_{(1)}(\theta_{ij}) = (\bar{\theta}_1 - \bar{\theta}_2)\theta_{ij} - \frac{1}{2}(\bar{\theta}_1^2 - \bar{\theta}_2^2) \text{ For } \sigma_1 = \sigma_2 \quad (10)$$

The discriminant functions $\hat{l}_{(1)}$ and $\hat{l}_{(2)}$ are effective in classifying a household as having spent catastrophically on health or not if μ_1 is significantly different from μ_2.

Application: To apply the methodology proposed by this study, data on health expenditure for was taken from the fifth round of the Ghana Living Standard Survey (GLSS 5) conducted by the Ghana Statistical Service (GSS). The data is representative of a nationwide sample of 8687 households in 580 enumeration areas, involving 37,128 household members. Detailed information was collected on demographic characteristics of respondents and several aspects of living conditions including, health, education, household income, consumption and expenditure, employment, housing, agricultural activities, remittances, savings, credits and assets. Sections, namely Tourism, Migration and Remittances were introduced. The survey spans over one-year (twelve months) of data collection involving the years September 2005 to September 2006.

The dependent variable used in this study was computed from the following household variables:

- Frequent non-food expenditure
- Household Expenditure on Health

The original data on health expenditure were in the following format. Expenditure on illness or injury for 2 weeks, on immunization for 12 months, on sickness for 12 months, on prenatal care for 12 months, on contraceptive for one month and on health insurance for 12 months. To standardize these expenditures, all values were converted to expenditure for 2 weeks. The total expenditure v_i for each household on health was then computed for 2 weeks. (i.e., $q = 2$). For each household 2 weeks of non-food expenditure (w_i) was computed from the nominal non-food expenditure. The proportion (z_i) of non-food expenditure accounted for by expenditure on health was computed for each of the 3100 households who were found to have spent on health. These households were found to have come from the Western, Central and Greater Accra regions of Ghana.

The study used a threshold value of $c = 0.10$; and so all households with $z_i > 0.10$ were classified to have spent catastrophically on health. Thus, the dependent variable for the binary logistic regression is defined by $Y_i = 1$ for spending catastrophically on health and $Y_i = 0$, otherwise. Of the 3100 households 622 were found to have spent catastrophically on health.

After multicollinearity diagnosis, household size (X_1), ecological zone (X_2), age of the head of household (X_3), socio-economic group of the head of household (X_4) and the education level of the head of household

Table 1: Fitted binary logistic regression

Variable	β	SE	Odds-ratio	p-value
Constant	-1.013	0.232	0.363	0.000
Household size	-0.157	0.021	0.855	0.000
Ecological zone	-0.392	0.092	0.676	0.000
Age of household head	0.005	0.003	1.005	0.018
Socio-economic group	0.112	0.027	1.118	0.000
Education	0.039	0.017	1.040	0.021

Authors' computation using GLSS (Round V)

Table 2: Frequency distribution for classification of 3100 households with reported nonzero health expenditure

Classification	N	Mean	S.D.	S.E.
Catastrophic	662	0.2460	0.07990	0.0311
Non-Catastrophic	2438	0.2093	0.07698	0.0016
Total	3100	0.2171		

Authors' computation using GLSS (Round V)

(X_5) recorded a tolerance levels close to one; hence these variables were used in setting up the logistic regression model. The logistic regression of the dependent variable (Y_i) on the above variables was found to be significant ($\chi^2 = 231.206,\ df = 7, p-value < 0.001$) and correctly classifies 78.6% of the cases. Table 1 shows the estimates of the fitted binary logistic regression model for the data. The fitted logistic regression equation is given by:

$$logit(\widehat{\theta}) = -1.013 - 0.157X_1 - 0.392X_2 + 0.005X_3 + 0.112X_4 + 0.039X_5$$

This implies that:

$$\widehat{\theta}_i = \frac{exp\{\widehat{\boldsymbol{\beta}}\boldsymbol{X}'\}}{1 + exp\{\widehat{\boldsymbol{\beta}}\boldsymbol{X}'\}} \quad (11)$$

where, $\hat{\beta}$ = (-1.013 -0.157 -0.392 0.005 0.112 0.039)' and X = $(1, x_1, x_2, x_3, x_4, x_5)'$.

Equation (11), was used to compute the likelihood $\hat{\theta}$ of spending on health for all 3,100 households found to have spent on health.

Table 1 shows the statistically significant binary logistic regression model for the GLSS 5 data. The variables Household size, Ecological zone, Age of household head, Socio-economic group and Education level of the head of household was found to be significant at determining the likelihood of catastrophic expenditure for households captured in the data.

Table 2 displays the descriptive statistics of likelihood of spending on health by catastrophic spending status on health. The Levene's test for the equality of variance of likelihood of spending on health by the two groups of households was not significant (F = 0.864, p-value = 0.353). Hence, there is no difference in variances for two household classifications and this implies that a linear discriminant function $\hat{l}_{(2)}$ Eq. (10) is appropriate.

Based on the data, an estimate of the common variance called the pooled variance for the two groups of households was found to be 0.006024. And hence the linear discriminant function for the likelihood of spending on health is given by; $\hat{l}_{(2)} = (\bar{\theta}_1 - \bar{\theta}_2)\theta - \frac{1}{2}(\bar{\theta}_1^2 - \bar{\theta}_2^2) = 0.037\theta$ -0.008355 and $\hat{c}_{(2)} = s^2 ln\left[\frac{c(1|2)p(2)}{c(2|1)p(1)}\right] = 0$, on the assumption of equal cost of misclassification and equal prior probabilities for both groups of households. Therefore a household with θ likelihood of spending on health is said to have spent catastrophically on health if $\hat{l}_{(2)} > 0$. The independent sample t-test for equal mean likelihood of spending on health is significant (t = 10.78, df = 3098, p-value <0.001) and so the $\hat{l}_{(2)}$ is effective in discriminating a household who spent catastrophically on health from one which did not.

This result is applied to the whole GLSS5 data of 8,687 households irrespective of whether or not a household reported expenditure on health after estimating the likelihood of health expenditure for each of the households in the entire data set. The number of households in the analysis reduced from 8,667 to 7,448 as a result of the exclusion of households who failed to report on some of the variables used in the model.

Estimates of Catastrophic Expenditure (CE) and Share of Catastrophic Expenditure (SCE) were computed by sex of household head and, also by the region of household.

RESULTS

Female headed households (41.10%) seem to spend more catastrophically on health when compared to their male (19.59%) counterparts.

Table 3 shows catastrophic spending and share by the sex of the head of household and the regions for all

Table 3: Catastrophic Expenditure (CE) and Share of Catastrophic Expenditure (SCE) by region and sex of household head (n = 7448)

		Male		Female		Total	
Region	n	CE (%)	SCE (%)	CE (%)	SCE (%)	CE (%)	SCE (%)
Western	686	25.77	11.95	49.24	11.41	32.51	11.71
Central	572	44.04	15.09	66.83	16.59	52.45	15.76
Gt Accra	982	44.43	29.13	71.13	24.35	52.34	26.99
Volta	625	28.93	12.05	50.54	11.06	35.36	11.61
Eastern	747	21.8	10.34	38.06	11.06	27.17	10.66
Ashanti	1393	15.81	14.04	31.73	17.07	21.03	15.39
BrongAhafo	721	12.63	5.79	22.26	6.24	15.81	5.98
Northern	748	1.67	1.04	1	1.06	2.67	1.05
Upper East	536	0.67	0.28	7.78	0.82	1.87	0.52
Upper West	438	0.79	0.28	4.91	0.35	1.37	0.31
Total	7448	-	100	-	100	-	100

Authors' computation using GLSS (Round V)

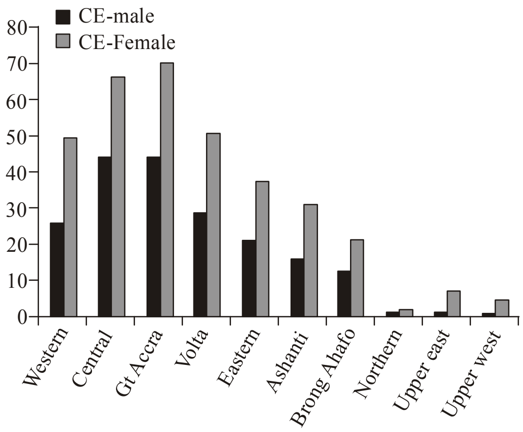

Fig. 1: Bar chart of the Catastrophic Expenditure (CE) in percentage by the regions and the gender

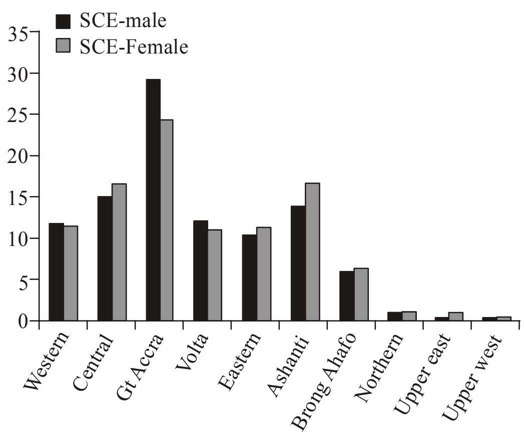

Fig. 2: Bar chart of the Share of Catastrophic Expenditure (SCE) in percentage by the regions and the gender

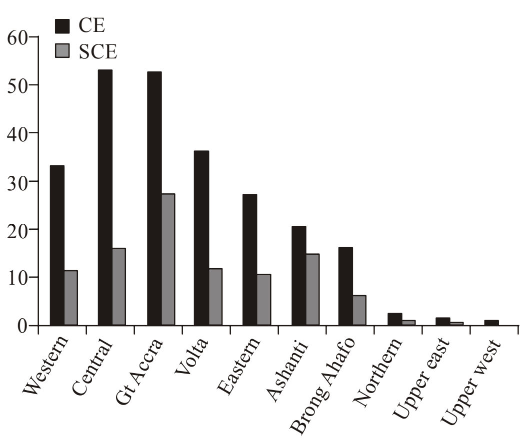

Fig. 3: Bar chart of the share of Catastrophic Expenditure (CE) and Share of Catastrophic Expenditure (SCE) in percentage by the regions

households in the data set. Figure 1 gives a graphical view of catastrophic spending by sex of head of household and the regions for all households. Catastrophic spending on health is highest (irrespective of gender) in the Greater Accra (52.34%) and Central (52.44%) regions of Ghana whilst catastrophic spending is least in the Upper West (1.37%) region of Ghana. For females, the region with the highest proportion of catastrophic expenditure is the greater Accra region with 71.13% whilst their male counterparts also have the highest proportion being 44.44%. Figure 2 shows a bar chart of the Share of Catastrophic Expenditure (SCE) in percentage by the regions and the gender. Greater Accra recorded the highest Share (26.99%) of catastrophic expenditure on health and is followed by the central (15.75) and Ashanti (15.39%), respectively. Upper West region recorded the lowest share (0.32%) of catastrophic expenditure on health. For both sexes the highest share of catastrophic expenditure is found in the Greater Accra and Central regions. From the GLSSS5 study, it can be inferred that the country has about 25.56% of its households spending catastrophically on health. Figure 3 together gives a side by side pictorial view of the every region's catastrophic spending and Share of catastrophic expenditure.

DISCUSSION AND CONCLUSION

The methodology suggests another procedure that can be adopted in research on catastrophic expenditure in health. The ability of the procedure to correctly predict at least 78% of a household's likelihood to spend on health for real data is encouraging. Apart from determining the catastrophic spending status of households who did not report on health expenditure during the survey, the methodology may be used to monitor catastrophic health spending of households until the next survey on the socio-economic status of households is done.

In general, female headed households spend more catastrophically on health when compared to their male counterparts and this is consistent for all the regions of Ghana. This supports the findings most national surveys that female-headed households are generally low income earners (Ghana Statistical Service, 2006, 2003). The cultural phenomenon of the female to take care of the sick at home might also be a contributing factor to the catastrophic expenditure on health by a larger proportion of females when compared to their male counterparts.

The region with the largest proportion of households with catastrophic spending on health is the Greater Accra region. Central, Ashanti, Volta, Western and Eastern follow in that order. On the contrary, the five other regions that span the northern part of Ghana seems to report relatively smaller segment of their households as having spent catastrophically on health. The reverse is what was expected since these regions are considered poorer than those reporting higher proportions of households to have spent catastrophically on health. The larger populations in these regions together with the broad spectrum of the income ranges of the regions resident might be a key factor. The GLSS 5 report also indicates that the number of health facilities located in the regions found at the northern part of Ghana are lesser when compared to those in the southern part of the country and hence residents in those regions are less likely to visit facilities when ill. According to Coulombe (2004), households found in the Savannah zone are the poorest of which the three regions in the northern part of Ghana are found in the Savannah zone. One might just conclude that poorer households are less likely to spend catastrophically on health. The reverse is true since a head of household's ability to spend catastrophically on health is dependent on his ability to afford such a healthcare.

Just like other publications on health expenditure, this study has some limitations. The conclusions made from this methodology were based only on the data captured by the GLSS 5. Unfortunately, the data captured under the GLSS 5 failed to capture data from expenditure on transportation to health facilities or centers, expenditure on traditional or herbal health care, which forms an integral part of health care delivery in such a developing country like Ghana. Further validation work should be done on the model using out-of-sample data such as the next round of the GLSS survey, to further confirm the strength of the model and to determine its structural stability over time. However, this study clearly demonstrates that the methodology being proposed is efficient in classifying all households as having catastrophically spent on health or otherwise irrespective of the reported expenditure (including zero) on health for the survey period and may reveal more on dynamics of catastrophic spending should socio-economic surveys capture the kinds of information discussed in the limitations.

REFERENCES

Alan, A., 2012. An Introduction to Categorical Data Analysis. 2nd Edn., Wiley Series in Probability and Statistics, pp: 137-172.

Berki, S., 1986. A look at catastrophic medical expenses and the poor. Health Affairs, 5(4): 138-145.

Blankenau, J., J.M. Bailey and J. Hudson, 2009. The Causes and Consequences of the Rural Uninsured and Underinsured. Center for Rural Affairs, Newsletter.

Coulombe, H., 2004. Construction of the Census-based Ghana Poverty Map: District Level Results: Ghana's Economy at the Half Century. ISSER and Cornell University, pp: 18-20.

Daneshkohan, A., M. Karami, F. Najafi and M.B. Karami, 2011. Household catastrophic health expenditure. Iran. J. Public Health, 40(1): 94-99.

Filmer, D. and L. Pritchett, 2001. Estimating wealth effects without expenditure data-or tears: An application to educational enrollments in states of India. Demography, 38(1): 115-132.

Ghana Statistical Service, 2006. Ghana Demographic and Health Survey (GDHS), report. Accra, Ghana.

Ghana Statistical Service, 2003. Core Welfare Indicators Questionnaire (CWIQ) Survey, Accra, Ghana.

Johnson, R.A. and D.W. Wichern, 2007. Applied Multivariate Statistical Analysis. 6th Edn., Pearson Prentice Hall, Upper Saddle River, NJ, pp: 575-582.

Kim, Y. and B. Yang, 2011. Relationship between catastrophic health expenditures and household incomes and expenditure patterns in South Korea. Health Policy, 100: 239-246.

Lara, J.L.A. and F.R. Gómez, 2011. Determining factors of catastrophic health spending in Bogota, Colombia. Int. J. Health Care Fi., 11: 83-100.

Lave, J.R., C.R. Keane, C.J. Lin, E.M. Ricci, G. Amersbach and C.P. La Vallee, 1998. The impact of lack of health insurance on children. J. Health Soc. Policy, 10(2): 57-73.

Nguyen, H., R. Ivers, S. Jan, A. Martiniuk and C. Pham, 2013. Catastrophic household costs due to injury in Vietnam. Injury, Int. J. Care Injured, 44: 684-690.

Pal, R., 2010. Analysing catastrophic OOP health expenditure in India: Concepts, determinants and policy implications. India Ghandi Institute of Development Research, Mumbai.

Russell, S., 1996. Ability to pay for health care: Concepts and evidence. Health Policy Plann., 11(3): 219-237.

Wagstaff, A. and E. van Doorslaer, 2003. Catastrophe and impoverishment in paying for health care: With applications to Vietnam, 1993-98. Health Econ., 12: 921-934.

Wagstaff, A., 2008. Measuring Financial Protection in Health. World Bank Policy Research Working Paper No. 4554, The World Bank, USA.

Wyszewianski, L., 1986. Financially catastrophic and high cost cases: Definitions, distinctions and their implications for policy formulation. Inquiry, 23(4): 382-394.

Xu, K., D. Evans, K. Kawabata, R. Zeramdini, J. Klavus and C. Murray, 2003. Household catastrophic health expenditure: A multicounty analysis. Lancet, 362: 111-117.

Xu, K., D. Evans, P. Kadama, J. Nabyonga, P. Ogwal, P. Nabukhonzo and A. Mylena, 2006. Understanding the impact of eliminating user fees: Utilization and catastrophic health expenditures in Uganda. Soc. Sci. Med., 62: 866-876.