Construction of Poverty Index Identifying Potential Household Factors Impacting poverty in Pakistan

Wajahat Sarwar[1], Naeem Bukhari[2], Muhammad Akbar Awan[3], and Muhammad Altaf[4]
[1]Lecturer, Department of Allied Sciences, University of Haripur, Haripur, Pakistan, Mobile: +923465317998, Email: wajahat@uoh.edu.pk
[2]M. Phil Research Scholar, Centre for Human Genetics, Hazara University, Mansehra, Pakistan. Mobilel: +923315493532, Email: naeembukhari20@gmail.com
[3]Assistant Professor, Department of Statistics, International Islamic University, Islamabad, Pakistan. Mobile: +923009781648, Email: mmakbar@gmail.com
[4]Assistant Professor, Muhammad Altaf, Department of Basic Sciences, University of Engineering and Technology, Taxila, Mobile: +923315621107, Email: dm.altaf@uettaxila.edu.pk

Abstract: Currentlymultitudinous regional factors including cultural system, political system and economic systems are impacting poverty globally. Major causes impacting poverty includes lack of improvements in human capital, gender inequalities, political interventions and institutional policies. It’s a common phenomenon to overcome poverty making enough improvements in political, economic, institutional and social welfare systems. Current study aims to examines the variety of poverty led factors in Pakistani population and to highlight the poverty determinants role. The analysis is based on micro data taken from households’ integrated survey of Pakistan Social and Living Measurement (PSLM) 2013-14. Current analysis found that the region, household size, paid employees’ status and spouse employment has a significant and positive effect in poverty and age of household head, income, education, assets, spouse age and spouse education have a significant and negative effect in poverty.

Keywords: Construction, Household size, Pakistan, Poverty index.

Abbreviations:

__AUC__ = household size; __MCMC__ = Markov chain Monte Carlo; __MCSE__ = Monte Carlo Standard Errors; __MDG__= Millennium Development Goals; __MPI__ = Multidimensional Poverty Index; __ROC__ = Receiver Operating Characteristics; __2SRI__ = Two-Stage Residual Inclusion__; TSLS__= Two Stage Least Square.

Introduction

Poverty an attainment of the minimum level of living standardsinvolves the lack, deficiency, and loss of social, economic, cultural, political and other entitlements, rights and benefits [1]. Generally discussing, a poor household is defined as an entity without productive assets, income, and necessities. As Kanbur*et al*states that a poor in one dimension turns out to be poor in further dimensions; they cannot have a vast way into public services such as education, health, as compared to the non-poor[2].Inorderto measure poverty, it is necessary to study thoroughly the magnitude of a country to construct a statistical basis for analyzing the nature and characteristics of Poverty to evaluate monitoring trends systematically in the incidence of poverty, and to influence the structure of policies in order to address poverty problems [3]. Pakistan is a developing country and is facing the problem of poverty. Although different sectors are performing well but the growth rate of population has been nearly 3% for the last many years. Our report determines the figures of poverty by enlarging the investigation of poverty from household expenditures and consumption to several aspects of the betterment of human being.

Methodology

Theoretical Framework and Model Specification

According to Deaton and Muellbauertheory for rigorous data presentation empirical research techniques are followed [4-5].The components of the expected expenditure (Represented by X) include goods and services cost services($p_1,\ldots p_n$), household members with respect of their characteristics including such agesand gender($a_1,\ldots, a_m$), and the expected favourite utility level (U). which can be represented as follow;

$$Xh = E\,(U;\ p1\ldots, pn; a_1^h\ldots, a_m^h) \quad (1)$$

In this equation represents a household. It's true that household cannot be represented by pricing. As in majority of cases households face equal pricing issues.

Household characteristics are considered variable in comparison to expenditure functions. Whereas hidden utility can be measured by distinct expression of expenditure levels (X^h).

$$X^h = E\,(U;\ p_1\ldots,\ p_n; a_1^h\ldots, a_m^h)$$

$$= m(a_1^h\ldots, a_m^h;\ p_1,\ldots\ldots,p_n)E_1(U;\ p_1,\ldots\ldots,p_n) \quad (2)$$

E_1 represents existing per capita expenditure functions, m represents specific regulatory multiplicative factor. By dividing households with dissimilar characteristic comparative utility can be achieved mentioned as below.

$$X^h/\ m(a_1^h\ldots, a_m^h;\ p_1,\ldots\ldots,p_n) = E_1(U;\ p_1,\ldots\ldots,p_n) \quad (3)$$

Monotonic expenditures at the left-hand side of equation (3) are the measure of utility.

Exact cost of living can be determined by ratios of their expenditure functions. This can be achieved by (4) where the A and B superscripts represent the related prices in regions A and B.

True Cost-of-Living Index as;

$$\equiv \frac{E1(U; p_1^B\ldots p_n^{\mathbf{B}})}{E1(U; p_1^A\ldots p_n^A)} = P(U; p_1^B\ldots, p_n^{\mathbf{B}}; p_1^A\ldots, p_n^A). \quad (4)$$

For a specific desired region estimation of relative differences in prices is considered of prime importance and can be represent in a vector quantity as follows:

$$\overline{P}_j = \exp\left[\sum_{i=1}^{n} w_i \log\left({P_i^j}/{\overline{P}_j}\right)\right]\overline{P} \;=\; s_j\overline{P}\ . \quad (5)$$

Where $w_i$re presents expenditure rate at national level, i represents good and $\overline{P}_j$ is the price level vector for a specific region j. based on true cost of living can break into true cost of living index (4).

$$\frac{E1(U; p_1^B\ldots p_n^{\mathbf{B}})}{E1(U; p_1^A\ldots p_n^A)} = \frac{s_B E1(U; p_1^B/s_B\ldots p_n^{\mathbf{B}}/s_B)}{s_A E1(U; p_1^A/s_A\ldots p_n^A/s_A)} =$$

$$\frac{s_B}{s_A} P_r(U;\ p_1^B/s_B\ldots, p_n^{\mathbf{B}}/s_B; p_1^A/s_A\ldots, p_n^A/s_A) \quad (6)$$

The ratio $\frac{s_B}{s_A}$ in (6) controls for differences in price levels are independent of U, while $P_r(U;\ p_1^B/s_B\ldots, p_n^{\mathbf{B}}/s_B; p_1^A/s_A\ldots, p_n^A/s_A)$ controls for differences in relative household prices.

By Using equation (6) we can compare two different price levels impact on utility followed by two different regions.

$$\frac{X_B^h[m(.)s_B]}{P_r(U^B;.;.)} = \frac{E1(U^B; p_1^B\ldots p_n^{\mathbf{B}})}{P_r(U^B;.;.)} =$$

$$E_1(U^B;\ p_1^A\ldots, p_n^{\mathbf{A}}) \ \substack{\geq \\ <} E_1(U^A;\ p_1^A\ldots, p_n^A =$$

$$\frac{X_A^h/s_A}{m(.)} \text{ as } U^B \substack{\geq \\ <} U^A \quad (7)$$

Eq. (7) in case of many regions equation 7 can be followed to assess comparative utility level differences. while from an empirical point of view the functional form of $P_r(U^B;.;.)$ is unknown and U is unobservable. Intriguingly, in case of minor and major price difference in two adjacent areas equation 7 can be ex[pressed as; $p_1^B/s_B\ldots, p_n^{\mathbf{B}}/s_B$is equal to the normalized vector$p_1^A/s_A\ldots, p_n^A/s_A$. Thus,

$$\text{So, } E(U;\ p_1^B/s_B\ldots, p_n^{\mathbf{B}}/s_B) \approx E\left(U; \frac{p_1^A}{s_A}\ldots, \frac{p_n^A}{s_A}\right)$$

Which implies(U;.;.) $\approx$ 1. (8)

Various important household structural components estimation $X^h/[m(.)s_j]$ is predetermined. Which includes household compositional, regional, physical, human capitalization and expertis in a community ((R^h; K^h; E^h; and C^h). By Using Eq. (8) equation 7 can be expressed as;

$$\frac{X^h}{m(a_1^h\ldots a_m^h;\ p_1^j\ldots p_n^j)s_j} = F\left(a_1^h\ldots, a_m^h;\ R_1^h\ldots, R_r^h;\ K_1^h\ldots, K_k^h;\ E_1^h\ldots, E_e^h; C_1^h\ldots, C_c^h\right).\varepsilon \quad (9)$$

Here ε represents a multiplicative term accounting for unobserved effects.

According to [Deaton and Muellbauer (1986)]. No one can estimate exact m(.) without making some expected testable assumptions.

Specifically, multiplying both sides and taking logarithm F(.) of equation(7) becomes;

$$\log\ (X^h/s_j) = \sum_{j=1}^{r}\sum_{i=1}^{m}\alpha_{ij}a_i^h + \sum_{i=1}^{m}\beta_{ai}a_i^h + \sum_{i=1}^{r}\beta_{ri}R_i^h + \sum_{i=1}^{k}\beta_{ki}K_i^h + \sum_{i=1}^{e}\beta_{ei}E_i^h + \sum_{i=1}^{c}\beta_{ci}C_i^h + e. \quad (10)$$

Certain α'sparameters of m(.) and β'sparameters of F(.) and e=log(ε). When estimating (10) we can identify $\alpha_{ij} + \beta_{ai}$ within any region j, not α_{ij} or β_{ai} separately as well.

Estimations of parameters represented in equation (10).

A clear and specific poverty estimation model based on above all determinants;

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+\ldots\beta_{11}X_{11}(3.1)$$

Where Y represents Poverty status of the selected household.

Data and Description of Variables

Data of all variables are taken from the PSLM (2013-14) published by the Federal Bureau of Statistics, Government of Pakistan.

Construction of Poverty Index

The Global Multidimensional Poverty index uses different indicators about three dimensions, i.e. health, living circumstances, and education. Data of these indicators are taken from PSLM 2013-14 conducted and published by Federal Bureau of Statistics.

Alkire and Foster method of poverty measurement allowsus to estimate the multidimensional poverty index. Each household is divided into two categories as non-deprived and deprived. If the sum of deprivations of the household is greater than this cut-off point (33% or 1/3), it is measured as multidimensional poor, if less than this cut-off points then it is considered as non-poor [6-8].

Methodology of Analysis

Estimation is done by the classical and Bayesian framework.

Durbin–Wu–Hausman Test of endogeneity

By applying Durbin–Wu–Hausman test on our poverty index model hypothesis y = bX + e. where y represents a dependent variable, X is consideredvector of repressors, b represents vector coefficients and e represents expected the error. We have two estimators for b: b_0 and b_1. In case of null hypothesis both of these estimators are consistent, whereas, in case of alternative hypothesis, b_0 is consistent, whereas b_1 are not consistent. Now, the Wu–Hausman statistic becomes:

$$H = (b_1 - b_0)' (var(b_1) - var(b_0))^{\dagger}(b_1 - b_0)$$

Where † denotes the Moore Penrose pseudo inverse. Under the null hypothesis, this statistic has asymptotically the chi-squared distribution with the number of degrees of freedom equal to the rank of matrix $Var(b_0) - Var(b_1)$.

Derivation of the test can be shown as below. Assuming joint normality of the estimators

$$\sqrt{N}\begin{bmatrix} b_1 \\ b_0 \end{bmatrix} \xrightarrow{d} N\left(\begin{bmatrix} b \\ b \end{bmatrix}, \begin{bmatrix} var(b_1) & cov(b_1, b_0) \\ cov(b_1, b_0) & var(b_0) \end{bmatrix}\right)$$

Consider the function: $q = b_1 - b_0 \Rightarrow \operatorname{plim} q = 0$

By the Delta method

$$\sqrt{N}\,(q-0) \xrightarrow{d} N\left(0, \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} var(b_1) & cov(b_1, b_0) \\ cov(b_1, b_0) & var(b_0) \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$$

$$var(q) = var(b_1) + var(b_0) - 2\,cov(b_1, b_0)$$

By following obtained results after applying Hausman, that the covariance of an efficient estimator with its difference from an inefficient estimator yields zero;

$$var(q) = var(b_0) - var(b_1)$$

The chi-squared test is based on the Wald criterion can be express as follow;

$$H = \chi^2\,[K - 1] = (b_1 - b_0)' (var(b_1) - var(b_0))^{\dagger}(b_1 - b_0)$$

Where † denotes the Moore–Penrose pseudo inverse.

There are two methods commonly used to deal with endogeneity. Two-Stage Residual Inclusion:2SLS is only consistent when the Stage 2 equation is linear. If Stage 2 is nonlinear, use the two-stage residual inclusion (2SRI) method. Suppose we have a regression equation

$$y = a + b_1x_1 + b_2x_2 + e$$

Suppose that x_2 is endogenous to y. An instrumental variable is one that (a) is correlated with the endogenous variable x_2 (b) is uncorrelated with error term e (c) should not enter the main equation (i.e., does not explain y).

Stage 1: as in 2SLS, leading to predicted x_2^p and develop residuals $v = x_2 - x_2^p$

Stage 2: Predict y as a function of x_1, x_2 (not x_2^p) and the new residuals v:

$$y = f(a + b_1 x_1 + b_2 x_2 + b_3 v) + e$$

Where f(.) is a nonlinear function.Note that if Stage 2 is linear, then 2SRI yields the same results as 2SLS.

Two-Stage predictor substitution: 2SPS is the rote extension (to nonlinear models) of the popular linear two-stage least squares estimator.

Stage 1: as in 2SLS, leading to predicted x_2^p

Stage 2: the endogenous variables are replaced by first-stage predictors

$$y = f(a + b_1 x_1 + b_2 x_2^p) + e$$

Logistic regression under classical approach

CDF of logistic distribution can be used to set parameters, estimates and probability of the dependent variable and the maximum likelihood estimation method are applied to find the unknown parameter which maximizes the probability of attaining an observed set of the data.

$$P(y) = \frac{e^Y}{1+e^Y} = \frac{e^{\beta 0+\beta 1X1+\beta 2X2+\cdots+\beta 11X11}}{1+e^{\beta 0+\beta 1X1+\beta 2X2+\cdots+\beta 11X11}} = \frac{e^{X\beta}}{1+e^{X\beta}} \quad (3.2)$$

The first step in this method is to build a function, known as the Likelihood function. As Y_j is categorized as “0” and “1”, then statement for P in equation (3.2) offers the conditional probability P(Y=1/x) and the probability (1-P) offers conditional probability P(Y = 0 /x). Hence, for the pairs where Y_j = 1 will be contributed to likelihood function and for the pairs where Y_j = 0, 1-P will be contributed to the likelihood function. An easy method to express the likelihood functions from the expression.

$$L(\beta) = \prod_{i=1}^{n} P_i^{yi} . (1 - P_i)^{1-yi} \quad (3.3)$$

$$\mathrm{Ln}\{L(\beta)\} = \sum_{i=1}^{n} [yi.(X\beta) - \ln\{1 + e^{X\beta}\}] \quad (3.4)$$

The MLE obtain by taking the derivative of Equation (3.4) and equating it to zero, i.e.

$$\frac{\partial}{\partial\beta}\mathrm{Ln}\{l(\beta)\} = 0$$

we apply different diagnostics test, which includes linear relationship of independent vs dependent variables to assess multicollinearity relative to collinearity. Whereas multicollinearity depends on tolerance andinflation factor. Which can be represent as; $1\text{-}R^2$, once tolerance comes to 0, and the value of variance inflation becomes large and variables seems to closely relate to other variables.

For logistic regression models, we use Hosmer–Lemeshow test for goodness of fit. In the analysis of variance, one of the components into which the variance is partitioned may be a lack-of-fit.

ROC curve estimating fit of logistic regression model which is based on specificity (True Negative) sensitivity (True Positive) for all probable cutoff values. (1- specifities) are plotted on x-axis and sensitivities are plotted on the y-axis. This curve obtained is named as ROC curve. It helps to determine the accuracy of the diagnostic test. An AUC = 1 signifies a perfect test, i.e perfect specificity and perfect sensitivity and an AUC ≤ 0.5 represent a worthless test.

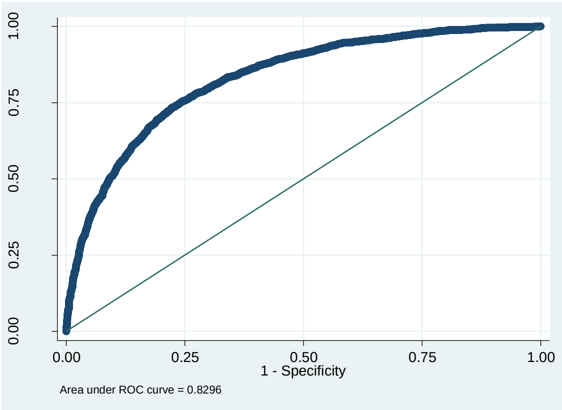

Fig. 4.1. ROC Curve.

Logistic regression under Bayesian approach

The Bayesian technique is based on Bayes Theorem is considered as superior to classical technique. While essential elements of Bayesian inference [9] 1. Construction of likelihood function.2.Formulation of distribution of unknown θ, P(θ).Derivation of posterior distribution P(θ/Y) by updating our beliefs about θ by combining information from likelihood function and prior distribution.
likelihood function derived for logit model is given as follow.

$$P(\beta/\text{data}) = \left[\left(\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{y_i} \times \left(1-\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{1-y_i}\right]$$

Meanwhile, distinct subjects are supposed, and they are self-governing from each other, the likelihood function over a data set of "n" subject is given as follows:

$$L(\beta/\text{data}) = \prod_{i=1}^{n} \left[\left(\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{y_i} \times \left(1-\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{1-y_i}\right]$$

Where $N_1 = \sum_i^n y_i$ is the number of (y = 1) and

$N_2 = \sum_i^n (1-y_i) = n - N_1$, is the number y = 0.$L(\beta/\text{data}) = \left[\left(\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{N_1} \times \left(1-\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{N_2}\right]$

(3.5)

Assuming that $B_j \sim N(\mu_j, \delta_j^2)$

Most mutual selections for μ is 0, and δ^2 's frequently select as a large enough to be considered as non-informative prior, corporate selection is being in the large form δ = 10 to 1000. So we use non-informative normal prior, N(0,1000) for all variables.

$$P(\beta) = \prod_{j=0}^{p} \frac{1}{\sqrt{2\pi\delta^2}} e^{-1/2}\left(\frac{\beta_j-\mu_j}{\delta_j}\right) \quad \textbf{(3.6)}$$

The posterior distribution is consequent by multiplying the prior distribution overall parameters by the filled likelihood function. The final form of the posterior distribution under normal prior to logit model is

$$P(\beta/\text{data}) \propto \left[\left(\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{N_1} \times \left(1-\frac{e^{X\beta}}{1+e^{X\beta}}\right)^{N_2}\right] \times \left[\prod_{j=0}^{p} \frac{1}{\sqrt{2\pi\delta^2}} e^{-1/2}\left(\frac{\beta_j-\mu_j}{\delta_j}\right)\right] \textbf{(3.7)}$$

In direction to estimate vector of the parameter or, there is a difficulty in algebraic expression.

By estimated Auto-Correlation Function, we check the level of dependence. The purpose of integer's j that provides evaluated the association between θ^k and θ^{k+j}. After burning this, correlation must depend on the lag j but not on k. It is calculated as the sample correlation between the pairs (θ^k, θ^{L+k}), L = 1,2……s^{-j}. If the autocorrelation is close to zero except to say, the 1st two j =1, 2 then we could take each third θ^{3k}, L =1, 2….s, after the burn in this sample be nearly uncorrelated its throw away information, unless there is extreme autocorrelation e.g. high association even with say j = 30, by drawing the graph in win bugs of ACF on the top to look good. Kernel density plots indicate that the Bayesian point estimates typical (posterior mean or posterior median) and the range among the 2.5th to 95.5th percentile signifies 95%.

The researcher is 95% confident that in the HBDI β_j lies within the HBDI. Consider, for illustration logistic regression model as (Y = Xβ) and that point of significant whether the explanatory variable must be involved. Consequently, the two models under attention remain

$$W_0 : \beta_j = 0$$
$$W_1 : \beta_j \neq 0$$

By using t-distribution properties of HBDI can be designed for β_j.

Results:

Estimation Results under Classical Framework

To estimate the parameters for logit model we use the MLE method which shows the ability of the households to the poverty status and compare the results to select the best estimation technique.

Estimation Results under Bayesian Framework

Model using the non-informative normal prior for logit model have been calculated in table 4.5 after running the MCMC algorithm for 11,000,000 iterations and discarding additional 1,000,000 iterations as a burn-in period and the thinning interval is 100 and the final posterior summaries on 1,000,000 samples are presented here. Based on 95% credible interval coefficient estimate of the region is significant as the interval does not contain zero and the positive relationship between the region of the household head and poverty. The expected odds ratio of the region is equal to 1.8366214.

Similarly, the expected odds ratio of age is 0.9888048. It indicates that increase of age of one year in the age of household head decrease the chances of poverty by 0.9888048 times or 1.2 %.

As Employment status divided into two categories, the first one is Employer (E_EM) and the second one is Paid Employee (E_PE). The expected odds ratio of the paid employee is 1.1836207.

The expected odds ratio of education of household head is equal to 0.9137602, which suggested that the educated household head have 0.9137602 or 8.7%less his or her chances of poor.

Based on 95% credible interval coefficient estimate of the Assets, the odds ratio of assets is equal to 0.2753641. As the increase of assets of household head decreases the poverty 0.2753641 or 72.47%. while, the odds ratio of spouse employment is equal to 1.2358869.

The expected odds ratio of spouse education is equal to 0.9031942, which suggested that the educated spouse of the household head have 0.9031942 or 9.69%reducetheir chances of poverty.

Diagnostic Test plots under Bayesian framework

Different diagnostics test including Kernal densities plot, Histograms, and Trace plots. Trace plots of all parameters presented in figure 4.2 which shows the good mixing and convergence of estimates [10]. Histograms of all parameters are presented in figure 4.3 show that all parameters follow the normal distribution. Kernal density plots of all parameters are presented in figure 4.4.

Conclusion:

Our analysis found that the region, household size, paid employees' status and spouse employment has a significant and positive effect in poverty and age of household head, income, education, assets, spouse age and spouse education have a significant and negative effect in poverty.

Current model is completely specified and evidence good overall fit. As a rule, that if tolerance is 0.1 or less, or equivalently VIF of 10 or more. All our variable's values of VIF are less than 10 which represent that there is no multicollinearity in the model. The plot of ROC curve and calculate AUC (area under the curve) which is typical performance measurement for the binary classifier. Our model's area under the curve is 0.8296 (82.96%) which is a good predictive ability.

References

1. World Bank. World Development Report 2000/2001: Attacking Poverty. 2000, New York: Oxford University Press.
2. Kanbur,R., & Squire, L. (1999). The evolution of thinking bout poverty; Exploring the interaction. Department of agriculture resource and managerial economics paper 92-94. Cornell University.
3. World Bank. World Report 2004, New York: Oxford University Press.
4. Angus S. Deaton and John Muellbauer*Journal of Political Economy*, (1986),Vol. 94, No. 4 pp. 720-744.
5. Lawrence Haddad and Ravi Kanbur. How Serious Is the Neglect of Intra-Household Inequality, Economic Journal, 1990), 100, (402), 866-81
6. Alkire, S and Foster, J.Counting and Multidimensional Poverty Measures', Oxford Poverty & Human Development Initiative (OPHI). 2007, Working Paper 7, University of Oxford.
7. Chakravarty, S., Mukherjee, D. and Renade, R. On the Family of Subgroup and Factor Decomposable Measures of Multidimensional Poverty', Research on Economic Inequality, 1998, 8: 175–194.
8. Foster, J., Greer, J. and Thorbecke, E. 'A Class of DecomposablePoverty Measures', Econometrica,1984,52(3): 761–766.
9. Hashmi, A. A., Sial, M. H., Hashmi, M. H., & Anwar, T.Trends and Determinants of Rural Poverty: A Logistic Regression Analysis of Selected Districts of Punjab [with Comments]. The Pakistan Development Review, 2008, 909-923.
10. UNDP Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development. 2010, New York: Palgrave Macmillan.