Socioeconomic segregation of the poor is associated with higher child mortality in major Indian cities

Tarani Chandola, Social Statistics, University of Manchester
Sita Mikkilineni, Indian Institute of Public Health, Hyderabad
Souvik Bandyopadhyay, Indian Institute of Public Health, Hyderabad
Anil Chandran, Indian Institute of Public Health, Hyderabad

Abstract:
Background: Although urbanisation is generally associated with poverty reduction in developing counties, in some contexts, it results in increased socioeconomic segregation of the poor. Some studies suggest that such segregation is associated with higher mortality and poorer health outcomes. The paper examines whether socioeconomic segregation of the poor is associated with higher child mortality risk in Indian cities and whether this association is confounded by other socio-demographic factors at the contextual (city) level and at the compositional (household) level.

Methods: Data from 39,427 households living in 1,876 urban wards within 59 districts (cities) with more than a population of 1 million from the third (2008) District Level Household Survey (DLHS-3) and 2001 Indian Census were analysed. The outcome was any child death reported by households in the preceding 4 years of the DLHS-3 survey. An index of isolation of the poor (measured using adult illiteracy as a proxy) was estimated at the neighbourhood and district levels. Potential confounding socio-demographic variables from the district and household variables were also analysed.

Results: The Index of isolation was associated with higher child mortality, even after taking into account socio-demographic confounders. A unit increase in the likelihood that a poor household lives next to another poor household in poor neighbourhoods is associated with an average increase in 1.3 child deaths in that district.

Conclusion: Increasing segregation of the poor may result in higher child mortality. As developing countries become increasingly urbanised, there is a risk that this leads to increasing segregation of the poor as well as increasing child mortality.
Introduction
Urbanisation in developing countries is often accompanied by an increasing spatial concentration of poverty in cities. Although urbanisation can play a positive role in overall poverty reduction, when it is accompanied by weak economic growth or when distributive policies are nonexistent or ineffective, it results in the socioeconomic segregation of poor people through the concentration of poor people living in poor localities, rather than significant poverty reduction (UN HABITAT 2011). Although levels of urbanisation in India are relatively low at around 31 percent in 2011, the rate of urbanisation is among the highest in the world. As a result of urbanisation, India contains the highest number of urban slum dwellers, accounting for 17 per cent of the world’s slum dwellers.

The association between poverty and ill health is well known. The risk of premature mortality and disease as a result of living in poor areas such as slums with inadequate housing, inaccess to safe water to adequate sanitation, is well known. Furthermore, there are now a multitude of studies showing associations between area level poverty rates and deprivation and poor health/mortality, even after taking into account individual and household level socioeconomic circumstances (Van Ham et al 2012, Pickett and Pearl 2001, Kawachi and Berkman 2003). This suggests an area or contextual effect on health associated with living in a poor area, over and above the compositional effect of the number of poor people in an area. What is perhaps less well known is the association between socioeconomic segregation and health. Studies have shown that there are ecological correlations between socioeconomic segregation by income and mortality rates (Ross et al. 2001, Waitzman and Smith 1998, Lobmayer and Wilkinson 2002, Szwarcwald et al 2002). These studies have shown that greater socioeconomic segregation of the poor within cities is associated with higher mortality rates. However, there are few, if any analyses that take into account potential confounders of this ecological correlation, namely the strong association between higher socioeconomic position and health among individuals and families. It is likely that as poor people live in poor areas that are socioeconomically isolated from the rest of the city, any association between segregation and health needs to take account of such confounding from socioeconomic factors at the compositional (household) and contextual (area) levels.

The main research question of this paper is to investigate whether being poor, living in a poor area that is socioeconomically segregated from the rest of the city is associated with greater mortality risk.

Methods

Data rom the third District Level Household Survey (DLHS-3, IIPS 2010) and 2001 Indian Census were analysed were analysed for this study. The District Level Household Survey was designed to provide
information on reproductive and child health in all the districts of India. The steps involved in the sampling of households were:

(i) the selection of primary sampling units (PSUs) within Census wards;
(ii) the selection of the households from each of the selected village/urban PSUs through a random selection of households within a census enumeration block (which we have used to define as a neighbourhood in this study).

The total DLHS-3 population was restricted to households living within urban wards of the largest cities (districts) with a population of more than 1 million. This was to ensure comparability of the district units within the analysis. Mega cities such as Kolkata, Mumbai and Delhi were disaggregated into their constituent administrative districts. DLHS-3 wards were merged with ward level socioeconomic indicators from the 2001 Indian census. Around 9% of the DLHS-3 wards did not have a unique identifier, and they were deleted from further analysis. This resulted in an analytical sample of 39,427 households living in 1,876 urban wards within 59 districts.

Variables:
Health: Child mortality (age 0-4) was reported for the preceding four years in the DLHS-3 (from 2004-2008). This was recoded into a binary variable (any child death in the household vs. no child death) for logistic regression analysis. The under 5 death rate was calculated at the district level by dividing the total child mortality in a district by the number children in the DLHS-3 sample (and multiplying by 100).

As there are no direct measures of income or poverty in the DLHS and census data, poverty was measured through the use of a proxy indicator namely illiteracy. Another reason for using illiteracy to measure poverty is the difference in the meanings and concepts of standard poverty measures (such as indicators of wealth and status) across Indian cities.

The Isolation Index of Segregation:
As data on the socioeconomic distribution of the population within neighbourhoods are not available across all major Indian cities, the study used estimates of neighbourhood socio-economic composition from DLHS-3. However, the DLHS-3 is a sample survey of primary sample units (neighbourhoods) within wards and districts, and does not have information on all neighbourhoods within a district. Hence it was not possible to use the survey alone to calculate the standard Isolation Index of Segregation (Massey & Denton 1988). Instead, the DLHS-3 data was linked to the 2001 Indian census data to enable the calculation of the Index as shown below.
\[
\text{Ind\_isolation} = \sum \frac{\text{illiterate}_{jk}}{\text{illiterate}_k} \times \frac{\text{illiterate}_{ijk}}{\text{population}_{ijk}}
\]

where

- \(\text{illiterate}_{jk}\) is the number of illiterate people living in ward_{jk} within district_{k} (from Census data)
- \(\text{illiterate}_k\) is the number of illiterate people living in a district_{k} (from Census data)
- \(\text{illiterate}_{ijk}\) is the number of illiterate people living in neighbourhood_{ijk} (from DLHS-3 data) within ward_{jk}
- \(\text{pop}_{ijk}\) is the population living in neighbourhood_{ijk} (from DLHS-3 data) within ward_{jk}

The difference between the standard Isolation index and the one used in this study lies in the different data sources in Equation (1) for the two terms after the summation- the concentration rate of the district’s illiterate population within a Census ward and the illiteracy rate within the neighbourhood of the same Census ward. In the standard Index, both the concentration rate and the illiteracy rate refer to the same area unit, but in this study, the concentration rate refers to the Census ward while the illiteracy rate refers to a neighbourhood within the same Census ward. As the DLHS-3 does not contain information on all neighbourhoods and wards within a district, the estimated Index may not accurately reflect the true distribution of the Isolation Index. However, the use of the DLHS-3 survey data does allow for the estimation of the Index at the neighbourhood level (through the random sample of households within a primary sampling unit), rather than the estimation of the Index at the ward level, which often refers to a very large area and population (around 25,000-75,000) for the largest Indian cities.

The Index can vary from 0 to 100 where higher values denote greater segregation of the poor (measured using the adult illiterate population). An interpretation of the index is the likelihood that an illiterate (which we use to proxy for poor) household will live next to another poor household within the neighbourhoods where poor people live.

Confounders:
The illiteracy rates at the District and household levels were calculated from the Census and DLHS-3 data respectively. Other potential socio-demographic confounders included the number of household members, household wealth (constructed by the DLHS-3 survey team from a factor analysis of consumption measures and housing quality indicators), religion of head of household, whether the head of household belonged to a scheduled caste or tribe, and whether the household accessed government or private health care when someone was ill.

Analysis:
A scatterplot was used to examine the ecological correlation between the district level Index of isolation of the poor and the district level child mortality rate. The district level Index of isolation was derived from
summing up the PSU level Isolation index scores (in equation 1) within a district. Multilevel logistic regression (with the outcome of child death within a household) was used to estimate the effect of the (PSU level) Index of isolation, after taking into account potential socio-demographic confounders at the district and household levels. Multilevel analysis is particularly appropriate as it takes into account the sample design of the DLHS-3 which results in clustered samples at the district and Primary Sampling Unit (PSU) levels. It is important to take this clustering into account in the analysis as the main exposure variable is at the area level, rather than at the household level. The multilevel regression coefficients were estimated using MCMC estimation procedures within MLwin. The Bayesian Deviance Information Coefficient (DIC) was used to compare model fit. Convergence of the MCMC model was assessed using kernel density plots for normality of the posterior estimates, plots for assessing autocorrelation (AFC) and partial autocorrelation (PACF), and Raftery–Lewis diagnostic and Brooks–Draper diagnostic as suggested number of iterations (Browne, 2004). Goodness of fit was graphically examined using plots for normality of residuals and homoscedasticity for each level in the model (Browne, 2004).

Results

The scatterplot of the district level associations between child death rates and the Index of Isolation (Figure 1) shows that there is a strong ecological correlation (R=0.56). Districts (cities) where the poor (illiterate population) are more isolated have higher child mortality rates. This correlation is equivalent to the following effect size: a unit increase in the likelihood that a poor household lives next to another poor household in poor neighbourhoods is associated with an increase in 1.3 child deaths in that district.

There are similar correlations between the Index of isolation and the illiteracy rate at the district level (R=0.54) and the child mortality rate and the illiteracy rate at the district level (R=0.55). This suggests that the cities where the poor are most isolated (Lucknow and Ghaziabad for example) are also the districts with the highest child mortality rates and highest illiteracy rates. Similarly, the cities where the poor are least isolated (Ernakulam, Thiruvananthapuram and Kozhikode) are also cities with the lowest child mortality rates and lowest illiteracy rates. So it is likely that the strong correlation at the district level between socioeconomic segregation and child mortality could be confounded by other socioeconomic factors like the illiteracy rate. Hence it is important to take account of other socioeconomic factors at the district and household levels when examining the association of socioeconomic segregation with child mortality. This is done in the following multilevel logistic regression models.
Figure 1: Scatterplot of district level child death rate with Isolation Index of poverty in 59 Indian urban districts - DLHS-3 (2008) and 2001 Indian census data

Table 1 displays the distribution of the key independent variables in the analysis by any child death in the household. 1.3% (N=524) of all households in the analytical sample reported the death of child aged 0-4 in the four year period preceding the survey. Households where there was a child death in the previous 4 years had higher Index of isolation scores, higher illiteracy rates at the district and household levels, poorer wealth scores, more household members and were more likely to be Muslim and Scheduled Caste/Scheduled Tribe households.
Table 1: Distribution (Mean and Percentages) of key independent variables in the analysis by child death in the household

<table>
<thead>
<tr>
<th>Variable</th>
<th>No child death in the household</th>
<th>Child death in the household</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>38922</td>
<td>524</td>
</tr>
<tr>
<td>Isolation Index (District level)</td>
<td>1.84</td>
<td>2.23</td>
</tr>
<tr>
<td>Illiteracy rate (District)</td>
<td>28.8%</td>
<td>31.5%</td>
</tr>
<tr>
<td>Illiteracy rate (Household)</td>
<td>15.7%</td>
<td>31.6%</td>
</tr>
<tr>
<td>% of Households which are Muslim</td>
<td>16.5%</td>
<td>23.5%</td>
</tr>
<tr>
<td>% of Households which are SC/ST</td>
<td>18.6%</td>
<td>23.7%</td>
</tr>
<tr>
<td>% Households using government clinics when ill</td>
<td>38.6%</td>
<td>42.4%</td>
</tr>
<tr>
<td>Number of household members</td>
<td>4.93</td>
<td>5.73</td>
</tr>
<tr>
<td>Wealth Index score</td>
<td>1.32</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table 2 displays the results (estimated log odds and 95% credibility intervals) of the multilevel logistic regression models with any child death in the household as the dependent variable. In Model 1, a variance components model is fitted which takes into account the clustering in child mortality at the state, district, primary sampling unit (PSU) and household levels. As the logistic model constrains the variance at the lowest (household) level to be one, there appears to be significant clustering in child mortality at the state level, followed by the PSU level. Model 2 adds in the Isolation Index as an independent variable. A unit increase in the isolation index (which corresponds to a 1% increase in the likelihood that a poor household will live next to another poor household in poor neighbourhoods) is associated with an increase of 0.21 in the log odds of child mortality. This corresponds to an increased odds ratio of 1.23. The variance at the state level reduces from model 1 and the model fit improves (reduction in the DIC statistic). This suggests that some of the clustering in child mortality within states is accounted by the association of the isolation index with child mortality.

Model 3 additionally adjusts for the district level illiteracy rate. While the coefficient for the isolation index reduces a little, it remains associated with child mortality. Model 4 adjusts for a number of potential socio-demographic confounders at the household level. This reduces the odds ratio associated with a unit increase in the Isolation Index to 1.16. Overall, there was a modest reduction in the odds ratio of child mortality from 1.23 to 1.16 after adjusting for potential socio-demographic confounders at the district and household levels. The variance associated with the state level reduced very slightly (from 0.25 in Model 3 to 0.22 in Model 4) when adjusting for a range of socio-demographic variables at the household level. This suggests that a large proportion of the clustering in child mortality at the state level may be accounted for contextual (state and district) level factors, rather than compositional (household) level factors.
Table 2: Log odds (and 95% confidence intervals) of nested Multilevel Logistic Regression Models of Child death: DLHS-3

<table>
<thead>
<tr>
<th>Fixed Part</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
<td>-4.68(-5.05,-4.28)</td>
<td>-4.65(-4.97,-4.38)</td>
<td>-4.71(-5.04,-4.42)</td>
<td>-4.45(-4.81,-4.11)</td>
</tr>
<tr>
<td>Isolation Index (District level)</td>
<td>0.21(0.04,0.37)</td>
<td>0.17(0.03,0.33)</td>
<td>0.15(0.01,0.29)</td>
<td>0.04(0.01,0.08)</td>
</tr>
<tr>
<td>Illiteracy rate (District)</td>
<td>0.06(0.03,0.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some household members are illiterate (ref: none are illiterate)</td>
<td>0.34(0.13,0.57)</td>
<td>0.34(0.13,0.57)</td>
<td>0.34(0.13,0.57)</td>
<td>0.34(0.13,0.57)</td>
</tr>
<tr>
<td>All household members are illiterate (ref: none are illiterate)</td>
<td>0.95(0.60,1.29)</td>
<td>0.95(0.60,1.29)</td>
<td>0.95(0.60,1.29)</td>
<td>0.95(0.60,1.29)</td>
</tr>
<tr>
<td>Muslim household (ref: Hindu)</td>
<td>-0.02(-0.26,0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other religion household (ref: Hindu)</td>
<td>-0.28(-0.82,0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other backward caste household (ref: SC/ST)</td>
<td>0.18(-0.08,0.44)</td>
<td>0.18(-0.08,0.44)</td>
<td>0.18(-0.08,0.44)</td>
<td>0.18(-0.08,0.44)</td>
</tr>
<tr>
<td>Other caste household (ref: SC/ST)</td>
<td>0.16(-0.07,0.39)</td>
<td>0.16(-0.07,0.39)</td>
<td>0.16(-0.07,0.39)</td>
<td>0.16(-0.07,0.39)</td>
</tr>
<tr>
<td>Missing/DK caste household (ref: SC/ST)</td>
<td>0.40(-0.40,1.10)</td>
<td>0.40(-0.40,1.10)</td>
<td>0.40(-0.40,1.10)</td>
<td>0.40(-0.40,1.10)</td>
</tr>
<tr>
<td>Number of people within household</td>
<td>0.08(0.04,0.11)</td>
<td>0.08(0.04,0.11)</td>
<td>0.08(0.04,0.11)</td>
<td>0.08(0.04,0.11)</td>
</tr>
<tr>
<td>Usually go to govt clinic when ill (ref: Private clinic)</td>
<td>-0.01(-0.19,0.18)</td>
<td>-0.01(-0.19,0.18)</td>
<td>-0.01(-0.19,0.18)</td>
<td>-0.01(-0.19,0.18)</td>
</tr>
<tr>
<td>Middle wealth tertile household (ref: poorest tertile)</td>
<td>-0.53(-0.75,-0.32)</td>
<td>-0.53(-0.75,-0.32)</td>
<td>-0.53(-0.75,-0.32)</td>
<td>-0.53(-0.75,-0.32)</td>
</tr>
<tr>
<td>Highest wealth tertile household (ref: poorest tertile)</td>
<td></td>
<td></td>
<td></td>
<td>-1.14(-1.43,-0.85)</td>
</tr>
</tbody>
</table>

Random Part

| Level: state (N=19)                  | 0.56(0.22,1.21)           | 0.39(0.13,0.90)           | 0.25(0.07,0.59)           | 0.22(0.06,0.52)           |
| Level: district (N=59)               | 0.08(0.01,0.20)           | 0.07(0.00,0.20)           | 0.06(0.01,0.17)           | 0.04(0.00,0.13)           |
| Level: primary sampling unit (N=1876) | 0.16(0.06,0.33)           | 0.30(0.14,0.54)           | 0.35(0.16,0.59)           | 0.03(0.00,0.10)           |
| Level: household (N=39427)           | 1.00                      | 1.00                      | 1.00                      | 1.00                      |
| DIC                                  | 5338.78                   | 5330.95                   | 5324.69                   | 5167.23                   |

Discussion

This study found that urban Indian districts where the poor are segregated and isolated have higher child mortality rates than districts where the poor are less segregated. A unit increase in the likelihood that a poor household lives next to another poor household in poor neighbourhoods is associated with an increase in 1.3 child deaths in that district. This ecological correlation appears to remain even after taking into account a range of contextual (district level) and compositional (household level) socio-demographic factors that could confound this association. There appears to be some evidence of a ‘triple health jeopardy’ (Ross et al. 2001) of being poor, living in a poor district that is socioeconomically segregated from the rest of the city.

There may be a number of reasons for the relationship between socioeconomic segregation and child mortality. Poor nutrition, sub standard housing conditions and overcrowding increase susceptibility and exposure to infectious diseases which children under 5 are extremely susceptible to. Poor segregated areas, in spite of their greatest needs, often have inadequate services and restricted access to primary care services (Lobmayer & Wilkinson, 2002; Ross et al., 2001, Hastings 2009). Poor neighbourhoods are more
likely to be exposed to environmental pollutants (Anderton et al 1994, Ellen et al 2001, Cohen et al 2003, Ash and Fetter 2004), Hassing et al 2009). Socioeconomic segregation may enable the affluent access to the greatest resources and reinforce the social exclusion of the disadvantaged (Massey 1996, Massey and Denton 1988, Jargowsky 1996). The overall effect of income segregation is to reinforce social inequality (Dwyer 2010) through segregating the poor away from resources that enable social mobility, such as employment, transport, education, health care and other services. It may thus capture dimensions of social inequality other than poverty or deprivation rates that merely reflect social inequality.

There are a number of limitations to this study. This is a cross-sectional study showing associations between child mortality and a single measure of socioeconomic segregation. We cannot infer causality from observational evidence as we have not observed whether increasing socioeconomic segregation results in higher child mortality. While we have adjusted for a range of potential socio-economic confounders at the district and household levels, there may be other unobserved factors that cause the association of socioeconomic segregation with child mortality. In addition, we have not estimated any lagged effects of socioeconomic segregation on child mortality. If the association is causal, we don’t know what is the duration of exposure to a socioeconomically segregated neighbourhood that will generate higher child mortality risks. The measure of socioeconomic segregation was estimated using a mixture of survey and census data and so may not be accurately estimated. Other dimensions of socioeconomic segregation could not be measured due to data limitations. In particular, we have not been able to measure the spatial dimension of socio-economic segregation which may be particularly relevant in relation to discussions of social inequalities.

As urbanisation increases in India, there is a risk of increasing socioeconomic segregation of the poor and associated health risks. Indian cities already contain the largest concentration of the world’s population living in slums. With increasing urbanisation and associated economic pressures, the poor living within Indian cities are at risk of increasing isolation such as through slum resettlement programs far from employment, transport and public services networks. The results from this study cautions against such policies that increase the socioeconomic segregation of the poor, as this may lead to increased levels of child mortality.
References:


