Does Place of Birth Matter?
Spatial Analysis of Infant and Under-five Mortality Rates in India

Ankush Agrawal

Assistant Professor
Institute of Economic Growth
University of Delhi Enclave
North Campus, Delhi-110007
Email: ankush@iegindia.org

Abstract

This study examines spatial patterns in infant and under-five mortality rates in India at the levels of NSS-regions and Census-districts. We find significant spatial correlation both at the national and local level meaning that both global and local environment influences the mortality rates. We identify Assam-East as a spatial outlier. Besides, there exist several hot- and cold-spots in the country. The study further examines determinants of under-five mortality using spatial regression models. Contrary to the existing evidence, we find neither female labor force participation nor general level of modernization help reducing under-five mortality significantly. However, we find importance of reducing poverty, improving provisioning of public health interventions like antenatal care to women and immunization of children, and educating women. Integrating health awareness with health policy might be helpful in improving health outcomes. Using OLS without adjusting for spatial heterogeneity may lead to biased and inefficient parameter estimates.

Does Place of Birth Matter?

Spatial Analysis of Infant and Under-five Mortality Rates in India[1]

1. Introduction

India accounted for about 21 per cent of the children dying worldwide before turning five years of age in 2009, much higher than its share of either under-five or overall populations (United Nations Children's Fund 2011). Under-five mortality rate (U5MR) as well as infant mortality rate (IMR) in India, although lower than the South-Asian average, are higher than all its neighbors barring Bhutan, Pakistan and Myanmar.[2] The two rates are higher in India than the countries with the same level of economic and social development. There are concerns that India is not on course to meet the MDG target of reducing U5MR and IMR respectively to 41 and 27 per 1000 live births by year 2015.

There is substantial variation in infant and under-five mortality rates even within India. IMR in the state of Madhya Pradesh was five times than that in Kerala in 2006 (Office of the Registrar General, 2007). About half of the infants' deaths were concentrated in 22 per cent of the districts in the country according to the second National Family Health Survey (Deolalikar 2005). These variations across various regions continue to persist since a long time (Kohli 1971). Besides these administrative boundaries, the mortality rates of infants and children vary across social and economic classes, and males and females. It is important to understand the socioeconomic factors underlying these variations from policy perspective.

This study examines spatial distribution of infant and under-five mortality rates in India. To be specific, we examine spatial patterns in IMR at the NSS-region and U5MR at the district level. Analysis of health outcomes aggregated at such levels as NSS-regions and districts is useful from public health perspective where characteristics of population rather than individuals assume more importance. Two measures of spatial correlation, Moran's *I* and Geary' *c*, are used. Spatial

---

[1] *This work is still in progress.* I thank S Chandrasekhar, Dipankor Coondoo, Indrani Gupta, M H Suryanarayana and Finn Tarp for comments and suggestions, to Neelkanth N T for providing several useful references, and to Indira Gandhi Institute of Development Research and Institute of Economic Growth for institutional support.

[2] In 2009, U5MR and IMR (per 1000 live births) respectively for India's neighbors were 87 and 71 in Pakistan, 19 and 17 in China, 48 and 39 in Nepal, 79 and 52 in Bhutan, 52 and 41 in Bangladesh, 71 and 54 in Myanmar, and 15 and 13 in Sri Lanka (United Nations Children's Fund 2011).

correlation helps us understand whether the locations with high or low mortality rates of infant and children are surrounded by the locations having similar mortality rates. In addition, we identify spatial outliers. The spatial outliers are the locations that are both extreme and geographically homogeneous – the districts or regions having low value of U5MR/ IMR surrounded by high value of these variables and vice-versa. We then model U5MR by incorporating its spatial dependence using spatial econometric methods.

The paper is organized as follows. Section 2 discusses the literature on spatial patterns in demographic and health outcomes, and socioeconomic correlates of IMR and U5MR in India. In Section 3 we describe the data set. Spatial patterns in IMR and U5MR are also discussed. Section 4 presents the results of spatial regression models. This is followed by conclusion.

2. Literature Review

Persistence of poverty in many regions that are geographically contiguous is well documented. Cultural endowments, geographical capital and local conditions of a region like provision of public goods and infrastructure play important role in explaining geographic variations in poverty in China (Jalan and Ravallion 2002). Bloom, Canning and Saliva (2003) while analyzing a cross section of more than 100 countries found that unfavorable climatic conditions of a region like temperature and rainfall, poor infrastructural facilities like transportation, and unequal access to various public goods may create 'poverty traps'. Similarly some regions lag behind the others in educational and health attainments. For instance, childhood mortality among different West-African countries differs by a factor of four (Balk *et al*. 2003).

In India too, the states are quite diverse in terms of culture, geography and population composition and there are substantial differences in terms of economic and social development too.[3] A large number of poor people are concentrated in four neighboring states – Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh – known as BIMARU states.[4] Poverty head count ratio in rural India varied between three per cent in Saurashtra region of Gujarat to 73 per cent in Southern-Orissa in 2004 (Lanjouw and Murgai 2009). Similar differences across various regions

[3] States are sub-national units in India. Each state is a separate administrative unit.
[4] The term BIMARU is an acronym for the four states and was coined by demographer Ashish Bose. These states have higher fertility and poor indicators of socioeconomic development than the rest of states.

are observed in educational and health outcomes (Dreze and Sen 1995). Hence, it is important to understand spatial patterns in these dimensions and their socioeconomic correlates. Only few studies like Palmer-Jones and Sen (2006) examine spatial patterns in poverty. The study found that poverty and agricultural growth in rural areas is spatially dependent.

As far as reproductive and health outcomes in India are concerned, there is fair understanding of the regional patterns. Dyson and Moore (1983) point out North-South divide in fertility and sex-ratio (0-9 years) with North-Indian states having high fertility and sex-ratio favoring male children. The authors attribute the divide to differences in cultural factors, kinship structure and extent of female autonomy between the North- and South-Indian states. There are spatial patterns in fertility rates and sex-ratio. Districts with high fertility happen to be neighbors and belong to North-Indian states (Guilmoto and Rajan, 2001). Using pooled district level information from five different censuses of India from 1951 to 2001, the authors conclude that neighborhood effect is felt even at the distances greater than 300 km. They further compute Moran's I for fertility and find it to be very high for the districts located with in short distances. A recent study by Guilmoto (2008) examines spatial patterns in sex ratio of children (0-6 years) across the districts of India using spatial econometric tools. He finds 'hot spots' comprising of districts in Punjab, Haryana, and their neighbors where number of male children (0-6 years) outnumbers the females by perceptible margin. He uses spatial autoregressive error model to model the child sex-ratio in the Indian districts based on information from Census of India, 2001.

Studies on determinants of infant mortality in India using the state level data have highlighted the importance of economic and social factors like poverty and standard of living, mother's education, and housing and sanitation, and demographic factors like mean age at marriage.[5] State was used as unit of analysis owing to non-availability of further disaggregated data which restricted the degrees of freedom that these studies could exercise in selecting explanatory variables.

[5] A study by Jain (1985) used path analysis to examine the determinants of IMR for rural India based on information from a special survey conducted by the government in 1978. Beenstock and Sturdy (1990) estimated two separate regression equations for male and female IMR for the same survey of 1978.

Availability of district level estimates of infant and under-five mortality from the censuses in 1990s means district could be used as a unit of analysis. These estimates are indirect as the census collects information only on number of children ever born and surviving which are then used to derive infant and under-five mortality rates (United Nations 1983). Focus of the studies based on district level data has largely been to examine the role of women empowerment in gender-bias in the mortality rates.[6] Crucial role played by women's agency in child survival, reducing gender disparity, and promoting social progress is argued more forcefully by Dreze and Sen (1995). Most studies except Bhattacharya (2006) find empowering women through literacy and participation in work force can help reducing U5MR. Bhattacharya weighs importance of economic development vis-à-vis women's agency in reducing fertility and U5MR in India and finds, contrary to other studies, that general level of development and modernization has a greater impact on fertility and U5MR than women's agency. All the above studies find that higher literacy of females is significantly associated with lower U5MR.[7]

However, most of the above studies ignore spatial dependence in health outcomes. This despite the fact that studies by Kishor (1993) and Murthi, Guio and Dreze (1995) estimated spatial autoregressive error models and found evidence of significant spatial autocorrelation in U5MR across the districts. As health outcomes vary with structural characteristics of a region like climatic conditions, physical infrastructure, population, and social conditions; they are likely to be spatially dependent. In such cases, ordinary least squares (OLS) estimator is inefficient and perhaps inconsistent too under certain conditions. We, therefore, use spatial regression models which account for the spatial dependence while examining determinants of U5MR.

3. Data and Data Descriptive

There are three main sources of data on infant and under-five mortality rates for Indian population. These are Census of India, Civil Registration System (CRS) and Sample Registration System (SRS). The Census information is available decennially and reliable estimates can be generated at the district level. The SRS which is conducted annually provides reliable estimates at the national and the level of states. The CRS too is annual and provides estimates at the district

[6] Studies in this category include Kishor (1993), Murthy, Guio and Dreze (1995), Dreze and Murthi (2001), and Bhattacharya (2006).
[7] Masuy-Stroobant (2001) presents a review of international studies in this context. Also see Caldwell (1986).

and sub-district level. The SRS is considered the most reliable source for fertility and mortality estimates at the national and state level.[8]

For IMR, we use estimates from SRS for the period 2004-07. These estimates are available for 71 NSS regions of India.[9] Estimates of U5MR are generally derived using indirect method proposed by Brass from the Census data collected on number of children ever born and surviving. This study uses indirect estimates of U5MR at the district level based on Census of 2001 (Bhat and Zavier 2004). There were 593 districts in India at the time of Census of 2001.

We now describe the data on outcomes of interest, namely, IMR and U5MR. Infant mortality rate in India for 2004-06 (three year average) is 58 per 1000 live births. The five NSS regions having the highest IMR (in the range of 85-89 per 1000 live births) are from the states of Madhya Pradesh, Orissa, and Rajasthan. On the other hand Goa, both the regions of Kerala, and two of North-Eastern states Nagaland and Manipur are the ones having lowest IMR (between 13 and 18 per 1000 live births). The ranking of the regions in the four years (2004-07) has been fairly stable: the Spearman's rank correlation coefficient for IMR was in excess of 0.97 between any two years. This indicates persistence of IMR in the regions.

Fig-1 shows the district-wise U5MR in 2001. There seems to be no irregularity or sharp gradient across the districts indicating very high positive spatial correlation although the districts differ widely in their U5MRs. Three districts in Arunachal Pradesh have one of the highest U5MR in the country ranging between 228 and 258. Shivpuri and Chattarpur in Madhya Pradesh have U5MR of 212 and 199, and Badaun and Hardoi in Uttar Pradesh have U5MR of 205 and 202 respectively. Most of the districts in central, North-Eastern and North-Western parts of the country, and in Orissa have high U5MR. On the other hand, all 15 districts in Kerala are among the one having lowest U5MR in country. Thiruvanantpuram in Kerala has the lowest U5MR of 12 per 1000 live births. Barring some districts in Andhra Pradesh and Tamil Nadu, U5MR is low

[8] In India, registration of births and deaths was voluntary till 1969 and there were problems of both under- and incomplete-registration. Therefore, it was not possible to generate reliable data on vital statistics using 'Civil Registration'. On the other hand SRS is nearly complete in terms of coverage (Bhat 2002).
[9] Based on agro-climatic conditions prevailing in the states, the National Sample Survey (NSS) Organization divides the country into 78 regions/ zones. Thus a region is a group of districts from the same state but is not an administrative unit. Owing to the availability of data only 71 such regions have been used in this analysis.

in most of the Southern part of the country and Maharashtra. In addition to these high and low U5MR districts, we observe some ‘spatial outliers’: few districts have U5MR that is either too low or too high compared to their neighbors. We address this issue in one of the subsequent sections.

3.1 Spatial Autocorrelation

Spatial autocorrelation can be used to examine coincidence of value similarity with locational similarity (Anselin and Bera 1998). It can be computed at both national and local level. The former takes into account the spatial configuration of the whole country whereas the latter measures spatial association of a variable at a location with that in its local neighborhood.

Computation of spatial autocorrelation requires constructing a matrix, known as spatial weights matrix (W), to quantify the spatial proximity between each possible pair of observational units. The matrix can be constructed in different ways depending on the definition of neighbor employed. Simplest way is to construct binary connectivity matrix. An element w_{ij} of a binary connectivity matrix W equals unity if district j adjoins district i and equals zero in all the other cases. Similarly, nearest neighbor and centroid definitions can be used to define the spatial weights matrix (Anselin 2003). According to the latter, the element w_{ij} assumes value unity if the distance between centriods of the two districts is less than a specific value (say, δ) selected by the researcher and is zero if the distance equals or exceeds δ.[10] The centroid approach gives more precise results but computation of centroids requires information on geographic coordinates corresponding to each district. We construct the spatial weights matrix using centroid approach for the districts whereas we rely on binary connectivity matrix due to lack of availability of geographic coordinates for NSS-regions.

3.2 Global Spatial Autocorrelation

We estimate two often used measures of spatial autocorrelation Moran’s I and Geary’s c. Moran’s I and Geary’s c for global spatial autocorrelation are defined as:

[10] The δ must be chosen such that no district remains without neighbor. As large values of δ may result in loss of information, we omit both the districts in Andaman and Nicobar Islands, and lonely district in the island of Lakshadweep in this analysis. The three districts are far away from the mainland. Thus, study is confined to analysis of 590 districts.

$$I = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(Y_i - \overline{Y})(Y_j - \overline{Y})}{\frac{\sum_{i=1}^{N}(Y_i - \overline{Y})^2}{N-1}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}}$$

$$c = \frac{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(Y_i - \overline{Y})^2}{2}}{\frac{\sum_{i=1}^{N}(Y_i - \overline{Y})^2}{N-1}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}}$$

Here Y_i is infant or under-five mortality rate at *i*th region or district, $\bar{Y}$ denotes its average over all regions/ districts, N is total number of regions/ districts and w_{ij} is an element of the matrix W defined above corresponding to the pair of locations *i* and *j*.

Table-1 shows estimates of Moran's I and Geary's c for IMR at the NSS region level.[11] Positive and significant values of both the statistics in the above table means high IMR regions are surrounded by high IMR regions and low IMR regions are surrounded by low IMR regions at the national level. Value of both the measures remains almost constant in the study period.

Table-1 (last row) indicates presence of significant positive spatial correlation in U5MR too. Values of both Moran's I and Geary' c indicate that districts with high value of U5MR are surrounded by the districts with high value of U5MR and similar is true for districts with low U5MR.

3.3 Local Spatial Autocorrelation

The global spatial autocorrelation does not reveal existence of regional spatial patterns. Local spatial autocorrelation can be used to identify the locations that contribute most to the overall pattern of spatial clustering and to detect significant spatial clustering around individual locations. Local versions of Moran's *I* and Geary's *c*, denoted as I_j* and c_j* in this study for *j*th location, are defined as (Anselin 1995):

[11] We use Stata/SE 10 for all computations and maps. A set of user written commands by Maurizio Pisati was used for spatial correlations and regressions (Pisati 2001).

$$I_j^* = (Y_j - \overline{Y}) \sum_{i=1}^{N} w_{ij}(Y_i - \overline{Y})$$

$$c_j^* = \sum_{i=1}^{N} w_{ij}(Y_i - \overline{Y})^2$$

We compute local Moran's *I* for three year averaged (2004-06) IMR series. We find that one region, Assam East, is a spatial outlier: it has significant and negative value of *I* which means that this high IMR region is surrounded by low IMR regions. On the other hand, many regions have high IMR and are surrounded by high IMR regions. Most of these regions belong to two states: Madhya Pradesh and Rajasthan. Both the regions in Kerala, North and South Kerala, are low IMR regions surrounded by low IMR regions in the neighborhood. These observations indicate that characteristics of local neighboring areas influence IMR. We also compute local Geary's c. Again, Assam East has significantly different IMR from its neighborhood. However, Geary's c is insignificant for the remaining cases of high IMR regions surrounded by high IMR regions and low IMR regions surrounded by low IMR regions.[12]

We review the case of Assam. We find that both the regions of Assam (East and West) have IMR substantially higher than the remaining states in the North-Eastern part of the country. Since some neighbors of Assam West happen to be regions with high IMR, it is not a spatial outlier. We present information on some socioeconomic indicators of eight North-Eastern states of the country in Table-2. Assam does not differ perceptibly from the remaining seven states in other indicators of socioeconomic development: its per capita state domestic product (SDP) is slightly higher than that of Manipur, literacy is higher than that of Arunachal Pradesh and Meghalaya, urbanization than that of Sikkim. However, its IMR is substantially higher than the states having comparable indicators of socioeconomic development. The State Development Report of Assam (Government of India 2002, Chapter 4) blames poor access to basic health services like immunization, antenatal and maternal care, and inefficiencies in delivery of these services for high rates of infant and child mortality. However, the table clearly indicates that both Meghalaya and Nagaland are worse than Assam in terms of child vaccination, and antenatal and

[12] As a robustness check, we compute local Moran's I for the IMR average for 2005-07 and find similar results.

maternal care. Proportion of deliveries assisted by skilled health personnel is low in both Assam and Meghalaya. In terms of geography, whereas many of these states are mostly hilly terrain, four-fifth area of Assam is plain (Nayak 2009) which rules out natural disadvantage in access to health care facilities as an explanation for its high IMR. Thus, these socioeconomic factors seem unable to explain the high IMR of Assam although it may quite be possible that a combination of all these characteristics explains high IMR in Assam.[13]

As far as U5MR is concerned, local I values indicate that eight districts are spatial outliers: they are surrounded by the districts having significantly (at the five per cent level) dissimilar values of U5MR. On the other hand, there are 281 districts surrounded by the districts having significantly similar values. This means neighborhood characteristics do matter for child health outcomes. The eight districts are Kanpur Nagar and Balia (both in Uttar Pradesh), Jind and Kaithal (both from Haryana), Kargil (Jammu and Kashmir), North Tripura (Tripura), Hailakandi (Assam) and Darjiling (West Bengal). Both Kanpur Nagar and Balia have low U5MR (57 and 56 per 1000 live birth respectively) and are surrounded by districts with high U5MR; Jind, Kaithal, Kargil and North Tripura are districts with high level of U5MR neighbored by districts with low U5MR. However, on the basis of Geary's c, only Kanpur Nagar and Kargil among the above are spatial outliers.

3.4 Moran Plot

A convenient way to present the local spatial autocorrelations is Moran scatter plot. Moran scatter plot plots spatial lag of a standardize variable on the vertical axis (Wz) against original standardize variable (z) on the horizontal axis (here z denotes standardized values of Y). The graph is divided into four quadrants: geographical units lying in North-East quadrant are those with high IMR and are surrounded by high IMR locations (known as hot-spots), South-West quadrant are those having low IMR and are surrounded by low IMR locations (known as cold-spots), South-East quadrant are those having high IMR and are surrounded by low IMR locations, and North-West quadrant are those having low IMR and are surrounded by high IMR locations (Pisati 2001). Based on local spatial autocorrelation, it is possible to identify locations

[13] We do not consider political factors and incidences of violence in this study.

that are both extreme and geographically homogeneous. The geographical units lying in South-East and North-West quadrants in the Moran plot belong to this category.

Fig-2 presents the Moran plot for IMR (2004-06) at the region level. By construction, slope of the oblique line is the value of Moran's I. The hot spots mainly belong to the states of Madhya Pradesh, Orissa, Rajasthan and Uttar Pradesh and the cold spots are from Kerala and North-East India. Both the regions in Assam belong to the South-East quadrant which means the state is a spatial outlier in the North-East India. One region of Tamil Nadu, Inland region, too belongs to this quadrant: Tamil Nadu has high disparity with IMR in 2004 ranged from 33 in coastal areas to 60 in inland. Delhi and Saurashtra region of Gujarat are the ones having low IMR but are surrounded by high IMR regions. We do not show Moran plot for U5MR because number of districts is too large to reveal any conclusion from the plot.

4. Socioeconomic Correlates of U5MR

The above discussion indicates presence of significant spatial autocorrelation in IMR and U5MR. As noted earlier, such correlation in dependent variable may lead to correlation among the error terms rendering the OLS estimator inappropriate owing to violation of its underlying assumptions. We therefore use spatial regression to examine determinants of U5MR.

4.1 Spatial Regression

The 'spatial effect' discussed earlier can be modeled in the following two ways:

- *Spatial Lag Model*: If dependent variable, Y, is correlated with weighted average of its value in the neighborhood and other locations, this relationship can be expressed as

  $$Y = \rho WY + \beta X + \varepsilon$$

  Here, ρ is spatial lag parameter, W is spatial weights matrix, X is vector of explanatory variables and β is corresponding coefficient vector. It is assumed here that the error terms εs are identically and independently distributed (iids) although one can correct for heteroscedasticity (Anselin 2003). The OLS would provide biased *and* inconsistent estimates of the model parameters due to simultaneity bias.

- *Spatial Error Model*: If the spatial dependence enters the model through the error term, ε, we have the spatial error model:

$$Y = \beta X + \varepsilon$$

$$\varepsilon = \lambda W \varepsilon + \mu$$

  Here, λ is spatial autoregressive parameter and the errors μs are iids. Thus, it is a special case of regression with non-spherical error term and OLS, although unbiased, is inefficient.

Thus, the spatial regression takes into account proximity among geographical units through the weights matrix W. Both spatial error and spatial lag models are estimated by maximizing the corresponding likelihood functions discussed further in Anselin (2003). The dependent variable (Y) is U5MR and explanatory variables (X) are discussed in the next sub section.

4.2 Explanatory Variables

Choice of explanatory variables in this study is guided by the existing literature and an analytical framework proposed by Mosley and Chen (1984) to study the determinants of child survival in developing countries.[14] Table-3 summarizes the variables used in this study, their definition, and descriptive statistics for the 590 districts. We classify the variables into four categories: those describing social strata, those capturing extent of economic development or modernization, indicator of health care services, and women empowerment.

We use proportion of 'Scheduled Castes' and 'Scheduled Tribes' in districts to control for social strata. The two groups respectively accounted for 16 and eight percent of India's total population in 2001. Their share in total population across the districts varied from highest of 54 and 98 (respectively for Scheduled Castes and Scheduled Tribes) to zero. Proportion of 'Muslims' is used to capture religious composition. Muslims are second largest religious group in the country. They accounted for 13.4 per cent of Indian population in 2001.

[14] The framework was an attempt to integrate research methods employed by social and medical scientists, and was based on the premise that all social and economic determinants of child mortality necessarily operate via a common set of proximate determinants to exert an impact on mortality (Mosley and Chen 1984, p. 25). Although the framework seems more appropriate to study the determinants at household level, the same has extensively been used by studies involving aggregate level analysis.

We control for extent of urbanization, poverty, population density, structure of workforce, female literacy and availability of drainage facility. Nearly 28 per cent of all population resided in urban areas in 2001. As reliable estimates of poverty (using official poverty lines) at the district level are lacking, we use proportion of households in a district who do not own any of the following assets as its proxy: Radio, Television, Telephone, Bicycle, Scooter, Motor cycle, Car, Jeep. 'Population density' is expected to capture the adverse effect of congestion on health outcomes. There is substantial variation in population density across the districts ranging from 30,000 persons per square km in North-East Delhi to 2 persons per square km in Lahul district of Himachal Pradesh. 'Proportion of agriculture workers in total workforce' would capture the extent of dependence on agriculture. We use 'proportion of households who do not have any drainage facility for waste water outlet' as a proxy for availability of basic amenities. All these variables have been used in the literature as a proxy for general level of development.

Availability and utilization of health care services is measured using proportion of women who received at least three antenatal checkups (ANC) during pregnancy. This information was obtained from Reproductive and Child Health Survey of 2002-04. Only the women who gave last birth on or after 1-1-1999 were considered. 'Proportion of female employed as main worker to those in the labor force' is used to measure female labor force participation. Main worker is defined as all those engaged in any economic productive activity.

Information on all the socioeconomic correlates except utilization of health care services at the district level is available from the Census of India, 2001. Information on availability of health care infrastructure is not available from the Census of India 2001 so far, hence the same has been obtained from the second wave of District Level Household Survey (DLHS) conducted during 2002-04.

4.3 Results

We estimate OLS as well as spatial lag and error models (Table 4). First column shows OLS estimates, columns 2a-2c show estimates obtained using spatial error model and columns 3a-3c show estimates of spatial lag model. We estimate three different models for each of the spatial regression: first column in each spatial regression (cols 2a and 3a) uses share of population in urban areas as a proxy for general level of development / modernization which is replaced by ratio of workforce in agriculture to that in non-agriculture sector in the second column and share of agriculture workers in the workforce in the third.

Given our earlier observations on significant spatial correlations in U5MR, we carry out diagnostic testing of the residuals obtained from the OLS estimation. Significant Moran's I indicate that the residuals of the OLS model are positively spatially correlated. We then carry out robust LM (Lagrange Multiplier) test for both spatial lag and spatial error models. These test statistics are significant. Although these tests do not indicate whether to use spatial error or spatial lag model, they provide enough evidence to use either spatial lag or spatial error model. We therefore estimate both the error and lag models. A crude look at the table indicates that OLS tend to overestimate the model parameter in the presence of spatial effects.

Coefficients of the terms capturing spatial effects, viz., λ in spatial error model and ρ in spatial lag model, are statistically significant. High positive value of λ indicates that the variables like cultural practices towards child care and mother's nutritional status that could not be controlled in the model (and are captured by the error term) are positively correlated across the neighboring districts. On the other hand, high value of ρ indicates substantial spatial dependence in U5MR across the neighboring districts. Although, we discuss both the spatial error and lag models here, we prefer the spatial lag model because of higher 'Sigma'. Moreover, pseudo R-square for the lag model is also substantially higher.

We find that U5MR is higher in the districts having higher share of Scheduled Caste population. Scheduled Castes and Scheduled Tribes are among the most depressed classes in India in terms of their standard of living.[15] Our finding on Scheduled Castes is in line with that of all the studies discussed in the literature section: U5MR is higher in the districts having high proportion of Scheduled Castes. It has been observed that group identities like castes are associated with differences in access to health care infrastructure (Banerjee and Somanathan 2007). As the coefficient for Scheduled Tribes was statistically insignificant, we do not find any support for the claim by Murthi, Guio and Dreze (1995) that tribal lifestyle has some healthy aspects like less

[15] Scheduled Castes are believed to be among the lowest in caste hierarchy system in India. Scheduled Tribes are typically indigenous people who still remain geographically isolated and depend on primitive practices for their livelihood.

crowding and low interface to pollution which reduces U5MR in the districts where their population is higher. Our finding, however, is consistent with that of Kishor (1993) and Bhattacharya (2006) for 1981.[16]

Higher proportion of Muslims too is associated with high U5MR across the districts. Muslims constitute second largest religious group in the country. Bhattacharya (2006) too observed positive and significant relationship between the two. Many studies on cross section of countries (for example, Caldwell 1986, and Filmer and Pritchett 1999) find that infant and child mortality rates tend to increase with proportion of Muslim population in the countries.[17] As we control for availability and utilization of health care, these positive associations of U5MR with proportion of SCs and Muslims may reflect the role of inherent cultural practices among these groups.

The role of health care infrastructure in improving the health status of the population is long known. Utilization of medical facilities, captured by proportion of women in a district who had antenatal checkups at least thrice during pregnancy, helps reducing U5MR significantly. This finding is consistent with the findings of studies by Murthi, Guio and Dreze (1995) and Bhattacharya (2006). Both the studies use proportion of villages having any medical facility in a district as its proxy. As the care received during antenatal period is more sensitive to the infant and neonatal mortality than the U5MR, we use proportion of children immunized for BCG, Polio, DPT and Measles as an alterative measure.[18] Sign of all the significant coefficients did not show any change.

Poverty is significantly associated with U5MR: the latter is higher in the districts where proportion of household not having any of the specified assets is high. Poverty leads to poor

[16] For 1991, Bhattacharya (2006) finds positive and significant association between proportion of the Scheduled Tribe population and U5MR.

[17] It may be noted that infant and child mortality rates are lower for Muslims as a group than those for Hindus in India (Shariff 1995, Bhalotra, Valente, and van Soest 2009). However, if Muslims happen to be residing in the districts having high U5MR; one might observe a positive relationship between proportion of Muslims and U5MR at the district level.

[18] These results have not been shown in the table.

health outcomes including higher infant and child mortality (Jain 1985 and Deaton 2003). The same relationship was observed by Murthi, Guio and Dreze (1995).[19]

Districts with high population density have significantly lower U5MR. The negative association between population density and U5MR may reflect the fact that availability and accessibility of health care services is higher in the areas where population density is high. Better availability of transportation and other infrastructure in such areas may help in seeking health care whenever required.

We now discuss the variables used to examine the role of general level of development. Positive association between economic development (i.e. improvement in standards of living, technological progress and availability of basic amenities) and better health outcomes is well known in the form of Preston curve (Preston 1975). We use share of district's population residing in urban areas, share of workforce employed in agriculture, and share of workforce in non-agriculture to that in agriculture to measure the level of economic development.[20] None of the three proxy variables used by us for general level of development is significantly associated with U5MR. This finding does not corroborate with all the earlier studies. Kishor (1993) and Murthi, Guio and Dreze (1995) find negative association between proportion of urban population and U5MR whereas Bhattacharya (2006) finds it positive and significant. Bhattacharya (2006) also finds positive and significant association between proportion of agricultural workers and U5MR. Kishor (1993) however finds no significant association between productivity in either agricultural or industry and U5MR. This finding means once extent of utilization of health care services, level of poverty, literacy and social group composition of population is controlled for, the districts with higher level of development (captured by the three variables) have no natural advantage over the districts with low level of development. In other words, it is possible to achieve low U5MR in less developed districts by targeting policy interventions on empowering the economic status of poor, improving health care accessibility and education. Thus there exists a huge potential for reduction of the under-five mortality in India. Availability of drainage facility too is not a significant predictor of U5MR. Since, it is viewed as an indicator of general

[19] Kishor (1993) uses proportion of agricultural laborers as a proxy for landless individuals and finds it is *not* significantly associated wither with the male or female child mortality ratio.

[20] Last two of these three variables also capture the structure of local economy.

level of modernization, it endorses the above results. Magnitude of significant coefficients in all the three cases (columns suffixed a, b and c in Table-4) for a given model do not differ markedly which indicates that the parameter estimates are relatively stable.

If the females are empowered enough to take decisions related to daily household activities, they would have more say in child bearing. Such women can avail health care facilities, if the need arises, even in the absence the male members of households.[21] On the other hand, participation in work means women will have less time to take care of child. If the two effects cancel each other, one may not observe a significant relationship as happens in our case. It may be noted that as women may be forced to work due to poverty, it is necessary to control for poverty (Kishor 1993). Both the studies based on Census of 1981 by Murthi, Guio and Dreze (1995) and Kishor (1993) find higher extent of participation in labor force by women to be associated with higher U5MR in 1981. Bhattacharya (2006) finds negative association between the two for 1991. We however do not find significant association between proportion of women participating in labor force and U5MR.

Higher proportion of literate females in a district helps reducing U5MR. Connection between literacy and child mortality is the most obvious because of its effects on health related behavior. Therefore, female literacy is perhaps the most significant socioeconomic characteristic of any region that explains variations in child mortality rates. This finding corroborates the findings of all the earlier studies.

In sum, we find that neither the female labor force participation (i.e., women's agency, if the same has been used to measure it) nor the general level of modernization help reducing U5MR significantly. However, we do find an important role of reducing poverty, educating women and providing health care to women.

[21] Female labor force participation has been used in the literature as a proxy for extent of women empowerment. However, this holds true only to the extent female labor force participation correctly captures the extent of autonomy among women. As we are unsure about it, we shall prefer using the word 'female labor force participation'.

5. Conclusion

This study examines the spatial patterns in infant and under-five mortality rates across the NSS regions and Census districts for Indian population. We find evidence for significant global and local spatial correlation in IMR and U5MR across sub-national units. Structural characteristics of a region like population and physical infrastructure, geographical, climatic and cultural factors may provide the linkages between the two. Besides these socioeconomic endowments, role of health care services is also important: to the extent such deaths in one geographic unit is affected by provision of curative and preventive health care services of neighboring units; one may get significant spatial association in mortality.

The local spatial autocorrelation helped us in identifying 'geographic mortality traps' in some areas of country. Therefore, it is quite possible that various risks to health tend to cluster in some of areas. We identify Assam East as one such region which requires further exploration. Presence of significant local and global spatial autocorrelation means both immediate and wider social environment influence infant mortality.

The study further examines the determinants of U5MR in the country. We use spatial lag model for this purpose. Contrary to the existing evidence, we find that neither the female labor force participation nor the general level of modernization help reducing U5MR significantly. However, we do find the importance of addressing poverty. Public health interventions like providing antenatal care to women and immunization to children were significant. This means ensuring accessibility and availability of these services is critical for child health outcomes. Multiple deprivations in the forms of poverty and poor health co-exist across the districts. Significant association of female literacy with U5MR indicates that integrating health awareness with health policy might be helpful in improving health outcomes. We also find that OLS overestimates the model parameters.

References

Anselin, Luc. 1995. "Local Indicators of Spatial Association-LISA." *Geographical Analysis*, 27:2, pp. 93-115.

Anselin, Luc and Anil K Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics," in *Handbook of Applied Economic Statistics*. Aman Ullah and David E A Giles eds: Marcel Dekker, Inc., pp. 237-90.

Anselin, Luc. 2003. "Spatial Econometrics," in *A Companion to Theoretical Econometrics*. Badi H. Baltagi ed: Blavckwell Publishing.

Balk, Deborah, Tom Pullum, Adam Storeygard, Fern Greenwell, and Melissa Neuman. 2003. "Spatial Analysis of Childhood Mortality in West Africa." *DHS Geographic Studies 1*. ORC Macro and Center for International Earth Science Information Network (CIESIN), Columbia University: Calverton, Maryland, USA.

Banerjee, Abhijit and Rohini Somanathan. 2007. "The Political Economy of Public goods: Some Evidence from India." *Journal of Development Economics*, 82, pp. 287-314.

Beenstock, Michael and Patricia Sturdy. 1990. "The Determinants of Infant Mortality in Regional India." *World Development*, 18:3, pp. 443-53.

Bhalotra, Sonia, Christine Valente, and Arthur van Soest. 2009. "The Puzzle of Muslim Advantage in Child Survival in India." *No 2009–13*. Tilburg University.

Bhat, P N Mari. 2002. "Completeness of India's Sample Registration System: An Assessment Using the General Growth Balance Method." *Population Studies*, 56:2, pp. pp. 119-34.

Bhat, P N Mari and A J Francis Zavier. 2004. "Some Indirect Methods for Estimation of Fertility and Contraceptive Use at District Level." Population Resaerch Centre, Institute of Economic Growth: Delhi.

Bhattacharya, Prabir C. 2006. "Economic development, gender inequality, and demographic outcomes: Evidence from India." *Population and Development Review*, 32:2, pp. 263-92.

Bloom, David, David Canning, and Selvuilla. 2003. "Geography and Poverty traps." *Journal of Economic Growth*.

Caldwell, John C. 1986. "Routes to Low Mortality in Poor Countries." *Population and Development Review*, 12:2, pp. 171-220.

Deaton, Angus. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature*, pp. 113-58.

Deolalikar, Anil B. 2005. "Infant and Under-Five Mortality in India: Levels, Patterns and Correlates."

Dreze, Jean and Amartya Sen. 1995. *India: Economic Development and Social Opportunity*. Delhi: Oxford University Press.

Dreze, Jean and Mamta Murthi. 2001. "Fertility, Education and Development: Evidence from India." *Population and Development Review*, 27:1, pp. 33-63.

Dyson, Tim and Mick Moore. 1983. "On Kinship Structure, Female Autonomy, and Demographic Behavior in India." *Population and Development Review*, pp. 35-60.

Deolalikar, Anil B. 2005. "Infant and Under-Five Mortality in India: Levels, Patterns and Correlates."

Filmer, Deon and Lant Pritchette. 1999. "The impact of public spending on health: does money matter?" *Social Science and Medicine*, 49, pp. 1309-23.

Frisbie, W Parker. 2005. "Infant mortality," in *Handbook of Population*. Dudley L Poston and Michael Micklin eds. New York: Kluwer Academic.

Government of India. 2002. "State Development Report: Assam." Planning Commission: New Delhi.

Guilmoto, C. Z. and S. Irudaya Rajan. 2001. "Spatial Patterns of Fertility Change in Indian Districts." *Population and Development Review*, 27:4, pp. 713-38.

Guilmoto, C. Z. 2008. "Economic, Social and Spatial Dimensions of India’s Excess Child Masculinity." *Population*, 63:1, pp. 93-122.

International Institute for Population Sciences (IIPS). 2006. "Reproductive and Child Health: District Level Household Survey (2002-04) & Facility Survey (2003) India ". International Institute for Population Sciences: Mumbai.

International Institute for Population Sciences (IIPS) and ORC Macro. 2000. *National Family Health Survey (NFHS-2), 1998–99: India*. Mumbai: IIPS.

Jain, A K. 1985. "Determinants of Regional Variations in Infant Mortality in Rural India." *Population Studies*, 39:3, pp. 407-24.

Jalan, Jyotsna and Martin Ravallion. 2002. "Geographic poverty traps? A micro model of consumption growth in rural China." *Journal of Applied Econometrics*, 17, pp. 329-46.

Kishor, Sunita. 1993. "May God give Sons to All: Gender and Child Mortality in India." *American Sociological Review*, 58:April, pp. 247-65.

Kohli, K L. 1971. "Spatial variations of mortality in India; 1951-61."

Lanjouw, Peter and Rinku Murgai. 2009. "Poverty decline, agricultural wages, and nonfarm employment in rural India: 1983–2004." *Agricultural Economics*, 40, pp. 243-63.

Masuy-Stroobant, Godelieve. 2001. "The determinants of infant mortality: how far are conceptual frameworks really modelled?" Département des Sciences de la Population et du Développement, Université catholique de Louvain.

Mosley, W Henry and Lincoln C Chen. 1984. "An Analytical Framework for the Study of Child survival in Developing Countries." *Population and Development Review, Supplement to Volume 10*, pp. 25-45.

Murthi, Mamta, Anne-Catherine Guio, and Jean Dreze. 1995. "Mortality, Fertility and, Gender Bias in India: A District-Level Analysis." *Population and Development Review*, 21:4, pp. 745-82.

Nayak, Purusottam. 2009. "Human Development Reports on North-East India: A Bird's Eye View." *MPRA Paper No. 17015*. Munich Personal RePEc Archive.

Office of the Registrar General, India. 2007. "Sample Registration System Statistical Report 2006." Ministry of Home Affairs, Government of India: New Delhi.

Palmer-Jones, Richard and Kunal Sen. 2006. "It is where you are that matters: the spatial determinants of rural poverty in India." *Agricultural Economics* 34, pp. 229-42.

Pisati, Maurizio. 2001. "Tools for spatial data analysis." *Stata Technical Bulletin*, 60, pp. 21-37.

Preston, Samuel H. 1975. "The Changing Relation between Mortality and Level of Economic Development." *Population Studies*, 29:2, pp. 231-48.

Reserve Bank of India. 2007. "Handbook of Statistics on Indian Economy." Reserve Bank of India: Mumbai.

Shariff, Abusaleh. 1995. "Socio-Economic and Demographic Differentials between Hindus and Muslims in India." *Economic and Political Weekly*, 30:46, pp. 2947-53.

United Nations. 1983. "Indirect Techniques for Demographic Estimation, Manual X." Department of International Economic and Social Affairs, United Nations: New York.

United Nations Children's Fund. 2011. "The State of the World's Children 2011."

Table 1: Global Spatial Autocorrelation in IMR/ U5MR

Variable (Year)	Moran's I	Geary's c
IMR (2004)	0.48	0.49
IMR (2005)	0.51	0.48
IMR (2006)	0.50	0.49
IMR (2007)	0.52	0.46
IMR (2004-06)‡	0.50	0.48
IMR (2005-07) ‡	0.52	0.47
U5MR (2001)	0.33	0.64

Note: p-value for all estimates is 0.00. ‡ indicates IMR is average of three year estimates.

Table 2: Selected Socioeconomic Indicators: North-Eastern States

State	IMR (per 1000 live births)	Per capita SDP (FC) (Rs)	Literacy (%)	Urban Population (%)	Delivery (%)	Vaccination BCG (%)	Vaccination DPT (%)	Women receiving ANC (%)	Women receiving IFA (%)
Year	2001	2000-01	2001	2001	1998-99				
Arunachal Pradesh	39	15003	54.74	27.30	31.9	54.2	41.8	40.5	47.6
Assam	73	12447	64.28	12.90	21.4	53.5	37.5	30.8	45.3
Manipur	20	12157	68.87	26.58	53.9	71.0	59.1	54.4	38.0
Meghalaya	56	15200	63.31	19.58	20.6	46.1	25.4	31.3	40.6
Mizoram	19	16635	88.49	49.63	67.5	88.2	69.5	75.8	62.0
Nagaland	na	15746	67.11	17.23	32.8	46.1	29.6	23.1	26.7
Sikkim	42	15305	69.68	11.07	35.1	76.5	62.5	42.6	50.4
Tripura	39	14933	73.66	17.05	na	na	na	na	na
India	66	16172	65.20	27.81	42.3	71.6	55.1	43.8	47.5

Notes: 'Delivery' indicates proportion of deliveries assisted by skilled personnel; 'Vaccination' is proportion of children (12-23 months) receiving the BCG vaccination for BCG and at least three doses of DPT; Women receiving ANC is proportion of pregnant women who received at least three ante natal checkups during pregnancy, and those receiving IFA is the proportion receiving iron and folic acid tables or syrups for duration of three months or more during pregnancy. Skilled personnel include doctors, nurses, ANMs or other health personnel. 'na' indicates information not available. SDP per capita is at 1999-00 prices.
Sources: IMR: SRS (2001), Per capita SDP: Reserve Bank of India (2007), Literacy and Urban population: Census of India (2001), and remaining characteristics: IIPS (2000).

Table 3: Description of the Variables

Characteristic / Variable	Definition	Mean	SD	Min	Max
U5MR	Proportion of children dying before the age of five years per 1000 live births	105.09	40.77	12	258
Social Strata					
Scheduled Caste	Proportion (%) of 'Scheduled Caste' population	14.92	8.84	0	54.03
Scheduled Tribe	Proportion (%) of 'Scheduled Tribe' population	15.81	25.59	0	98.09
Muslim	Proportion (%) of 'Muslim' population	11.65	15.12	0	98.49
Economic Development / Modernization					
Urbanization	Share (%) of population residing in urban areas	25.75	20.47	0.00	100.00
Poverty	Proportion (%) of households not owing any of the following assets: Radio, Transistor, Telephone, Bicycle, Scooter, Motor cycle, Car, Jeep.	7.02	3.20	1.05	22.27
Population Density	Log of population density (population residing per square km)	5.77	1.22	0.00	10.29
Agriculture	Proportion (%) of agriculture workers in total population	12.83	8.51	0.05	51.07
Non-agriculture	Proportion (%) of non-agriculture to agriculture workers in total population	36.94	30.70	0.07	172.45
Female Literacy	Proportion (%) of literate females in a district (crude female literacy rate)	52.48	15.45	18.58	96.26
No drain	Proportion (%) of households having without any drainage facility for waste water outlet	57.13	23.64	2.64	96.34
Health Care Services					
Antenatal checkups	Proportion (%) of women who received at least three antenatal checkups (ANC) during pregnancy	52.02	26.57	5.40	100.00
Women Empowerment					
Female work participation	Proportion (%) of females categorized as 'main workers'.	28.91	11.97	4.66	65.20

Notes: The statistics presented here corresponds to sample of 590 districts for which analysis is carried out and is not weighted by their respective populations. Hence, mean values may not equal to those for the country.
Data Sources: For U5MR: Bhat and Zavier (2000), for ANC: IIPS (2006), for remaining variables: Census of India, 2001.

Table 4: Determinants of U5MR

	OLS	Spatial Error			Spatial Lag		
	(1)	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Proportion of 'Scheduled Castes'	0.50***	0.28*	0.29*	0.24	0.44***	0.46***	0.44***
	(0.00)	(0.07)	(0.06)	(0.13)	(0.00)	(0.00)	(0.00)
Proportion of 'Scheduled Tribes'	-0.00	-0.03	-0.04	-0.03	-0.00	-0.01	-0.00
	(0.95)	(0.59)	(0.49)	(0.61)	(0.95)	(0.85)	(1.00)
Proportion of Muslims	0.07	0.33***	0.34***	0.32***	0.19***	0.17**	0.18**
	(0.43)	(0.00)	(0.00)	(0.00)	(0.01)	(0.03)	(0.01)
Urbanization	0.24***	-0.03			0.05		
	(0.00)	(0.62)			(0.43)		
Ratio of agriculture to			-0.02			-0.08**	
non-agriculture workers			(0.63)			(0.04)	
Share of agricultural workers				0.24			-0.08
in labor force				(0.15)			(0.56)
Proportion of population	-0.06	1.40***	1.44***	1.32***	1.12***	1.43***	1.21***
without any specified assets	(0.89)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Population density (log)	-6.11***	-3.35**	-3.63***	-3.29**	-4.99***	-4.36***	-4.69***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Female literacy	-0.89***	-0.89***	-0.91***	-0.87***	-0.63***	-0.68***	-0.63***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Proportion of population	0.15**	-0.04	-0.03	-0.03	0.02	0.02	0.01
with no drainage facility	(0.01)	(0.53)	(0.65)	(0.62)	(0.69)	(0.65)	(0.79)
Antenatal checkups	-0.76***	-0.35***	-0.35***	-0.35***	-0.27***	-0.24***	-0.25***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female work	-0.05	-0.01	-0.00	0.05	-0.05	-0.09	-0.09
participation rate	(0.73)	(0.94)	(0.98)	(0.70)	(0.67)	(0.43)	(0.45)
Intercept	205.69***	174.98***	175.73***	168.98***	90.68***	90.74***	91.11***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R-sq‡	0.588	0.520	0.524	0.515	0.718	0.720	0.718
Moran's I	27.83***						
Robust LM (error)	317.40***						
Robust LM (lag)	28.97***						
λ		0.92***	0.92***	0.92***			
		(0.00)	(0.00)	(0.00)			
ρ					0.69***	0.69***	0.69***
					(0.00)	(0.00)	(0.00)
Sigma (Maximum		20.12***	20.13***	20.09***	21.64***	21.57***	21.63***
Likelihood root MSE)		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No of observations	590	590	590	590	590	590	590

Notes: p-values in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$; ‡ indicates that pseudo R-square is used in spatial models. Pseudo R-square is squared correlation between observed and predicted value of the U5MR. Note that it is not comparable to the (adjusted) R-square of the OLS.

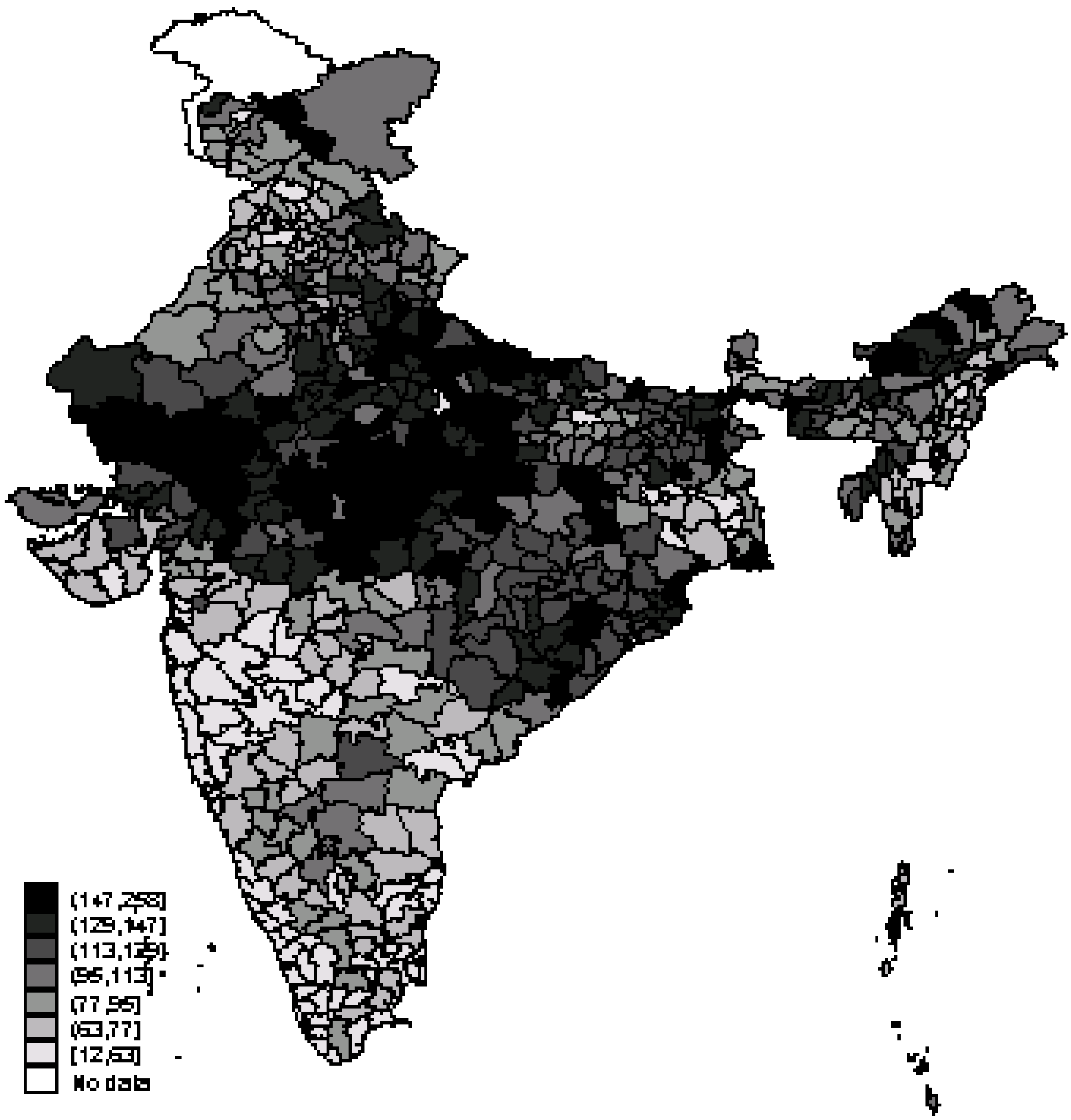

Figure 1: Under-five Mortality Rate in Indian Districts: 2001

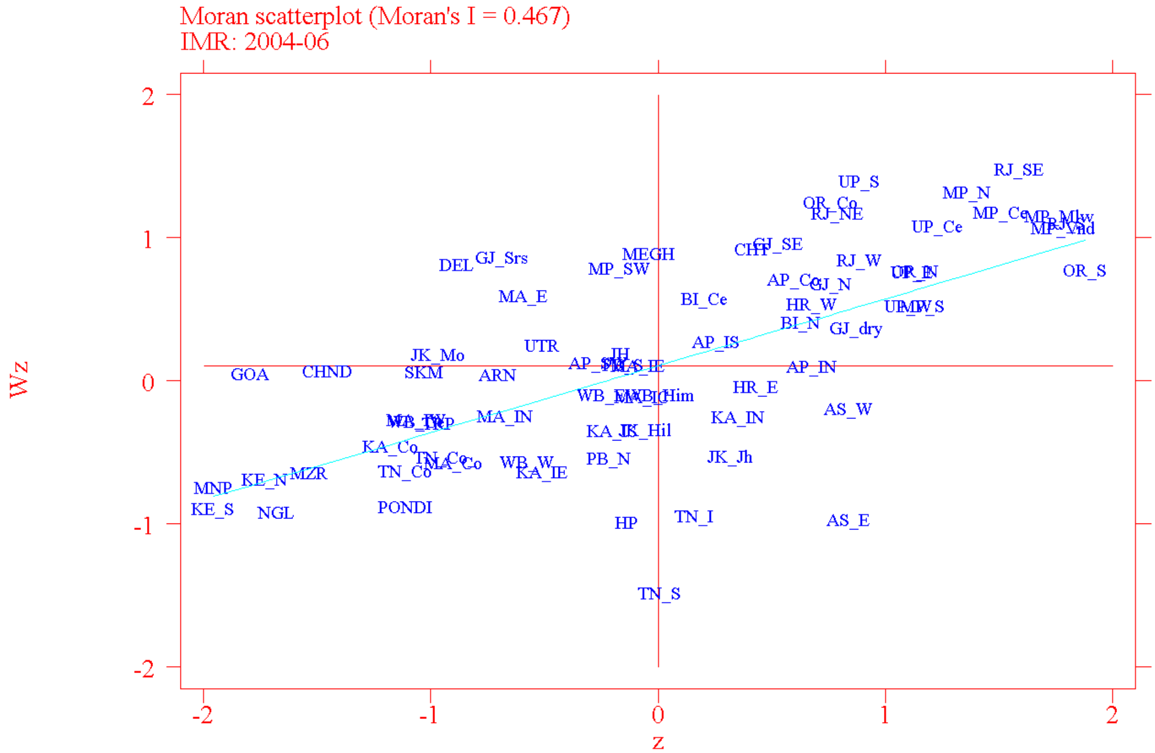

Figure 2: Moran scatter plot for IMR (2004-06): NSS regions