Understanding the Geographical heterogeneity of children undernourished and its risk factors across the districts of India: Geographically Weighted Regression Analysis

Apoorva Nambiar[1,2,3], Ashish Singh[2], Dharmalingam Arunachalam[3], Satish B Agnihotri [2]

[1]IIT Bombay- Monash Research Academy; [2] Indian Institute of Technology, Bombay; [3]Monash University, Australia

---

Background

Undernutrition among children always has been and continues to be one of the critical development issues globally. And India has been identified as one of the countries among the Low- and Middle-Income countries where prevalence of undernutrition is alarmingly high. Even though the recent evidence shows a nominal reduction in the rates in India, the goals are far to be reached. Hence, fresh approaches are necessary to analyse the problem and its risk factors in order to produce robust evidence to find solutions. And the improved spatial analytical methods and models entrenched into the Geographic Information Systems (GIS) facilitate more in-depth examination of the causes and consequences behind the problem of undernutrition among children. These applications have significant policy implications because they aid in understanding the relationships between geography and child undernutrition and its various determinants, and are particularly important in planning impactful programmes to help address such issues. Drawing on which, the objective of the present study is to unravel the location-specific (local) spatial heterogeneity and spatial dependency, in associations between children undernourished and its selected risk factors across the districts of India.

Data & Methods

The latest 2019-20, the fifth round of the National Family Health Survey, was utilised in the current study. The prevalence of children stunted, children underweight and children wasted were the three outcome variables studied.

Analytical Approach

Firstly, Principal Component Analysis (PCA) was applied on the selected 22 independent households, maternal and child related variables. Principal Components (PCs) with eigenvalues larger than one were considered. The Kaiser–Mayer–Olkin (KMO) sampling adequacy test value with 0.83, suggested that the variables were suitable for the PCA analysis. Once the complexity of the data set was reduced to its principal components, the scores were predicted and the bivariate, multivariate geo-spatial techniques and models were applied on the principal components with the outcome variables to identify the association and the regional patterns of the magnitude of association across the 705 districts of India.

Local Indicators of Spatial Association (LISA) statistics, Univariate and Bivariate Local Moran's *I*, were employed to confirm the spatial clustering and dependency in the nutritional status of children along with their risk factors. The stationary regression models like the Ordinary Least Square (OLS), and various spatial econometric models like the Spatial lag and spatial error models, Geographically Weighted Regression (GWR), were applied to check for spatial dependency and examine the spatial influence of various correlates on the health of the child, across the geographical space, after adjusting for the endogeneity. GWR, a non-stationary regression model, was utilised to study the local relationships between these variables, their direction and magnitude, location-specific, across all 705 districts. The spatial influences between neighbourhoods by considering local regression parameters rather than global regression parameters were estimated independently for each location.

Results

The first PC was heavily loaded on IFA supplementations, antenatal care visits, institutional delivery and postnatal care given to both mother and child. The second component was loaded heavily on education and nutritional status of the mother, sanitation facilities at home, prevalence of anaemia in both women and child. The third component was loaded on child's pre-primary education, early marriage and early pregnancy, adequate diet and prevalence of diarrhoea. and the fourth component on households using clean fuel for cooking, improved drinking water and having health insurance. Hence, the four retained principal components were broadly termed as, PC1: maternal care during pregnancy and child-birth, PC2: Social determinants, PC3: Child and adolescent care, and PC4: Household characteristics.

Table 1 depicts the aspatial, OLS, model results. PC1, primarily influenced by the characteristics of maternal care during pregnancy and child-birth showed the strongest negative association with stunting, denoting the effect of intergenerational cycle on the child being stunted. That is, if the mothers have higher antenatal care visits, IFA supplementations, delivers in a health facility and if the mother and child both receive post-natal care, might lead to a reduction in the stunting prevalence. Adjusted R^2 explaining the significant variance in the outcomes, was the highest for PC1 with stunting. PC2, heavily influenced by the social determinants namely schooling and the nutritional status of the mother, usage of improved sanitation facilities, and anaemia in children and women, showed high and positive association with children wasted. That is, with poor sanitation facilities, low levels of schooling and BMI of mothers, and if the child and women suffers from anaemia, will lead to an increasing rate of wasting among children. The underweight parameter, which is a comprehensive measure of both stunting and wasting, had almost the same levels of association with both PC1 and PC2, i.e, maternal care during pregnancy and child-birth as well as the social determinants.

Table 1: Results of bivariate and multivariate linear regression model applied on the four extracted PCs and the undernutrition parameters, NFHS 5, 2019-20, India

<table>
<tr><th colspan="3" rowspan="2">Results</th><th colspan="4">Components</th></tr>
<tr><th>PC1
(Maternal care during pregnancy and child-birth)</th><th>PC2
(Social determinants)</th><th>PC3
(Child and adolescent care)</th><th>PC4
(Household characteristics)</th></tr>
<tr><td rowspan="15">Model 1 (Bivariate)</td><td rowspan="5">Stunted</td><td>Coeff.</td><td>-2.14***</td><td>0.93***</td><td>-0.37</td><td>0.29</td></tr>
<tr><td>CI</td><td>[-2.33, -1.95]</td><td>[0.59, 1.26]</td><td>[-0.80, .05]</td><td>[-0.21, .80]</td></tr>
<tr><td>R-squared</td><td>0.41</td><td>0.04</td><td>0.004</td><td>0.001</td></tr>
<tr><td>Adj R-squared</td><td>0.41</td><td>0.04</td><td>0.002</td><td>0</td></tr>
<tr><td>Prob > F</td><td>0</td><td>0</td><td>0.09</td><td>0.26</td></tr>
<tr><td rowspan="5">Underweight</td><td>Coeff.</td><td>-2.05***</td><td>2.57***</td><td>-0.24</td><td>0.12</td></tr>
<tr><td>CI</td><td>[-2.28, -1.80]</td><td>[2.22, 2.91]</td><td>[-.73, .25]</td><td>[-.45, .71]</td></tr>
<tr><td>R-squared</td><td>0.28</td><td>0.23</td><td>0.001</td><td>0</td></tr>
<tr><td>Adj R-squared</td><td>0.28</td><td>0.23</td><td>0</td><td>0</td></tr>
<tr><td>Prob > F</td><td>0</td><td>0</td><td>0.34</td><td>0.674</td></tr>
<tr><td rowspan="5">Wasted</td><td>Coeff.</td><td>-0.67***</td><td>1.24***</td><td>0.03</td><td>-0.05</td></tr>
<tr><td>CI</td><td>[-.85, -.48]</td><td>[.99, 1.49]</td><td>[-.29, .36]</td><td>[-.43, .34]</td></tr>
<tr><td>R-squared</td><td>0.067</td><td>0.1211</td><td>0</td><td>0</td></tr>
<tr><td>Adj R-squared</td><td>0.0657</td><td>0.1198</td><td>0</td><td>0</td></tr>
<tr><td>Prob > F</td><td>0</td><td>0</td><td>0.84</td><td>0.82</td></tr>
<tr><td rowspan="15">Model 2 (Multivariate)</td><td rowspan="5">Stunted</td><td>Coeff.</td><td>-2.14***</td><td>0.93***</td><td rowspan="15"></td><td rowspan="15"></td></tr>
<tr><td>CI</td><td>[-2.33, -1.95]</td><td>[0.67, 1.18]</td></tr>
<tr><td>R-squared</td><td colspan="2">0.45</td></tr>
<tr><td>Adj R-squared</td><td colspan="2">0.45</td></tr>
<tr><td>Prob > F</td><td colspan="2">0</td></tr>
<tr><td rowspan="5">Underweight</td><td>Coeff.</td><td>-2.05***</td><td>2.57***</td></tr>
<tr><td>CI</td><td>[-2.24, -1.84]</td><td>[2.29, 2.84]</td></tr>
<tr><td>R-squared</td><td colspan="2">0.52</td></tr>
<tr><td>Adj R-squared</td><td colspan="2">0.51</td></tr>
<tr><td>Prob > F</td><td colspan="2">0</td></tr>
<tr><td rowspan="5">Wasted</td><td>Coeff.</td><td>-0.67***</td><td>1.24***</td></tr>
<tr><td>CI</td><td>[-.83, -.49]</td><td>[1.00, 1.48]</td></tr>
<tr><td>R-squared</td><td colspan="2">0.19</td></tr>
<tr><td>Adj R-squared</td><td colspan="2">0.19</td></tr>
<tr><td>Prob > F</td><td colspan="2">0</td></tr>
</table>

*** P<0.001 level of significance

Table 2: Results: Aspatial OLS, the Spatial Lag and Spatial Error Models, children stunted with the highest loaded variables of the extracted four PCs. NFHS 5, 2019-20

District- level correlates	OLS β	Spatial Lag Model β	Spatial Error Model β
Households using improved drinking water %	0.05	0.03	0.03
Women whose Body Mass Index (BMI) is below normal (BMI < 18.5 kg/m2) %	0.64***	0.43***	0.56***
Women having early pregnancies %	0.06	0.03	0.08
Children who received postnatal care within two days of delivery %	-0.21***	-0.14***	-0.17***
R- squared value	0.43	0.52	0.52
Lambda Value (Lag Coefficient)			0.46***
Rho Value (Lag coefficient)		0.41***	
Log likelihood	-2305	-2256	-2263
AIC value	4621	4524	4536
No. of regions	705	705	705

All three aspatial models reported that mother's having low BMI showed a strong positive correlation with all three indicators of undernutrition among children. Correlates like mother's nutritional status and early pregnancy were positively associated with children underweight, whereas, postnatal care for children showed a negative association with underweight. With respect to children stunted, mother's nutritional status and postnatal care for children showed significant association, where mother having low BMI had positive association and postnatal care for children had negative association. Children wasted had significant association with households using improved drinking water and mother's BMI level. Households having improved drinking water to consume, showed significant negative association with only one nutritional parameter—children wasted.

With respect to stunting, lag coefficient vale of 0.41 (p-value <0.001), indicated that a variation in the children underweight prevalence in a specific region may significantly affect the rate of the children underweight prevalence by 41% in the neighbouring regions.

Results: Geographically Weighted Regression

GWR, an extension of typical standard regression models, captured geographic influences among neighbourhoods by considering local regression parameters rather than global regression parameters that were calculated separately for each location. And below are the results of the spatial mapping of the significant local coefficients as derived from the GWR model, at the unit of districts for all three indicators of undernutrition. The local estimated coefficients for each of the retained components (risk factors) were mapped in quintile scale. Results clearly show how these retained PCs differ with respect to geographical regions.

The positive association meant that with the increase in the values of the predictor variables (the particular retained components), undernutrition among children reduced in those particular locations, and higher the value of the coefficient, stronger the association in those locations.

The relationship between PC1 (i.e, maternal care during pregnancy and child-birth, and the social determinants) and children stunted was negative and highly location-specific. The local association or the strength of the local coefficients, varied across the districts. The districts having darker shades show strong negative association between the maternal care during pregnancy and child-birth with the child undernourished. Few districts from the north-east, eastern coastal belt from Odisha to Andhra Pradesh, and a few districts of UP, Madhya Pradesh, Himachal Pradesh show strong negative association, indicating that in these districts, with the improvement in the maternal care during pregnancy and child-birth, the stunting levels will decrease. The relationship between the PC2, i.e., the social health determinants (namely education and nutritional status of the mother, sanitation facilities at home, prevalence of anemia in both women and child) and children stunted was positive and highly location-specific. The districts having darker shades show strong positive association between the social health determinants and the nutritional status of the child. Districts of Gujarat, Maharashtra, northern Karnataka, Telangana and Andhra Pradesh show strong positive association, indicating that in these districts the deteriorating rates of social health determinants might lead to poor stunting levels. Madhya Pradesh shows a very interesting pattern, having various levels of association across its districts.

It is also observed that the clusters that are formed when one PC is examined with the children undernourished without adjusting for other correlates versus after adjusting, change with the change in the magnitude of association between them.

Similar results with different clusters are observed when these retained PCs are studied with respect to children wasted and underweight. Below are the GWR local coefficients maps showing location specific coefficients of association between PC1 and PC2 with undernutrition parameters:

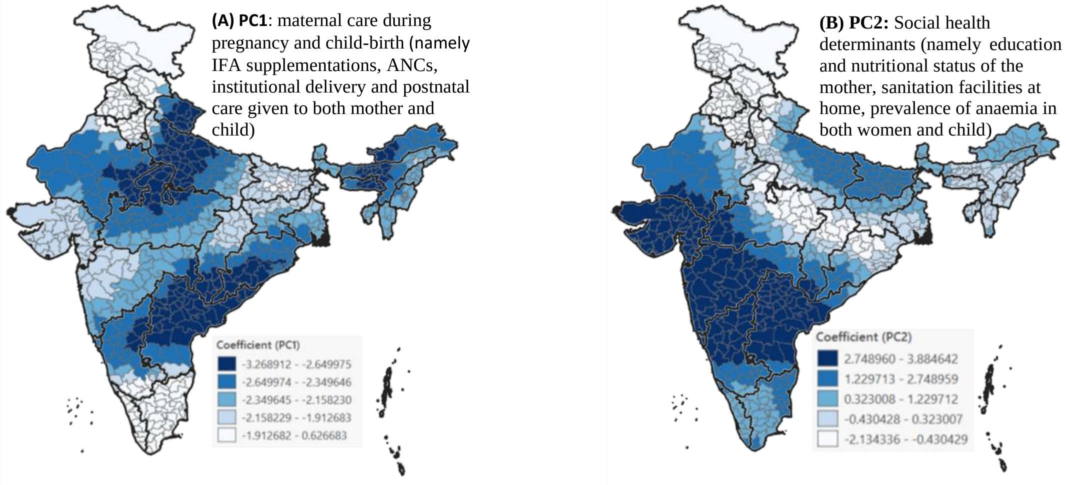

Figure 1: District-wise Geographical variation of the local coefficients of ***children stunted*** *with (A) PC1, (B) PC2*, as derived from the GWR model results. NFHS 5, 2019-20, India

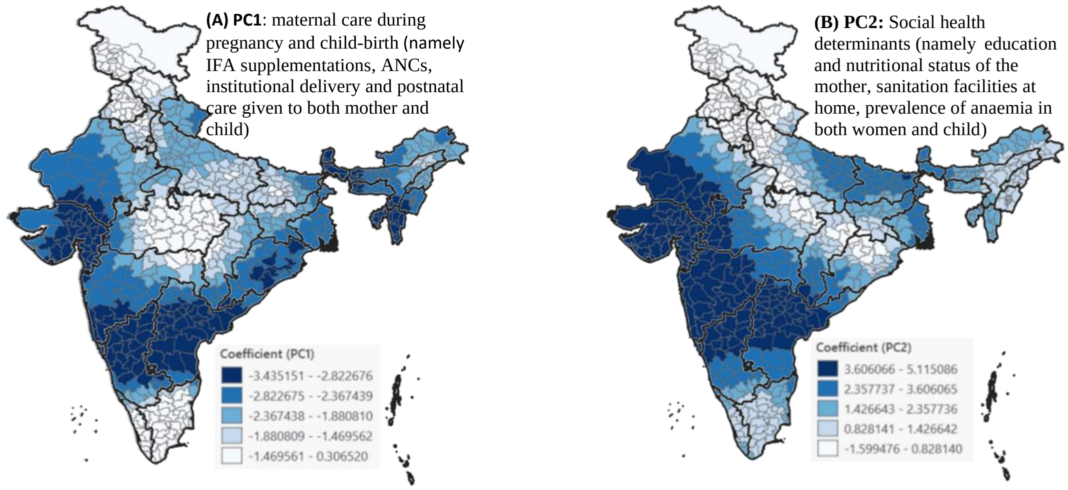

Figure 2: District-wise Geographical variation of the local coefficients of ***children underweight*** *with (A) PC1, (B) PC2*, as derived from the GWR model results. NFHS 5, 2019-20, India

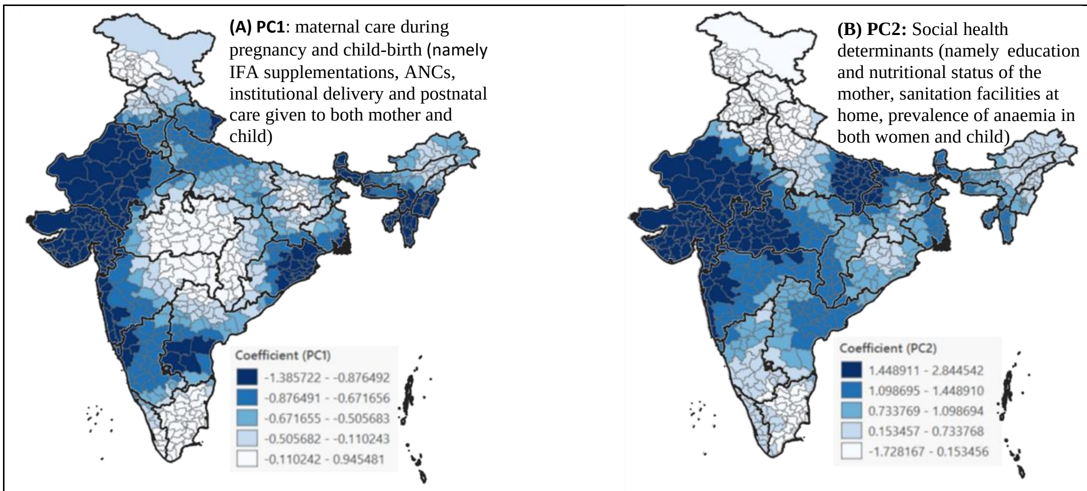

Figure 3: District-wise Geographical variation of the local coefficients of ***children wasted*** *with (A) PC1, (B) PC2*, as derived from the GWR model results. NFHS 5, 2019-20, India

Conclusion

The GWR findings revealed that the geographical variability in correlations between undernutrition among children and its risk factors across the districts were substantially region-specific, with the direction and strength of these relationships varying between districts, thus allowing for a more nuanced understanding of the potential local determinants of the problem that may otherwise be hidden over the examined space. Within a state itself the association between the variables vary. Thus, the identification of geographical hot spots, and the contributing factors to the high undernutrition levels across and specific to the regions, can assist policymakers in developing robust programs and interventions specific to nutrition in order to tackle the problem and produce undernutrition-free enclaves. These results from the analysis facilitates the allocation of resources effectively to reduce health inequities between and within districts.

While child nutrition in India is a major development agenda, special focus is arguably required on the poorly performing regions, especially by addressing the maternal health care needs, improvement of sanitation facilities, better health care services, improving educational levels of women, and enhancing women empowerment. Only by addressing these causes and consequences, the SDG goals in terms of addressing health and wellbeing and the National Nutrition Mission goals will be achieved. The evidence gathered from this study can be used by decision-makers for developing better strategies at the micro-level and long-term planning to find solutions to mitigate the problem of undernutrition in India and can also be used as a case study for other LMICs.