लखनऊ विश्वविद्यालय
University of Lucknow
(Accredited A++ by NAAC)

WORLD CONGRESS OF EPIDEMIOLOGY 2024

Developing Flexible statistical models to study undernutrition among under-five children in India

DR. SHAMBHAVI MISHRA

DEPARTMENT OF STATISTICS, UNIVERSITY OF LUCKNOW, LUCKNOW, UTTAR PRADESH, INDIA

Email: mishra_shambhavi@lkouniv.ac.in

WORLD CONGRESS OF EPIDEMIOLOGY WCE 2024

25-09-2024

Introduction

- Under-nutrition among under-five children is a serious public-health problem in the World. Poor nutrition in the first ***1,000 days*** of a child's life can also lead to stunted growth, which is associated with impaired cognitive ability and reduced school and work performance.
- The interaction between undernutrition and infection can create a potentially lethal cycle of worsening illness and deteriorating nutritional status.
- While the 2023 edition of the UNICEF-WHO-World Bank Group Joint Malnutrition Estimates shows that stunting prevalence has been declining since the year 2000, more than one in five – 148.1 million children under 5 –were stunted in 2022, and at least 45.0 million suffered from wasting at any given point of time in the year.
- The indicators of malnutrition are alarming in India as well. The National Family Health Survey 5 (NFHS-5, 2019-21) found that while some nutrition indicators for children under five improved, severe acute malnutrition (SAM) is a public health emergency in India
- Undernutrition, as we know is a complex interplay of factors. Low Birth weight, lack of breastfeeding, undernutrition, air pollution, hazardous drinking water & food and poor hygiene practices are some of the important risk factors associated with two leading causes of death among under-five children worldwide: pneumonia and diarrhea.

Introduction

- The most conventionally applied binary logistic regression model uses the dichotomized version of the nutrition scores. The dichotomization of the continuous outcome variable is done by defining a fixed cut-off point which results in loss of information. By using the response variable in the continuous form, it coarsens the outcome and enables in providing a detailed picture of its determinants.
- When undernutrition is considered as the outcome variable, where it is assumed that the shape of the response variable is skewed i.e. it does not follow a well-known distribution, use of ***flexible models*** may be adopted to study effect of covariates.
- For undernutrition, when considering continuous variable, mean of the response variable is not of interest as researcher will not be concerned about average undernutrition. However, areas of interest would be lower quantiles of the response variable.
- ***Quantile Regression Modelling*** approach is a distribution-free approach useful when shape of the response variable depends on covariates.
- The focus of this research is to model conditional quantiles of the outcome variable and comparing it with the conventionally applied models. The research further extends to examine if covariates exert their effect differentially in a non-linear way.

Methodology
Quantile Regression Modelling

- Quantile regression modelling was introduced by Koenker and Bassett in 1978 as an extension to the conventionally used mean regression modelling to model conditional quantile function of a continuous response variable Y depending on a set of explanatory variables X. The linear quantile regression model is written as:

$$y_j = {x_T}^t \beta_j + \epsilon_{\tau j}$$

where, j=1, 2,…,n are the observations. yj is the j[th] value of the outcome variable.

xj = (1, x_{j1}, … … … , x_{jq})T is the covariate vector for j[th] observation, $\beta\tau$ = ($\beta_{\tau 0}$, $\beta_{\tau 1}$, … … … , $\beta_{\tau q}$)T is quantile-specific linear effects & τ ∈ (0,1) indicates quantile (fixed in advance), $\varepsilon_{\tau j}$ is the unknown error term with CDF $F_{\varepsilon\tau j}$ and density $f_{\epsilon\tau \mathrm{j}}$

- Further, to also account for possible non-linear relationships between covariates and quantiles of the response variable, the above model is further extended to ***additive quantile regression model:***

$$Q_{Y_j}(\tau \mid xj,zj) = x_j^T \beta\tau + \sum_{i=1}^{p} f_{\tau i}(z_{ji})$$

where, $x_j^T \beta\tau$ denotes linear term including the intercept and $f_{\tau i}$ denotes smooth functions of continuous covariates z_{ji} , (i=1,2,…..,p) assumed to be non-linearly related with the quantile function of the response variable

Results

Fixed effects on underweight for under-five children in India

Covariates	0.1 quant	0.3 quant	0. 5 quant
Intercept	-3.789***	-2.647***	-1.973***
Sex [Female as reference]			
Male	-0.095***	-0.072***	-0.057***
Birth Order [1 as reference]			
2-3	-0.034	-0.115***	-0.075***
4-5	-0.042	-0.135***	-0.150***
6+	-0.147	-0.128*	-0.179***
Residence [Rural as reference]			
Urban	-0.174***	-0.109***	-0.087***
Mother Education [Higher as reference]			
No Education	-0.392***	-0.321***	-0.297***
Primary	-0.191***	-0.187***	-0.229***
Secondary	-0.119**	-0.112***	-0.148***

Results

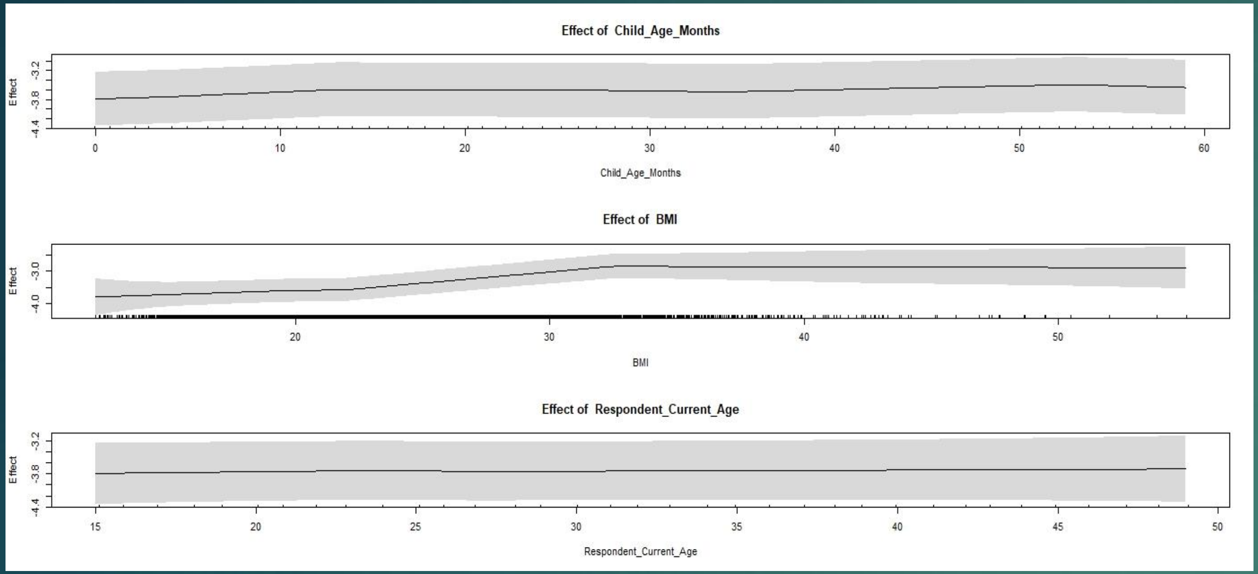
Fixed effects on underweight for under-five children in India (contd.)

Covariates	0.1 quant	0.3 quant	0. 5 quant
Wealth Index [Rich as reference]			
Poor	-0.15012***	-0.164385***	-0.14754***
Rich	0.10660*	0.133121***	0.11014***
Birth Type [Multiple Births as reference]			
Singleton	0.54899***	0.375498***	0.33907***
Had Diarrhea in last two weeks [No as reference]			
Yes	0.09311	-0.007583	-0.01718
Had Fever in last two weeks [No as reference]			
Yes	0.02924	-0.044604	-0.02892

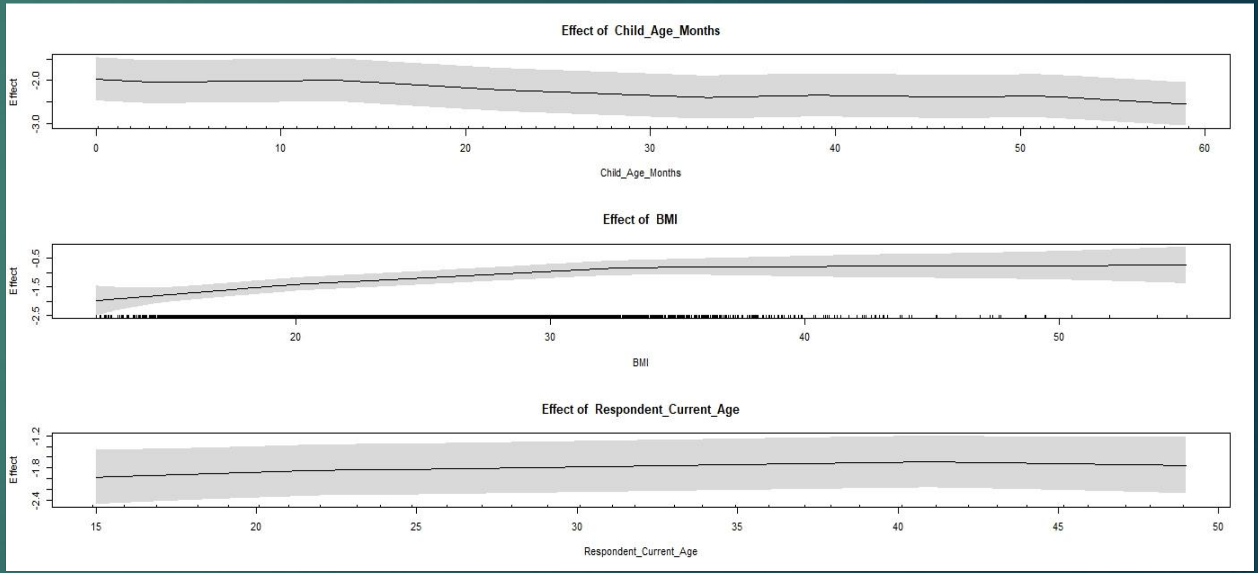
Results

Nonlinear effects on weight-for-age for under-five children

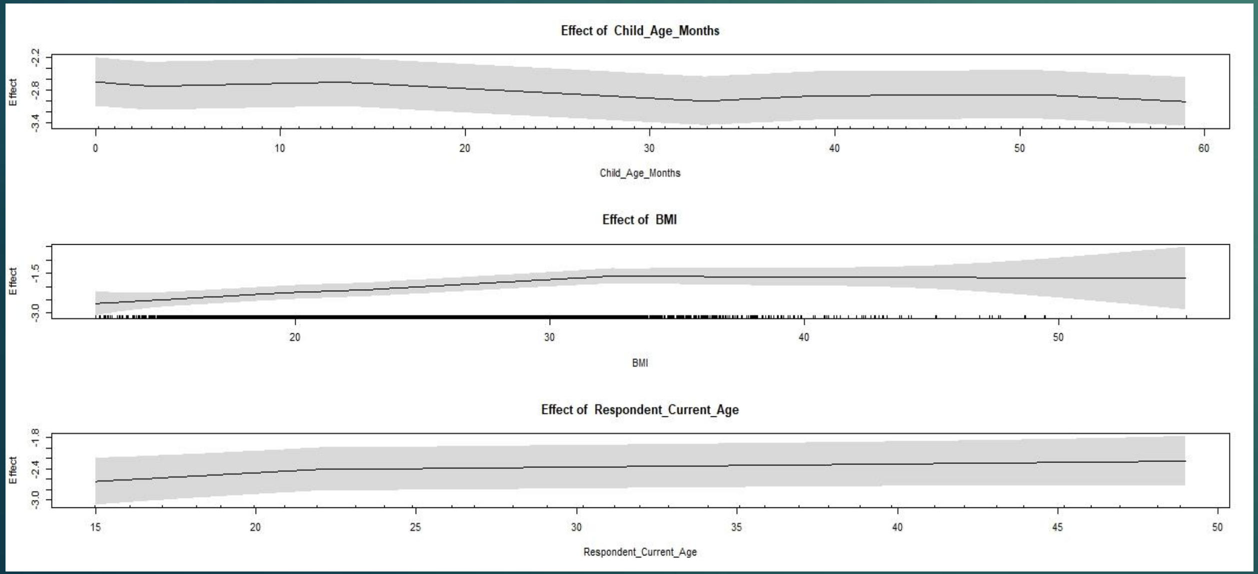
q = 0.10



q = 0.50



q = 0.30



SHAMBHAVI MISHRA

Results

Akaike Information Criterion (AIC) for the fitted models

Quantiles (q)	Linear Quantile Regression	Additive Quantile Regression Smoothing
STUNTING		
0.20	125586.1	124782.5
0.40	121507.6	119668.3
0.50	122140.2	119802.2
UNDERWEIGHT		
0.10	122855.4	122264.3
0.30	108473.3	107642.8
0.50	104902.4	103646.8
WASTING		
0.07	132375.5	130939.2
0.17	121148.5	119841.9
0.50	108536.6	107988.3

Conclusions

- For stunting, Child' s age and Mother's Body Mass Index were among the continuous factors exerting non-linear effect on stunting, with mother's BMI showing maximum effect size at lower end of the distribution.
- For underweight also these two covariates affected the outcome in a non-linear way, however mother's nutritional status had the maximum effect size in the lower quantiles especially for wasting. The results were in accordance with the results of similar studies performed.
- For weight-for-height z-score, significance of covariates varied across quantiles showing their differential impact. Child's age, mother's BMI and mother's age were among the continuous covariates affecting the outcome in a non-linear way. The magnitude of the effect was maximum for lower quantile and a reduction was observed thereafter.
- Although widely applicable, logistic regression model enables the researcher to have an idea of the determinants of undernutrition, providing only a preliminary basis.
- To study variables such as nutritional status of children, where lower quantiles are of main interest, focus should be on how factors affect the entire conditional distributional of the outcome variable taken as is rather than summarizing the distribution at its mean. This can be achieved by applying quantile regression modeling.
- An extension to it further enables to non-parametrically estimate the linear or potentially non-linear effects of continuous covariates differentially on the outcome using penalized splines

Thank You

Any questions??