

---

# Expected Poverty Approach to Estimate Poverty Vulnerability-Considerations for Doing Estimation

Md. Touhidul Alam<sup>1</sup>, Mohammad Ehsanul Kabir<sup>2</sup> &  
Pk. Md. Motiur Rahman<sup>3</sup>

<sup>1</sup>Lecturer, Dhaka School of Economics, University of Dhaka, Bangladesh

<sup>2</sup>Assistant Professor, Dhaka School of Economics, University of Dhaka, Bangladesh

<sup>3</sup>Supernumerary Professor, Institute of Statistical Research and Training, University of Dhaka, Bangladesh

---

**Abstract:** Households, in rural Bangladesh, have differentiated level of susceptibilities in response to shocks. Households with lack in physical and human capital are more susceptible to shocks and are less likely to recover from them. This leads households poverty vulnerable with shortfall of socially accepted welfare by making large variation in income. This overarching aspect of vulnerability to shock has been taken in consideration in estimating poverty vulnerability using cross sectional data with a simple method called vulnerability as expected poverty (VEP) which is subject to model estimation. But, doing estimation of such model requires some considerations which are subject matter of this study. This study, from a survey data, introduces loss due to cyclone as climate change shock and estimate VEP model for Satkhira, Bangladesh using a cross sectional survey of 154 households .

**Key Words:** Poverty Vulnerability, VEP model, Shocks

## 1. Introduction

Household poverty is an ex-post measure of household wellbeing. Again, whether a household is prone to be poor i.e., poverty vulnerable, is an ex-ante measure of household well-being. For many policy purposes to reduce poverty, it becomes important to find out the ex-ante risk that a household will, if currently non-poor, fall below the poverty line, or if currently poor, will remain in poverty. Now, the question is that how a non-poor household can be poverty vulnerable and fall prey to poverty. Households have differentiated level of susceptibilities to response shocks (both idiosyncratic and covariate). So, people suffer from large income variation when they face shocks. In turn, many of them, who have less susceptibility to response shock, become poverty vulnerable. Ideally, poverty vulnerability assessment largely warrants panel data due to presence of inter-temporal variation from year to year. Most of the

methods to assess poverty vulnerability (e.g., Vulnerability as expected low utility (VEU), vulnerability as uninsured exposure to risk etc.) rely on panel data. But, such data are seldom available in developing countries with required features eg. data on shock and loss due to shock and relevant coping strategies. Hence, the second best method is to assess poverty vulnerability through the method of vulnerability as expected poverty (VEP) as devised by Chaudhury and others [4]. This method works quite well with cross-sectional data and in this paper, methodical considerations have been reflected in estimating VEP model proposed, conceptualizing vulnerability as the probability that a household, regardless of whether it is poor today, will be consumption poor tomorrow.

## 2. Methodology

### 2.1 Data and their Characteristics

Data, used in this study, has been taken from the survey under the research project “Chronic Poverty in Rural Bangladesh-2009”, which was carried out under the supervision of Professor P.K. Md. Motiur Rahman, Institute of Statistical Research and Training (ISRT), university of Dhaka.

In its sampling design, a three-stage stratified random sampling design was followed for selecting the final sampling unit (FSU). At the first stage, eight of the most vulnerable and least developed districts were selected. At the second stage, 32 villages were selected at random and at the third stage, households were selected at random. At the third stage, the criterion of stratification was the economic status of households, i.e. (i) non-poor, (ii) descending non-poor (iii) ascending poor and (iv) chronically poor [1].

The survey was carried out during January 28 – February 28, 2010 on 1,282 rural households (320 non poor households, 225 ascending poor, 227 descending non-poor and 510 chronically households) selected at random from 32 villages

spread over rural areas of 8 districts of Bangladesh-Panchagarh, Kurigram, Sherpur, Madaripur, Satkhira, Barguna, Sunamgonj and Khagrachari. But, in this study, 154 households from only Satkhira have been selected because of its being victim of Cyclone Aila in 2009 so that shock and related coping strategy can be modeled in vulnerability estimation. Table-1 indicates list of location considered in the survey for Satkhira.

**Table 1: List and location of selected villages**

Sl. no.	District	Upazilla	Union	Village
01	Satkhira	Ashasuni	Ashasuni	Shitol Pur
02		Kolaroa	Jogi Khali	Paik Para
03		Sadar	Bolle	Mukun dapur
04		Shyam nagar	Munshiganj	Mou khali

Source: [1]

## 2.2 Poverty Line

Assessment of poverty vulnerability pivots poverty line and hence it is very important to refer a reliable poverty line. It acts as a threshold of wellbeing of the households. However, it has been frequent in the literature to anchor poverty lines in nutritional requirement, basically in calorie norm [2]. Within the domain of calorie norm, we can use Direct Calorie Intake (DCI) method and Food Energy Intake (FEI) method. But, Cost of Basic Need (CBN) method is a method of computing poverty lines that respects local consumption patterns and prices while retaining the normative calorie anchor [3]. Hence, we have taken CBN based poverty lines (upper poverty line, **Zu** and lower poverty line, **Zl**) into consideration for our study. Bangladesh Bureau of Statistics (BBS), in its Household Income and Expenditure Survey, 2010 (HIES, 2010), referred CBN based poverty line as in the following table-2:

**Table- 2: CBN based poverty line of HIES- 2010**

Division	zf	znfl	znfu	Zl	Zu
Khulna	884	308	551	1192	1435

Source: [6]

Here, zf = Food poverty line, znfl = lower non-food allowance, znfu = upper non-food allowance, zl = Lower poverty line, zu = Upper poverty line

## 2.3 Vulnerability as Expected Poverty Model

Estimation of poverty vulnerability is drawn extensively from methodology developed by Chaudhuri [4]. The underlying idea of their methods is to construct appropriate probability

distribution of consumption expenditure conditional on household characteristics. The method is discussed in detail below:

Let  $y_n$  be the per capita consumption level of  $i^{\text{th}}$  household at time  $t$  and  $z$  be the poverty line (CBN based poverty line in our study). Then the household is poor if

$$y_{it} < Z$$

Again, let  $V_{it}$  is the vulnerability index of the  $i^{\text{th}}$  household at time  $t$  given the probability(ex-ante) that the household will be poor in time  $t+1$  (next year) is as  $V_{it} = P_{\text{rob}}(Y_{i,t+1} < Z)$ .

Now, for cross section data let us assume that for the  $i^{\text{th}}$  household the stochastic process generating income/consumption expenditure is as follows:

$$\log y_i = X_i \beta + \epsilon_i, i=1,2,3,\dots,n$$

.....(1)

where  $y_i$  is the per capita income/ consumption expenditure (welfare indicator);  $X_i$  represents a set of observable household characteristics which are assumed to be the determinant of a household's vulnerability index;  $\beta$  is a vector of parameter and  $\epsilon_i$  represents other shocks or crisis that differentiates per capita income/expenditure between two households having same characteristics. Also  $\epsilon_i$  being the usual disturbance term used in the statistical model along with the additional assumption of zero mean to secure the unbiasedness property of the estimates of  $\beta$ .

However, the independence and normality assumptions of  $\epsilon_i$  are also obligatory, so that any objectionable systematic pattern doesn't endanger our estimate's validity. But, most importantly, the  $\epsilon_i$ s are not supposed to be identical among households at various levels of expected income/expenditure. On the other hand, the assumption of homogeneity of variance of  $\epsilon_i$  is not expected to be maintained.

Moreover, variance of  $\log y_i$  is less than that of  $y_i$ , and variance of  $\log y_i$  equals the variance of  $\epsilon_i$ . However, the problem of heteroscedasticity is not as bad for  $\log y_i$  as of simple model with  $y_i$ . This is one of two basic reasons why we prefer the model with  $\log y_i$ . The other reason is to acquiesce the model with the assumption of following normal distribution of explained variable for the application of least square estimation procedure.

To dissolve the quandary of heteroscedasticity further, several parametric models are suggested concerning the variance of  $\epsilon_i$ , denoted by  $\sigma^2$ , relating to the set of observable household characteristics  $X_i$ . for simplicity, we can select the following model:

$$\sigma_i^2 = \alpha X_i \dots \dots \dots (2)$$

The estimation of these parameters  $\alpha$  and  $\beta$  from models can be carried out by the three-step feasible generalized least squares (FGLS) procedure, as suggested by Amemiya [5]. The estimation procedure is outlined as below:

First, we apply the ordinary least square (OLS) method on model (equation 1). At this stage, we need to construct the estimated residual index using parameter values.

Second, we use the squared estimate of  $\varepsilon_i$ /OLS as the crude estimate of the variance of  $\varepsilon_i$ , that is,  $\sigma^2$ . These are used in model (equation 2) for the parameter estimation purpose. Then the model for estimation becomes as follows:

$$\widehat{\varepsilon}_i^2 \text{ OLS} = \alpha X_i + \xi_i \dots \dots \dots (3)$$

Here,  $\xi_i$  is the disturbance term for the model which allows the lack of homoscedastic problem and another major issue, the measurement error in the survey data which inflates volatility.

Now, using the ordinary least square (OLS) method on (equation 3), we get the estimate  $\widehat{\alpha}_{OLS}$  of the parameter  $\alpha$ .

Next, using the estimate, we transform the model as:

$$\frac{\widehat{\varepsilon}_i^2 \text{ OLS}}{\widehat{\alpha}_{OLS} X_i} = \alpha \left[ \frac{X_i}{\widehat{\alpha}_{OLS} X_i} \right] + \frac{\xi_i}{\widehat{\alpha}_{OLS} X_i} \dots \dots \dots (4)$$

Now, the parameter is estimated once more by using ordinary least squares. This estimate has the enviable statistical property that is an asymptotically efficient FGLS estimate, which solves our inefficiency problem as a consequence of heteroscedasticity. Using the estimate in model (equation 2), we get the consistent estimate of  $\sigma^2$  as,

$$\widehat{\sigma}_i^2 = \widehat{\alpha}_{FGLS} X_i \dots \dots \dots (5)$$

Similarly, to estimate  $\beta$ , we construct another form of the model using the estimate of  $\sigma$  as follows:

$$\frac{\log y_i}{\widehat{\sigma}_i} = \beta \left[ \frac{X_i}{\widehat{\sigma}_i} \right] + \frac{\varepsilon_i}{\widehat{\sigma}_i} \dots \dots \dots (6)$$

The estimate  $\widehat{\beta}_{FGLS}$  also has the asymptotic efficiency property.

Therefore, by using the FGLS estimate of  $\beta$  and  $\alpha$ , we get the expected value of log per capita income/consumption expenditure and its variance is as follows:

$$\widehat{E}(\log y_i | X_i) = \widehat{\beta}_{FGLS} X_i \dots \dots \dots (7)$$

and the variance of log per capita income/expenditure:

$$\text{Var}(\log y_i | X_i) = \widehat{\alpha}_{FGLS} X_i, \text{ for each household } i \dots \dots \dots (8)$$

Now, assuming that income/expenditure follows log-normal distribution, we can use these estimates to form an estimate of the probability that a household with the characteristics  $X_i$  will be poor after a shock or crisis. In other words, a

household's vulnerability level ( $v_i$ ) which is originally corrected by the expected value of the model and scaled up by its volatility is obtained as follows:

$$\widehat{V}_i = \widehat{\text{Prob}}(\log y_i < \log Z | X_i) = \Phi \left[ \frac{\log Z_i - \widehat{E}(\log y_i | X_i)}{\sqrt{\text{Var}(\log y_i | X_i)}} \right] = \theta \left\{ \frac{\log Z - X_i \beta}{\widehat{\sigma}_i} \right\} \dots \dots \dots (9)$$

where  $\theta(\cdot)$  is the cumulative density of the standard normal distribution and Z is poverty line.

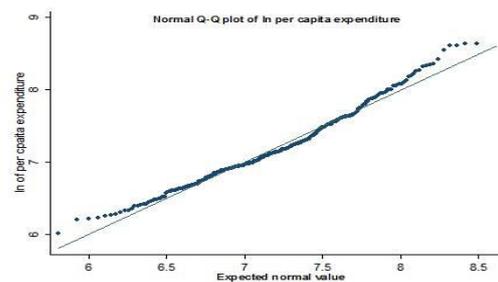
The value,  $\widehat{V}_i$  lies between 0 - 1 which captures the predicted probability of poverty for every household. Though the choice of vulnerability threshold is quite arbitrary, it is standard practice to define vulnerability, as the ex-ante risk (probability,  $\widehat{V}_i \geq 0.5$ ) that a household will, if currently non-poor, fall below the poverty line, or if currently poor, will remain in poverty.

### 3. Data Exploration And Assumption Checking

The assumption of randomness of the sample is essential for all tests and it is checked for the study through the 'Run test'. The test results showed that the sample data is random.

To test the normality assumption (test of goodness of fit) of dependent variable, log of per capita expenditure, normal Q-Q plot has been used. The normal Q-Q plot of log per capita monthly expenditure in figure-1, shows that observed points almost fall into normally distributed straight line.

**Figure-1: Normal Q-Q plot of log per capita expenditure**

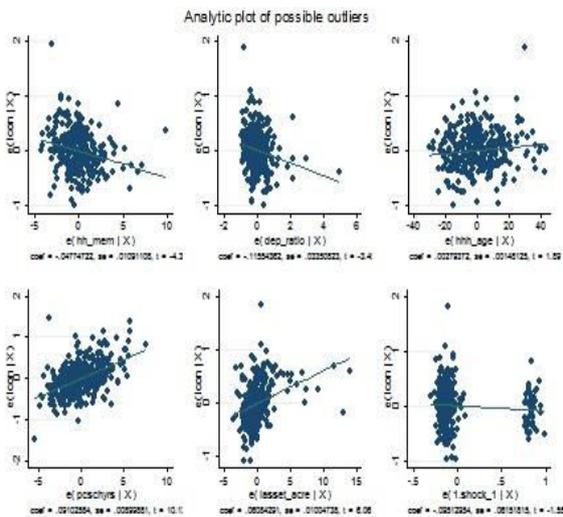


**Source: Own compilation**

Therefore, this study can perform parametric tests with the dependent variable of log per capita monthly expenditure.

In case of detecting outliers both analytic plot and statistical tools e.g., Cook's Distance and leverage values have been used. Analytic plots of possible outliers in figure- 2, suggest that all data points seem to be in range. Perhaps, there is no incidence of outliers in our data.

Figure- 2: Added-variable plots



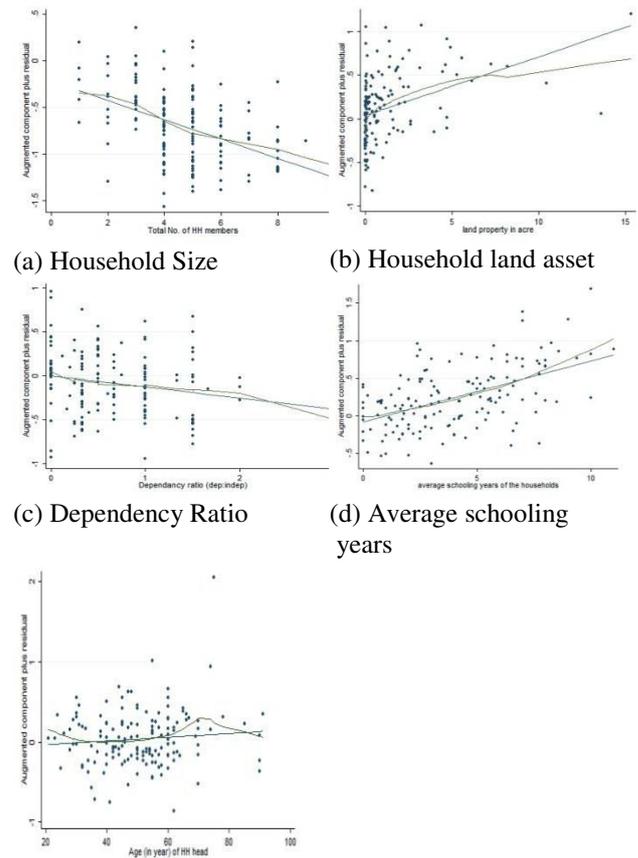
Source: Own compilation

Again, from Cook's D distance, it is apparent that a few values is more than  $(4/310=0.0129)$  but none of them are greater than 1 i.e.,  $D > 1$ . Hence, it can be said that, there is no big outlier problem. Moreover, from leverage values, no value is found to be over 0.05 or closer to 1. Hence, using the rule of thumb, we can say that this study is not overarched with outlier problem.

#### 4. Econometric Modeling Of Poverty Vulnerability

The focal assumption of Vulnerability as Expected Poverty (VEP) approach is that much of the variation across the households, vis-à-vis inter temporal variation in case of panel data, can be attributed to the observable characteristics of the households. Predicted probabilities of poverty are then generated based on household characteristics e.g., demographic and socio-economic indicators and shock and options available to mitigate that shock i.e., coping strategies. However, in estimating equation (1), it is required to find whether the regressors are linearly related to regressand. We can do it using augmented component-plus-residual plot. Augmented component-plus-residual plot in figure-3, suggests all linear relationship between regressors and the regressand.

Figure- 3: Augmented component-plus-residual plot



Source: Own compilation

Again from Table- 3, to test multi collinearity, from VIF and 1/VIF, it can be concluded that the model is free of multi collinearity problem.

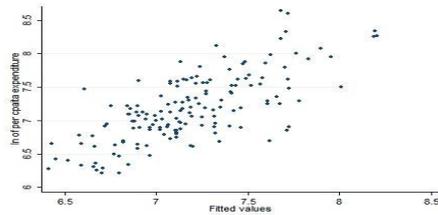
Table- 3: VIF and 1/VIF of covariates

Variable	VIF	1/VIF
Household Size	1.41	0.7084
Dependency Ratio	1.37	0.7284
Sex of household head dummy (female)	1.28	0.7806
Age of household head	1.06	0.9446
Average schooling years	1.41	0.7097
Household land asset	1.18	0.8479
Shock dummy (yes)	2.36	0.4241
Coping strategy dummy (Yes)	2.15	0.4650

Source: Own compilation

A 45 degree pattern from figure-4 of observed and predicted log per capita expenditure suggests that, the model seems to be doing a good job in predicting log per capita expenditure.

**Figure-4: Scatter plot of observed and predicted log per capita expenditure**

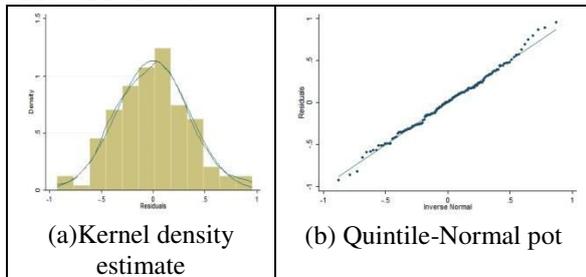


Source: Own compilation

Again, Ramsey RESET test (P-value, 0.7961), in present case suggests the independence of error term in the model i.e., there are no omitted variables in the model.

As usual, from figure-5, from kernel density and quintile-normal plot of estimated residuals, it is apparent that the residuals are normally distributed.

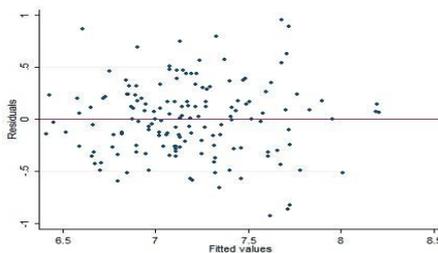
**Figure 5: Kernel density and Quintile-Normal Plot of residuals**



Source: Own compilation

We can't expect homoscedasticity in cross sectional data. So, it is very important to check heteroscedasticity when estimating equation (1) using OLS. About heteroskedasticity from figure-6, it is hard to conclude about any systematic pattern (outward-opening funnel, double bow or non-linear).

**Figure 6: Residuals vs predicted values (Y hat)**



Source: Own compilation

But, from White's pure heteroskedastic test, ( $H_0$ : Homoskedasticity & P-value < 0.5) it may be concluded that, our model is not free of heteroskedastic problem. Hence, from 3-step FGLS as suggested in the model, results have been incorporated in table-4.

**Table 4: Results of log per capita expenditure**

Variable	Coefficients	Std. error	t-value	95% CI	
				Lower limit	Upper limit
$\beta_{FGLS}$					
Household size	0.039*	0.02	1.7	-.004	.08
Dependency ratio	-0.41**	0.08	-5.02	-.577	-.25
Sex of HH dummy (female)	0.44***	0.10	4.16	.231	.649
Age of HH	0.01***	0.02	8.37	.014	.023
Household average schooling yrs	0.08***	0.01	5.24	.053	.11
Household land asset (in acre)	-.02***	0.01	-1.52	-.065	.008
Shock dummy (yes)	.44***	0.11	3.74	.21	.68
Coping Strategy dummy (yes)	-1.2***	0.15	-8.06	-1.57	-.95
Total HHs	154				
Adj- R <sup>2</sup>	0.69				

Source: Own compilation

Legend: \* p<0.10, \*\*p<0.05, \*\*\*p<0.01

Now, using two poverty line (lower & upper poverty line based on Cost of Basic Need approach for Khulna), from equation (9), probability of each household of being poor or remaining poor in near future can be calculated giving each household an index of vulnerability ranging from 0-1.

## 5. Conclusion

Poverty is a stochastic phenomenon. Today's non-poor household may fall prey to poverty if it faces unbearable shock by itself. So, it is important, for forward-looking anti-poverty policy interventions, to look for households who are likely to be poor in the future and their relevant characteristics. But, while doing estimation using ordinary least square (OLS), it requires following different assumptions, applied in this paper, which we need to be aware of.

## 6. References

- [1] **Rahman P.M.M., Matsui N., Ikemoto Y., 2009.** *The chronically poor in Rural Bangladesh: livelihood constraints and capabilities*, 1<sup>st</sup> edn., Routledge, London
- [2] **Gaiha, R. and K. Imai. 2008.** *Measuring Vulnerability and Poverty: Estimates for Rural India*, Research Paper No. 2008/40, UNUWIDER
- [3] **Jones, R.P.** *Can we set Poverty Lines using Calorie Norms in India or Bangladesh?* Paper 4 (draft for comment only) submitted to Economic Development and Cultural Change  
<[http://www.uea.ac.uk/~d280/other/india/Paper 4 - Can we Set Poverty Lines Using Calorie Norms.pdf](http://www.uea.ac.uk/~d280/other/india/Paper_4_-_Can_we_Set_Poverty_Lines_Using_Calorie_Norms.pdf)>
- [4] **Chaudhuri, S., J. Jalan, and A. Suryahadi. 2002.** *Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia*. Discussion Paper #:0102-52, New York, Columbia University.
- [5] **Amemiya, T. 1977.** *The Maximum likelihood Estimator and the Non-linear three stage least squares Estimator in the General Non-linear simultaneous Equation Model*. *Econometrica* 45: 955-968.
- [6] **HIES, 2010.** Bangladesh Bureau of Statistics *Report of the Household Income and Expenditure Survey 2010*, Dhaka: Planning Division, Ministry of Planning