IARIW 2021

Monday 23 – Friday 27 August

Statistisk sentralbyrå
Statistics Norway

Measuring Multi-dimensional Vulnerability to Poverty:

A study through construction of index and decomposition of influences with the help of Fuzzy Logic and Artificial Intelligence

Sugata Sen

(Panskura Banamali College)

Paper prepared for the 36th IARIW Virtual General Conference
August 23-27, 2021
Session 18: Recent Experiences in Both Official and Academic Approaches to Measuring Poverty
Time: Thursday, August 26, 2021 [14:00-16:00 CEST]

Measuring Multi-dimensional Vulnerability to Poverty:

A study through construction of index

and

decomposition of influences

with the help of

Fuzzy Logic and Artificial Intelligence

***Dr. Sugata Sen*[1]**

Abstract: *The traditional multi-dimensional measures have failed to properly estimate the vulnerability of households towards poverty. The reasons behind this inability are the failure of the existing measures to recognise the graduality inside the concept of poverty and the ex-post consideration of the idea of poverty. So this work wants to develop a measure to estimate the vulnerability in an ex-ante multidimensional perspective with the help of fuzzy logic.Decomposition of the composite measure is done through artificial intelligence. To estimate and to decompose the vulnerability an integrated mathematical framework is developed.The constructed index is tested and the dimensional influences are compared under different socio-economic clusters.*

1. Introduction

The traditional multi-dimensional measures have failed to properly estimate the vulnerability of households towards poverty. The reasons behind this inability are the failure of the existing measures to recognise the graduality inside the concept of poverty and the ex-post consideration of the idea of poverty. So this work wants to develop a measure to estimate the vulnerability

[1]Associate Professor of Economics, Panskura Banamali College, INDIA 721152, sensugata@gmail.com,
M: +91 7501730545

in an ex-ante multidimensional perspective with the help of fuzzy logic. Decomposition of the composite measure is done through artificial intelligence. To estimate and to decompose the vulnerability an integrated mathematical framework is developed.The constructed index is tested and the dimensional influences are compared under different socio-economic clusters. The study is based upon primary data and machine learning methode has been used extensively to find the dimensional influences. The application of Shapley Value Machine Learning within the studies of multidimentional poverty has given the poverty studies a newer dimension.

2. Review of literature

The idea of poverty first appeared in the thinking of Confusious nearly 500 years before the birth of Christ. Poverty also appeared in the philosophy of Aristotol through the idea of ditributive justice. But poverty started to gain importance in social policy formulations from the period of Mercentailism. The Marcentailists conceived poverty as an essential pre-condition for economic development. But instead of viewing poverty as a social problem, it was viewed as a necessity towards the stability of the ongoing social structure. Though the idea of re-distributive justice was thought by Aristotol, it started to draw social attention with the writings of Charles de Montesquieu in the 18th century. Following the ideas of Montesquieu states started to share the burden of poverty with the emergence of modern capitalism. Through the coming out of modern welfare state and the growing importance of re-distributive justice the estimation of poverty emerged as an important activity of the state (Ravallion, 2016).

Philosophically it is accepted that the idea of poverty is related to the notion of well being. So the assesment of poverty refers to the achievement of certain level of well being (Sen, Poverty: An Ordinal Approach to Measurement , 1976). With time this concept of well being is transformed to the idea of being well. This transformation opened a whole new era in poverty research through capability approach. Poverty started to be appear as capability deprivation. This shift in the perspectives in the poverty research gradually accepts capability based indicators for poverty measurement. With the emergence of capability approach the poverty also becomes multi-dimensional. Naturally the multi-dimensional poverty (MP) index is right now the most accepted indicator of poverty (Sen, Inequality Reexamined, 1995).

The idea of multi-dimensional poverty index is carried forward by the development of axiomatic identification process by Chakravarty and Burguignon (Bourguignon & Chakravarty, 2003). Alkire and Foster developed an index on the basis of this identification process (Alkire & Foster, Counting and multidimensional poverty measurement, 2011). But the Alkire and Foster index have tried to distinguish the poor from the non-poor through a well defined threshold. Their argument is based upon the classical Boolean logic of either yes or no. In their idea the concept of poverty is rigoriously defined as an ordinary proposition. But the idea of poverty suffers from vagueness and naturally cannot be defined through a well defined cut-off. So developing and discussing the multidimensional poverty on the light of Boolean logic is not correct (Qizilbash, 2006). The idea of poverty suffers from vagueness and naturally cannot be defined through a well defined cut-off.

The graduality within a vague concept can well be represented by the idea of fuzzy logic (Zadeh, 1965). Naturally the logic of fuzzy sets started to reshape the discourses on poverty. Cerioli and Zani first attemted to use the fuzzy logic on the measurement of multidimentional poverty (Cerioli & Zani, 1989).They have tried to estimate the strength of poverty in each dimension through a membership function. Then aggregated the strength of every dimentions and normalised through the number of dimensions to get the overall strength of multidimentional poverty of each individual. Their idea has been improved further by Cheli and Lemmi through the idea of Total Fuzzy and Relative (TFR)

(Chelli & Lemmi, 1995). After that a voluminous research appeared in this field to illustrate different forms of membership functions. Many of them have also depicted the aggregation and inference issues related with this type of fuzzy indices (Martinetti, 2006).On the basis of these works Betti *et. al.*have tried to develop an idea called "Integrated Fuzzy and Relative" (IFR) approach to the analysis of fuzzy multidimensional poverty. Under IFR the authors have tried to deliver a more acceptable membership function. They have also disscussed different executable operations of fuzzy poverty sets (Betti, Cheli, Lemmi, & Verma, 2006). Chakraborty has provided an axiomatic interpretation of fuzzy multi-dimensional poverty index (Chakravarty, 2006).

Apart from measuring the composite effect of the multi-dimensional poverty, a large volume of research appeared on the decomposition of composite index. The sub-group decomposibilty of MP index became very important due to its special importance in policy formulation. The Shapley Value Decomposition is a solution in the findings of influencial causal factors. This type of decomposition takes into consideration the average of the marginal contributions of a factor under different combinations. To that respect, the concerned factor is first withdrawn from the model and the rest of the factors are permuted to form different distributions. Gradually, the withdrawn factor is added to each of the combination and the marginal contribution of the added factor in a specific distribution is counted. The average of marginal contributions of the stated factor from all the distributions is the influence of that very factor on the composite influence. In this way, the average contribution of all the factors are determined. The aggregation of all these factoral influences deliver the overall variation of the depedent factor. In this way, the Shapley Value Method decompose the overall variation of the composite dependent factor into the independent causal factors (Shorrocks, 2013).

To execute the Shapley Value Decomposition of the multi-dimensional poverty index machine learning can be used. Machine Learning (ML) is a technique of data analytics that instructs computer to learn from experience. Machine Learning algorithms use computational methods to “learn” information directly from data without depending on a pre-set equation as a model (Kubat, 2017). Understanding human learning and cognition is the aim of ML. Undoubtedly the key of human intelligence is their capability to learn. Thus an overall understanding of human learning process is very important to understand human intelligence. ML can help us to understand the basic principles of human learning and may lead to the invention of more fruitful learning techniques. Like human beings this technique also learns from the existing data and utilises that learning to draw conclusions from complex data (Theobald, 2017).

Supervised machine learning creates models that make predictions based on evidence in the presence of ambiguity. A supervised learning algorithm takes into consideration a known set of input output data and on the basis of that input output relationship trains a model to produce feasible predictions from a new data. Supervised learning applies classification and regression techniques to develop predictive models. Classification techniques under supervised learning predict discrete outputs. These models classify input data into different categories. Some important applications of classification procedure include medical imaging, credit scoring and speech recognition. Classification techniques are useful when the data are tagged, categorized or separated into specific groups or classes. Some common algorithms for classification techniques are support vector machine (SVM), boosted and bagged decision trees, discriminant analysis, Naïve Bayes, *k*-nearest neighbour, logistic regression and neural networks. Regression techniques under supervised learning predict successive reactions. This technique is used when the models are related with a data range and takes into consideration only the real numbers.

Some common regression algorithms are linear models, nonlinear models, boosted and bagged decision trees, neural networks and adaptive neuro-fuzzy learning (Chopra, 2018).

Unsupervised learning reveals hidden patterns or inherent structures in data. It is used to draw decisions from dataset consisting of input data without pre-set outputs. Clustering is the most common unsupervised learning technique. It is used for experimental data analysis to reveal hidden patterns or groupings within data. Some common applications for cluster analysis are market research, gene sequence analysis and object pattern recognition. Common algorithms for performing clustering include k-means and k-medoids, hierarchical clustering, Gaussian mixture models, hidden Markov models, self-organizing maps, fuzzy c-means clustering and subtractive clustering (Srinivasaraghavan & Joseph, 2019).

Shapley Value Machine Learning can be applied to decompose the individual influences of causal factors on multi-dimensional poverty through the techniques proposed by Shapley Decomposition. Shapley Value Machine Learning calculates the factoral contributions through Shapley decomposition method. One of the framework called SHAP executes the Shapley Machine Learning in reality. SHAP framework takes into consideration LIME, DeepLIFT and layer-wise relevance propagation. But in reality finding the exact outcomes through SHAP is really challanging. But the operational efficiency of SHAP can be improved by combining the Additive Feature Attribution Methods. Help of Max SHAP and Deep SHAP can also be taken for exact computation of SHAP outcomes (Lundberg & Lee, 2017).

Thus it appears that the estimation of multi-dimensional poverty through ex-post consideration delivers improper assessment. Again, this type of poverty also suffers from vagueness and naturally cannot be defined through a well defined cut-off. The graduality within a vague concept can best be represented through fuzzy logic. So it is beter to use membership function to forecast the extent of vulnerabilitty. The dimension specific levels of vulnerability can form the composite multi-dimensional vulnerability. To find the relative importance of the explanatory dimensions Shapley Value Decomposition can be applied. The decomposition method as prescribed by shapley can easily be executed through Unsupervised Machine Learning. The shapley value decomposition, with the use of unsupervised machine learning can help us to find the influences of different explanatory dimensions on the composite level of poverty. The proper estimation of relative importance of these factors can usher a new dawn in poverty eradication policies. SHapley Additive exPlanation (SHAP) can execute this decomposition through Local Interpretable Machine-agnostic Explanation (LIME) algorithms. This machine learning technology can play a potent role to find relative importance of different influencing factors. Thus the specific objectives of this study are:

3. Objectives

- Estimating the vulnerability to become poor multi-dimentionaly in ex-ante perspectives through fuzzy logic.
- Development of machine learning process to examine the dimensional influences on composite poverty.

4. Methodology

Poverty dimensions are selected following the OPHI methodology (Alkire & Santos, OPHI working paper no: 38, Acute Multidimensional Poverty: A New Index for Developing Countries, 2010). The strength of membership of a household to the poor set for each dimension is determined through appropriate membership function. Household level multi-dimensional vulnerability is determined through the average of dimension specific membership strengths of each household. Social vulnerability is determined through the average of household vulnerabilities. To find the dimensional influences on the composite multidimensional social vulnerability Shapley value decomposition is followed. This decomposition is executed through Shapley value machine learning (SHAP). To test the developed model data have been collected from 320 households. This set of households is chosen through stratified random sampling. Data are collected through questionnaire based household survey. The help of 10 point Likert scale was taken to quantify the respondent observations on different dimensions. 10 carried the highest weight towards poverty and 1 carried the lowest strength towards the membership. 0 is used to show the absence of membership towards the poor set. These dimensions were fixed through the OPHI framework (Alkire & Santos, OPHI working paper no: 38, Acute Multidimensional Poverty: A New Index for Developing Countries, 2010). To determine the strength of membership of a household under a particular dimension the chosen Likert scale value is deflated through 10. Thus it gave a partially continuous scale within 0 and 1, where 0 means the household is definitely not a member of the poor set related to the corresponding dimension and 1 means the household is definitely a member of poor set. The value in between 0 and 1 give the strength of partial membership of a household to a particular dimension.

5. Findings

Let there are i households where i=1,2,.n and j dimensions where j=1,2,…k. The performance of n households in k dimensions can be expressed through n×k real valued non-negative matrix. Here each row vector $y_i = \{y_{ij}\}$ interprets the performance of i^{th} household. The grade of membership to the poor set of the i^{th} household in j^{th} dimension is expressed through the membership function μ_p(ij). A household is definitely poor if his performance in dimension j is within 0 and y'_j. On the other hand if achievement is above y''_j then the individual is not poor on dimension j. For achievement between y'_j

and y''_j the membership function takes on values between 0 and 1 exclusively. More clearly it can be interpreted that if

- μ_p(ij)= 0 if the i^{th} household is certainly not poor in the j^{th} dimension.
- μ_p(ij)= 1 if the i^{th} household completely belongs to the poor set corresponding to j^{th} dimension.
- 0 <μ_p(ij) < 1 if the i^{th} household shows a partial membership to the poor set p of j^{th} dimension.

The grade of membership of the i^{th} household to the multi-dimensional poor set can be defined as

$$\mu_M(i) = \frac{\sum_{j=1}^{k} \mu_p(ij)}{k}$$

Then social vulnerability is

$$\lambda = \frac{\sum_{i=1}^{n} \mu_M(i)}{n}$$

If the desired value of λ is 0, then the difference between desired and observed vulnerability is λ. To decompose λ Shapley value decomposition has been used. This method calculates the average of marginal contributions of each dimension. In our model, in order to find the contribution of j^{th} dimension on the composite social vulnerability different combinations of K-1 dimensions are constructed. So, the total no of combinations among the different dimensions excluding the j^{th} dimensions is –

$$^{(K-1)}C_1 + {}^{(K-1)}C_2 + {}^{(K-1)}C_3 + {}^{(K-1)}C_4 + \ldots. + {}^{(K-1)}C_{K-1}$$

$$= \sum_{h=1}^{K-1} (k-1)_{C_h}$$

$$= \theta$$

If j^{th} dimension is added to each of the θ combinations we would get θ marginal contributions of the j^{th} dimension. Let, the marginal contribution of j^{th} dimension corresponding to the s^{th} combination from θ is φ_s^j. Then the set of marginal contributions of the j^{th} dimension arising out of θ combinations is-

$$H_j = (\varphi_1{}^j, \varphi_2{}^j, \ldots \varphi_s^j, \ldots, \varphi_\theta^j)$$

Then average of marginal contribution of the j^{th} dimension is

$$CON_j = \frac{1}{\theta}\sum_{s=1}^{\theta} \varphi_s^j$$

Let the polinomial form corresponding to the s^{th} combination under the j^{th} dimension can be chosen from the set of ψ alternative polinomials or from ψ^{sj}. Learning from the successive trials within ψ^{sj} the machine learning process chooses that polinomial from ψ^{sj} which minimises the error. In this way θ^j functional forms are determined. From these θ functions θ^j incremental influences are estimated.

This estimation of dimensional contributions is executed through Local Interpretable Machine-agnostic Explanation (LIME) algorithms. LIME deliberately perturbs a combination by accepting input variables from the neighbourhood and counts the effect of that perturbation on the output. Finally the relevance of the particular input is determined through the average of deviation in the output due to the perturbations. Technically here LIME initiates the process to locate

$$\min E_{\pi_{sj}} = \min[g(F_u) - f(F_v)]_{\pi_{sj}} \rightarrow F_{sj}^* \rightarrow \varphi_s^j$$

where u≠ v and u,v=1,2…..ψ. π_s is the neighbourhood of sth functional form under jth dimension. g(F_u) is the expected value and f(F_v) is the deserved value of the multi-dimensional poverty from a particular polinomial related to π_{sj}. F_{sj}^* is the chosen polynomial from ψ^{sj}. This process is used for all the combinations under θ^j to find θ^j incremental influences. Finally the average of the θ^j incremental incluences determines

$$CON_{\mathrm{j}} = \frac{1}{\theta}\sum_{s=1}^{\theta} \varphi_{\mathrm{s}}^{\mathrm{j}}$$

The application of the constructed index on our dataset finds that rural dwellers, female headed households, senior citizens, indigenous people, religious minorities and academically backward households are more vulnerable to multidimensional poverty in comparison to their counterparts. Decomposition of the index finds that the dimensional influences varied significantly under different socio-economic groups. The relative influence of health on vulnerability has been found to be highest in comparison to other dimensions for the female headed households, senior citizens and indigenous communities. The relative influence of education on vulnerability is highest within the set of dimensions for the religious minorities and the academically backward households. Interestingly it is also observed that the relative influence of health on the composite vulnerability also very high for the academically backwards. The relative influence of standard of living on composite vulnerability has been found to be highest for the rural dwellers.

6. Conclusion

This work has tried to develop an ex-ante idea of poverty instead of traditional ex-post considerations. To that respect the idea of poverty has been substituted by vulnerability to become poor. As the traditional uni-dimensional concept of poverty is becoming obsolete and the multi-dimensional poverty is gaining attention, this work has applied the ex-ante concept of vulnerability to multi-dimensional poverty. It is also assumed here that the idea of poverty as well as vulnerability suffers from vagueness. So the idea of fuzzy logic is applied here to develop a household level as well as social index of vulnerability to multi-dimensional poverty. To understand the dimensional influences the composite vulnerability index is decomposed with the help of Shapley value decomposition. This decomposition is implemented here with the help of Shapley value machine learning algorithms. The developed model is tested over a sample of 320 households. It is observed that rural dwellers, female headed households, senior citizens, indigenous people, religious minorities and academically backward households are more vulnerable to multidimensional poverty in comparison to their counterparts. Decomposition of the index finds that the dimensional influences varied significantly under different socio-economic groups. The relative influence of health on vulnerability has been found to be highest in comparison to other dimensions for the female headed households, senior citizens and indigenous communities. The relative influence of education on vulnerability is highest within the set of dimensions for the religious minorities and the academically backward households. Interestingly it is also observed that the relative influence of health on the composite vulnerability also very high for the academically backwards. The relative influence of standard of living on composite vulnerability has been found to be highest for the rural dwellers. Thus, this effort successfully establishes an departure of poverty measurements from ex-post considerations to ex-ante valuation with the substitution of poverty with vulnerability. The application of fuzzy logic in this work within the ex-ante measurement of poverty is also unique in the discourses of multi-dimensional poverty. Finally, the application of machine learning techniques in the decomposition of multi-dimensional poverty will open newer vistas within the ambit of computational economics.

References:

Agarwal, B. (1997). Bargaining and Gender Relations: Within and Beyond the Household. *Feminist Economics , 3* (1), 1-51.

Alkire, S., & Foster, J. (2009). *Counting and Multidimensional Poverty Measurement.* Oxford: OPHI.

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics* , 476-487.

Alkire, S., & Santos, M. E. (2010). *OPHI working paper no: 38, Acute Multidimensional Poverty: A New Index for Developing Countries.* Oxford: Oxford Poverty and Human Development Initiative.

Alkire, S., Kanagaratnam, U., & Suppa, N. (2018). *The Global Multidimensional Poverty Index (MPI):2018 Revision.* Oxford: OPHI.

Alkire, S., Roche, J. M., & Vaz, A. (2017). Changes Over Tme in Multidimensional Poverty: Methodology and Results for 34 Countries. *World Development* , 232-249.

Betti, G., Cheli, B., Lemmi, A., & Verma, V. (2006). Multidimentional and Longitudinal Poverty: an Integrated Fuzzy Approach. In A. L. Betti, *FUZZY SET APPROACH TO MULTIDIMENTIONAL POVERTY MEASUREMENT* (pp. 114-137). New York: Springer.

Bourguignon, F., & Chakravarty, S. R. (2003). The Measurement of Multidimensional Poverty. *The Journal of Economic Inequality* , 25-49.

Carroll, C. D. (1994). How does Future Income Affect Current Consumption? *The Quarterly Journal of Economics , 109* (1), 111-147.

Cerioli, A., & Zani, S. (1989). A Fuzzy Approach To The Measurement of Poverty. *Income and Wealth Distribution, Inequality and Poverty* (pp. 272-284). Italy: Springer.

Chakravarty, S. R. (2006). An Axiomatic Approach to Multidimentional Poverty Measurement via Fuzzy Sets. In A. Lemmi, & G. Betti, *FUZZY SET APPROACH TO MULTIDIMENTIONAL POVERTY MEASUREMENT* (pp. 49-72). New York: Springer.

Chelli, B., & Lemmi, A. (1995). "Totally" Fuzzy and Relative Approach to the Multidimentional Analysis of Poverty. *Economic Notes* , 115-134.

Chopra, R. (2018). *Machine Learning.* New Delhi: Khanna Book Publishing.

Cowan, R. S. (1976). "Industrial Revolution" in the Home: Household Technology and Social Change in the 20th. *Technology and Culture , 17*, 1-23.

Deutsch, J., & Silber, J. (2006). The "uzzy St" Approach to Multidimensional Poverty Analysis: Using the Shapley Decomposition to Analyze the Determinants of Poverty in Israel. In A. L. Betti, *FUZZY SET APPROACH TO MULTIDIMENSIONAL POVERTY MEASUREMENT* (pp. 155-174). New York: Springer.

Duflo, E. (2003). Grandmothers and Granddaughters: Old Age Pension and Intrahousehold Allocation in South Africa. *World Bank Economic Review , 17* (1), 1-25.

Haddad, L., & Kanbur, R. (1990). How Serious is the Neglect of Intra-Household Inequality ? *The Economic Journal* , 866-881.

Klasen, S., & Lahoti, R. (2016, March 11). *How Serious is the Neglect of Intra-Household Inequality in Multi-dimensional Poverty Indices ?* Retrieved March 23, 2019, from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2742083

Kubat, M. (2017). *An Introduction to Machine Learning.* Cham: Springer.

Lundberg, S. M., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model.* Retrieved March 19, 2019, from NIPS Proceedings: https://computer.org/csdl/proceedings-article/icpr/1994/00576879/12OmNx965uE

Martinetti, E. C. (2006). Capability Approach and Fuzzy Set Theory: Description, Aggregation and Inference Issues. In A. L. Betti, *FUZZY SET APPROACH TO MULTIDIMENSIONAL POVERTY MEASUREMENT* (pp. 93-113). New York: Springer.

Munshi, K., & Rosenzweig, M. (2016). Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *American Economic Review , 106* (1), 46-98.

Qizilbash, M. (2006). Philosophical Accounts of Vagueness, Fuzzy Poverty Measures and Multidimensionality. In A. Lemmi, & G. Betti, *FUZZY SET APPROACH TO MULTIDIMENSIONAL POVERTY MEASUREMENT* (pp. 9-28). New York: Springer.

Ravallion, M. (2016). *THE ECONOMICS OF POVERTY - HISTORY, MEASUREMENT AND POLICY.* New Delhi: Oxford University Press.

Sen, A. (1995). *Inequality Reexamined.* London: Harvard University Press.

Sen, A. (1976). Poverty: An Ordinal Approach to Measurement . *Econometrica* , 219-231.

Shorrocks, A. F. (2013). Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. *The Journal of Economic Ineuality* , 99-126.

Srinivasaraghavan, A., & Joseph, V. (2019). *Machine Learning .* Hoboken: Wiley.

Theobald, O. (2017). *Machine Learning For Absolute Beginners: A Plain English Introduction (Second Edition).* Seattle: Scatterplot Press.

Vargha, L., Gál, R. I., & Crosby-Nagy, M. O. (2017). Household production and consumption over the life cycle: National Time Transfer Accounts in 14 European countries. *DEMOGRAPHIC RESEARCH , 36* (32), 905-944.

Vijaya, R. M., Lahoti, R., & Swaminathan, H. (2014). Moving from the Household to the Individual: Multidimensional Poverty Analysis. *World Development , 59*, 70-81.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control , 8* (3), 338-353.