Discussion Paper ERU/2010 – 01
March, 2010

Decomposing the Difference in Poverty Incidences: A Spatial Reformulation

Somnath Chattopadhyay, Amita Majumder
(Indian Statistical Institute, Kolkata)

Hasanur Jaman[1]

Address for correspondence:

Somnath Chattopadhyay

Senior Research Fellow,

Economic Research Unit,

Indian Statistical Institute,

203 B. T. Road,

Kolkata 700108,

India

e-mail: pikluchatterjee@gmail.com

FAX: 91-33-25778893

[1] The author is an Assistant Teacher at Bidhannagar Municipal School, Salt Lake, Kolkata.

Summary

The Oaxaca decomposition methodology of finding the difference in the incidences of poverty between two regions in terms of characteristics and coefficients effects has been reformulated in a spatial framework. The incorporation of the spatial regression technique in the conventional regression framework provides a markedly different picture of the characteristics and coefficients effects for rural West Bengal, India.

JEL Classification: C20, I30, R22

Keywords: Poverty Incidence, Head Count Ratio, Spatial Regression, Delta Method.

1. Introduction

Poverty mapping is an important tool used in the developing countries to identify ways of improving living standards. There have been numerous studies linking incidence of poverty *(a binary variable*) to various socio economic explanatory variables using the logit/probit models worldwide (see e.g., Geda et al., 2001; Bokosi, 2007) as well as with reference to India (see e.g., Gang et al., 2002, Bigman et al., 2001). Using a slightly different approach, the World Bank methodology (World Bank, 2005) considers the *logarithm of the ratio of income to poverty line* as the dependent variable (which is a common way of allowing for the log normality of the variable) instead of a binary dependent variable in the logit/probit regression. The reason is that in the binary models some information is lost and the resulting logit or probit regression is relatively sensitive to specification errors.

In this paper we analyze the incidence of poverty in the context of rural West Bengal using the World Bank approach, which has also been used in Bhaumik et al., 2006 for studying difference in poverty incidences between Serbians and Albanians in Kosovo using Living Standard Measurement Survey. We relate incidence of poverty to various socio economic indicators using the

World Bank methodology in a spatial framework taking into consideration the spatial correlation based on the geographic and economic distances among the households within the districts of a region. We study the disparity in poverty estimates (in particular, the Head Count Ratio (HCR)) between North Bengal and South Bengal, a traditional division of West Bengal into two parts by the river Hoogly. We then decompose the difference between the poverty estimates into a *characteristics effect* (which can be interpreted in terms of resource constraint) and a *coefficients effect* (which can be interpreted in terms of efficiency constraint) using the Oaxaca decomposition method (Oaxaca, 1973). While the underlying assumption in the World Bank Methodology as implemented in Bhaumik et al., 2006, is that the probability of a person being poor is independent of the probabilities of his neighbours being poor, an assumption which may not be realistic in practice, in the spatial framework the assumption of independence is relaxed by suitably defining "neighbours" using geographical and economic proximity.[2]

To our knowledge, in the existing literature on spatial dimension of poverty, the spatial effect is incorporated either as an area characteristic entering the

[2] Our model is, in fact, a spatial reformulation of the model in Bhaumik et al. ,2006.

general logit/probit type regression analysis as an additional regressor (see Petrucci et al., 2003) or the spatial effect is incorporated in the *incidence of poverty*, the dependent variable in the regression of incidence of poverty (of a region) on a set of poverty correlates (see Voss et al, 2006; Chandrasiri et al., 2008).

Our paper is distinct in the sense that we *directly model the spatial dependence in household level monthly per-capita expenditure values* (in fact logarithm of monthly per-capita expenditure values divided by the poverty line) and calculate the probability of incidence of poverty at the household level, whence incidence of poverty of a region is estimated. Applying the Oaxaca decomposition methodology in the context of spatial regression analysis is a new dimension in poverty incidence analysis. The spatial effect introduced in the traditional Oaxaca decomposition methodology is particularly meaningful in the context of poverty analysis and anti-poverty policy formulation.

The plan of the paper is as follows. Section 2 discusses the concepts of spatial autocorrelation and the Oaxaca decomposition method. Section 3 formulates the regression models and derives the formulae for Oaxaca decomposition under the spatial framework. Section 4 describes the data and presents the empirical results. Finally, section 5 concludes.

2. Spatial Autocorrelation and the Oaxaca Decomposition Method

There may be situations in which the values observed in one location are dependent on the values observed in adjacent locations. This type of autocorrelation, i.e., the correlation pattern observed among observations ordered in geo-space is known as spatial autocorrelation. Following Anselin, 1999; spatial autocorrelation or the coincidence of value similarity with location similarity can be formally expressed by the moment condition: $cov[y_i, y_j] = E[y_i, y_j] - E[y_i]E[y_j] \neq 0$, for i≠j, where i, j refer to individual observations (locations) and $y_{i(j)}$ is the value of a random variable of interest at the respective locations. This covariance pattern becomes meaningful from the perspective of a spatial analysis when the particular configuration of nonzero $(i, j\ )$ pairs has an interpretation in terms of spatial structure, spatial interaction or the spatial arrangement of observations. Spatial dependence may be captured either through model specification error or through dependence of outcome variables on its spatially lagged value. *Spatial* *lag* is indicative of a possible diffusion process in the sense that events occurring in one place predict an increased likelihood of occurrence of similar events in adjacent places.

The idea of Oaxaca decomposition (Oaxaca, 1973) methodology is to explain the gap in the mean values of an outcome variable (*here, poverty*) between two groups (e.g., *between regions*). The gap is decomposed into a part that is due to group differences in the magnitudes of the determinants of the outcome in question, on the one hand, and group differences in the effects of these determinants, on the other. The technique is easy to apply and only requires coefficient estimates from linear regressions for the outcome of interest and sample means of the independent variables used in the regressions[3]. For comparison of poverty between two regions and for proper distribution of aid between the regions, it is necessary to extract the contribution of specific factors (poverty correlates) as well as the contribution of all the factors contributing to the overall characteristics effect. The part of the difference of poverty incidences between two regions unexplained by the characteristics

[3] See Fairlie, 2005.

effect is the coefficients effect, which will reveal the discriminatory effect of the characteristics on the dependent variable over the two regions.

3. Poverty Incidence Analysis in the Spatial Framework

In standard regression model spatial dependence can be incorporated in two distinct ways: as an independent regressor in the form of spatially lagged dependent variable or in the error structure (see Anselin et al. 1998).The former is referred to as *spatial lagged* model and is appropriate when the focus of interest is the assessment of the existence and strength of spatial interaction. The presence of the spatial lag is similar to the inclusion of endogenous variables on the RHS in systems of simultaneous equations. This model is, therefore, often referred to as the simultaneous spatial autoregressive model. In the context of poverty analysis the intuitive assumption that *consumption of the i^{th} household is influenced by the consumptions of the neighbouring similar households* leads to a spatial lagged dependent model.[4]

[4] The other type of model is known as spatial error model:

$$Y = X\beta + \varepsilon \quad (1)$$

$$\varepsilon = \lambda W\varepsilon + \xi \quad (2)$$

where Y is the $(n \times 1)$ vector of observations on the dependent variable, X is the $(n \times k)$ matrix of observations on the explanatory variables, β is the $(k \times 1)$ vector of regression coefficients, ε is the $(n \times 1)$ vector of error terms, $W\varepsilon$ is a spatial lag for the errors, λ is the autoregressive coefficient and ξ is a "well-behaved" error, with mean 0 and variance matrix $\sigma^2 I$.The consequences of ignoring spatial error

Assuming a spatial lagged dependence structure[5], our model is

$$\left(\frac{y}{z}\right)_i^* = \rho \sum_{j=1}^{n} w_{ij} \left(\frac{y}{z}\right)_j^* + X_i\beta + \varepsilon_i \,; \forall\, i = 1, 2, \ldots\ldots, n \qquad (1)$$

where $\left(\frac{y}{z}\right)_i^* = ln\left(\frac{y}{z}\right)_i$; y denoting the monthly per capita household consumption, z denoting the poverty line. The subscript i stands for the i^{th} household. The degree of spatial association (*among the* $\left(\frac{y}{z}\right)^*$ *values for the households*) is captured by the autoregressive coefficient, ρ and $\left(w_{ij}\right)$ is the weight vector for the i^{th} household. Assuming the relationship as in (1) to hold for each household, we get from (1)

$$\left(\frac{Y}{Z}\right)^* = \rho W \left(\frac{Y}{Z}\right)^* + X\beta + \varepsilon \qquad (2)$$

Here $\left(\frac{Y}{Z}\right)^*$ is the $(n \times 1)$ vector of spatially dependent variable, W $(\mathrm{n} \times n)$ is an exogenously determined symmetric weight matrix with positive elements. Corresponding to each observation, each row denotes its neighbourhood status with other locations. For non-neighbours, w_{ij}=0, while for neighbours the

---

dependence are not quite as severe as those of ignoring spatial lag dependence (see Patton et al. 2005). The main problem is that the OLS estimates become inefficient, but they are still unbiased (see Anselin, 1999).

[5] See Anselin et al., 1998, Lesage et al., 2008.

weights are either w_{ij}=1 (binary weights), or a function of something else, such as w_{ij}=1/d_{ij}; d_{ij} being the distance between observation i and observation j.

$W\left(\frac{Y}{Z}\right)^*$ is the *spatial lag* term.

From (2) we get

$$\left(\frac{Y}{Z}\right)^*=(I-\rho W)^{-1}X\beta+(I-\rho W)^{-1}\varepsilon\ . \qquad (3)$$

Each inverse can be expanded into an infinite series, including both the explanatory variables and the error terms at all locations[6]. Consequently, the spatial lag term in (2), should be treated as an endogenous variable and proper estimation method must account for this endogeneity[7]. Since OLS will be biased and inconsistent for this type of model due to the simultaneity bias, either a maximum likelihood or an instrumental variables estimator is needed for the estimation of a spatial lag model.

[6] $\left(\frac{Y}{Z}\right)^*=(I+\rho W+\rho^2W^2+\rho^3W^3+\cdots)\,X\beta+(I+\rho W+\rho^2W^2+\rho^3W^3+\cdots)\,\varepsilon$.

[7] See Anselin, 1999.

Following Anselin, 1992; the maximum likelihood estimation of the spatial lag model is based on the assumption that error terms are normally distributed and given this assumption, a likelihood function, that is a nonlinear function of the parameters, can be derived.[8] In general, a solution can be found by applying the technique of nonlinear optimization. It turns out that the estimates for the regression coefficients β and the error variance σ^2 can be expressed as a function of the autoregressive coefficient ρ. Now, substituting these expressions into the likelihood function, one obtains the concentrated likelihood function, containing only a single parameter, the autoregressive coefficient ρ.[9] By a simple search over values of ρ, the ML estimate is found. The other parameters can then be found from a least squares regression of $\left\{\left(\frac{Y}{Z}\right)^* - \rho \mathrm{W}\left(\frac{Y}{Z}\right)^*\right\}$ on X. A bisection search over values of ρ in the interval $\left(\frac{1}{\omega_{min}}, \frac{1}{\omega_{max}}\right)$ is implemented, where ω_{min} $and\, \omega_{max}$ are,

[8] The likelihood function is of the form:
$L = \sum_i \ln(1 - \rho\omega_i) - \frac{N}{2} ln(2\pi) - \frac{N}{2} ln(\sigma^2) - \left(\left(\frac{Y}{Z}\right)^* - \rho W\left(\frac{Y}{Z}\right)^* - X\beta\right)^t \left(\left(\frac{Y}{Z}\right)^* - \rho W\left(\frac{Y}{Z}\right)^* - X\beta\right)/2\sigma^2$;
Where ω_i is the i[th] Eigen value of the weight matrix W. For full details on the derivation and implementation considerations, see Ord ,1975, Anselin ,1980; Anselin ,1988a; Anselin et al. , 1992.

[9] The concentrated likelihood function takes on the form:
$L_C = -{}^{N}\!/_{2}\, ln[(e_0 - \rho e_L)^t (e_0 - \rho e_L)/N] + \sum_i \ln(1 - \rho\omega_i)$
With e_0 and e_L as the residuals in OLS regression of $\left(\frac{Y}{Z}\right)^*$ on X and $W\left(\frac{Y}{Z}\right)^*$ on X respectively. For further details see Anselin ,1988a; Anselin et al. , 1992.

respectively, the smallest and the largest eigen values of the weight matrix (see Anselin, 1992).[10] The acceptable values of ρ that would yield a stable specification for the autoregressive models are to be found in this interval outside which no other value (of ρ) is acceptable.

The poverty incidence, i.e., the probability of being poor for any particular household can be found as the probability of the estimated value of its income divided by the poverty line being less than one. Let p_i be the probability of being poor for the ith household[11].

Then, p_i= prob $\left(\left(\frac{y}{z}\right)_i^* < 0\right)$

$$=\text{prob}\left(\rho \sum_{j=1}^{n} w_{ij} \left(\frac{y}{z}\right)_j^* + X_i\beta + \varepsilon_i < 0\right) \qquad \text{[from(1)]}$$

$$= \text{prob}\left(\varepsilon_i < -\,\rho \sum_{j=1}^{n} w_{ij} \left(\frac{y}{z}\right)_j^* - X_i\beta\right)$$

$$=\text{prob}\left(\frac{\varepsilon_i - E(\varepsilon_i)}{\sqrt{var(\varepsilon_i)}} < \frac{-\,\rho \sum_{j=1}^{n} w_{ij} \left(\frac{y}{z}\right)_j^* - X_i\beta - E(\varepsilon_i)}{\sqrt{var(\varepsilon_i)}}\right)$$

---

[10] See Anselin et al., 1992 for technical details on the implementation of this bisection search.

[11] The ith household will be poor if its monthly per-capita income (*we have taken the monthly consumption data in absence of income data*), y is less than the poverty line, z. i.e. , $\left(\frac{y}{z}\right)_i < 1$; i.e., $\left(\frac{y}{z}\right)_i^* < 0$.

$$= \Phi\left(\frac{-\rho\sum_{j=1}^{n} w_{ij}\left(\frac{y}{z}\right)_j^{*} - X_i\beta - E(\varepsilon_i)}{\sqrt{var(\varepsilon_i)}}\right);$$

[Φ is the C.D.F of standard normal distribution][12]

$$=\Phi\left(\frac{-\rho\sum_{j=1}^{n} w_{ij}\left(\frac{y}{z}\right)_j^{*} - X_i\beta}{\sigma}\right); \qquad [\text{Assuming } var(\varepsilon_i)=\sigma^2]$$

$$=\Phi\left(\rho^{*}\sum_{j=1}^{n} w_{ij}\left(\frac{y}{z}\right)_j^{*} + X_i\beta^{*}\right); \qquad \left[\rho^{*} = -\frac{\rho}{\sigma} \text{ and } \beta^{*} = -\frac{\beta}{\sigma}\right] \quad (4)$$

The head count ratio, HCR for any region is asymptotically equal to the sample average of the household level poverty incidences[13], i.e., HCR for a region A will be,

$$H_A=\frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(\widehat{\rho_A^{*}}\sum_{j=1}^{n^A} w_{ij_A}\left(\frac{y}{z}\right)_{j_A}^{*} + X_{i_A}\widehat{\beta_A^{*}}\right) \quad ; [\text{using } (4)] \qquad (5)$$

n^A *is* the number of households residing in region A.

---

[12] See Bhaumik et al., 2006.

[13] See Bhaumik et al., 2006.

HCR for region B will be,

$$H_B = \frac{1}{n^B} \sum_{i=1}^{n^B} \Phi \left(\widehat{\rho_B}^* \sum_{j=1}^{n^B} w_{ij_B} \left(\frac{y}{z}\right)_{j_B}^* + X_{i_B} \widehat{\beta_B}^* \right) \tag{6}$$

n^B is the number of households in region B.

Now the difference of $HCR's$ between the regions A and B is given by:

$H_A - H_B$

$$= \left[\frac{1}{n_A} \sum_{i=1}^{n^A} \Phi \left(\widehat{\rho_A}^* \sum_{j=1}^{n^A} w_{ij_A} \left(\frac{y}{z}\right)_{j_A}^* + X_{i_A} \widehat{\beta_A}^* \right) \right]$$

$$- \left[\frac{1}{n_B} \sum_{i=1}^{n^B} \Phi \left(\widehat{\rho_B}^* \sum_{j=1}^{n^B} w_{ij_B} \left(\frac{y}{z}\right)_{j_B}^* + X_{i_B} \widehat{\beta_B}^* \right) \right] \quad \text{; [using (5) and (6)]}$$

$$= \overline{\Phi\left(\widehat{\rho_A}^* W_A \left(\frac{Y}{Z}\right)_A^* + X_A \widehat{\beta_A}^* \right)} - \overline{\Phi\left(\widehat{\rho_B}^* W_B \left(\frac{Y}{Z}\right)_B^* + X_B \widehat{\beta_B}^* \right)}$$

$$= \left\{ \overline{\Phi\left(\widehat{\rho_A}^* W_A \left(\frac{Y}{Z}\right)_A^* + X_A \widehat{\beta_A}^* \right)} - \overline{\Phi\left(\widehat{\rho_A}^* W_B \left(\frac{Y}{Z}\right)_B^* + X_B \widehat{\beta_A}^* \right)} \right\} +$$

$$\left\{ \overline{\Phi\left(\widehat{\rho_A}^* W_B \left(\frac{Y}{Z}\right)_B^* + X_B \widehat{\beta_A}^* \right)} - \overline{\Phi\left(\widehat{\rho_B}^* W_B \left(\frac{Y}{Z}\right)_B^* + X_B \widehat{\beta_B}^* \right)} \right\} \tag{7}$$

The first term in bracket is the *characteristics effect* showing the part of the difference of head count ratios between the regions attributable to the difference in characteristics of the explanatory variables between the regions and the other part is the *coefficients effect* showing the differential impact of the characteristics over the two regions.

Following Yun, 2003; we get:

Characteristics effect,

$$C = \overline{\Phi\left(\widehat{\rho_A}^{*} W_A \left(\frac{Y}{Z}\right)_A^{*} + X_A \widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_A}^{*}\right)} \quad (7.1)$$

$$= \left[\Phi\left(\widehat{\rho_A}^{*} \overline{W_A \left(\frac{Y}{Z}\right)_A^{*}} + \overline{X_A}\widehat{\beta_A}^{*}\right) + R_{M_1}\right] - \left[\Phi\left(\widehat{\rho_A}^{*} \overline{W_B \left(\frac{Y}{Z}\right)_B^{*}} + \overline{X_B}\widehat{\beta_A}^{*}\right) + R_{M_2}\right]$$

(R_{M_i}'s are approximation residuals resulting from evaluating the function $\overline{\Phi(\cdot)}$'s at the mean values)

$$= \left\{\left(\overline{W_A \left(\frac{Y}{Z}\right)_A^{*}} - \overline{W_B \left(\frac{Y}{Z}\right)_B^{*}}\right)\widehat{\rho_A}^{*}\right\} \times \Phi^1\left(\widehat{\rho_A}^{*} \overline{W_A \left(\frac{Y}{Z}\right)_A^{*}} + \overline{X_A}\widehat{\beta_A}^{*}\right) +$$

$$\{(\overline{X_A} - \overline{X_B})\, \widehat{\beta_A}^{*}\} \times \Phi^1\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}} + \overline{X_A}\widehat{\beta_A}^{*}\right)$$

$$+ \left(R_{T_1} + R_{M_1} - R_{M_2}\right) \tag{8}$$

[R_{T_i} is the approximation residual resulting from evaluating the difference of the function $\Phi(\cdot)$'s by using the first order Taylor expansion , Φ^1 being the first derivative of the function Φ.]

Thus from (8),

Aggregate characteristics effect=characteristic effect due to the spatial lagged dependent term + characteristics effect due to the explanatory variables.[14]

Again from (7) we get,

Coefficients effect,

$$D=\overline{\Phi\left(\widehat{\rho_A}^{*}W_B\left(\frac{Y}{Z}\right)_B^{*} + X_B\widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_B}^{*}W_B\left(\frac{Y}{Z}\right)_B^{*} + X_B\widehat{\beta_B}^{*}\right)} \tag{7.2}$$

[14] Since $R_{M_i}'s$ are of the same order, it can be assumed that $\left(R_{M_1} - R_{M_2}\right) \to 0$.
R_{T_1} contains higher order terms in Taylor series expansion involving powers of values less than one and hence can be assumed to be tending to zero.

$$=\left\{\Phi\left(\widehat{\rho_A}^*\overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_A}^*\right) + R_{M_1}{}'\right\} - \left\{\Phi\left(\widehat{\rho_B}^*\overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right) + R_{M_2}{}'\right\}$$

($R_{M_i}{}'$'s are approximation residuals resulting from evaluating the function $\Phi(\cdot)$'s at the mean values)

$$=\overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho_A}^* - \widehat{\rho_B}^*\right) \times \Phi^1\left(\widehat{\rho_B}^*\overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)$$

$$+ \overline{X_B}\left(\widehat{\beta_A}^* - \widehat{\beta_B}^*\right) \times \Phi^1\left(\widehat{\rho_B}^*\overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)$$

$$+ \left(R_{T_1}{}' + R_{M_1}{}' - R_{M_2}{}'\right) \qquad (9)$$

[R_{T_i} is the approximation residual resulting from evaluating difference of the function $\Phi(\cdot)$'s by using the first order Taylor expansion , Φ^1 being the first derivative of the function Φ]

Thus from (9),

aggregate coefficients effect = coefficient effect due to the spatial autoregressive coefficient term + coefficients effect due to the explanatory variables.[15]

Detailed decomposition analysis

The characteristics effect, C contains the effects of all the explanatory variables. The contribution of specific factors can be factored out from the overall contribution of all the variables in making the characteristics effect.

Following Yun (2003),

the weight [16](the share of the particular variable in the aggregate characteristics effect) of the k^{th} *explanatory variable* as derived from (8) is:

$$V_{\Delta X}^{k} = \frac{\left\{\left(\bar{X}_{A}^{k}-\bar{X}_{B}^{k}\right)\beta^{*k}_{A}\right\}\times\Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right)}{\left\{\left(\bar{X}_{A}-\bar{X}_{B}\right)\beta^{*}_{A}+\left(\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}-\overline{W_B\left(\frac{Y}{Z}\right)_{B}^{*}}\right)\widehat{\rho_A}^{*}\right\}\times\Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right)}$$

---

15 $\left(R_{T_1}{}' + R_{M_1}{}' - R_{M_2}{}'\right) \to 0$ (as explained in footnote :8)

16 See Bhaumik et al., 2006.

(10)

$$= \frac{\left(\bar{X}^k_{A-}\bar{X}^k_B\right)\beta^{*k}_{A}}{\left(\bar{X}_{A-}\bar{X}_B\right)\beta^*{}_A + \left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)_B^*}\right)\widehat{\rho_A}^*} \qquad (10.1)$$

The weight for the spatial factor, i.e., the weight due to the spatial lag term, $W\left(\frac{Y}{Z}\right)^*$ is:

$$V^{Spatial}_{Ch} = \frac{\left\{\left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)_B^*}\right)\widehat{\rho_A^*}\right\}\times\Phi^1\left(\widehat{\rho_A}^*\overline{W_A\left(\frac{Y}{Z}\right)^*_A} + \overline{X_A}\widehat{\beta_A}^*\right)}{\left\{\left(\bar{X}_{A-}\bar{X}_B\right)\beta^*{}_A + \left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)_B^*}\right)\widehat{\rho_A^*}\right\}\times\Phi^1\left(\widehat{\rho_A}^*\overline{W_A\left(\frac{Y}{Z}\right)^*_A} + \overline{X_A}\widehat{\beta_A}^*\right)} \qquad (11)$$

$$= \frac{\left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)_B^*}\right)\widehat{\rho_A}^*}{\left(\bar{X}_{A-}\bar{X}_B\right)\beta^*{}_A + \left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)_B^*}\right)\widehat{\rho_A}^*} \qquad (11.1)$$

Since the weights considered here are shares of respective variables in the aggregate characteristics effect, it follows that:

Hence,

$$\sum_{k=1}^{K} V^k_{\triangle X} + V^{Spatial}_{Ch} = 1 \qquad (12)$$

The contribution of a *specific coefficient* in the overall coefficients effect, D, capturing the differential impact of the given level of characteristics can also be factored out.

The weight (the share of the particular variable in the aggregate coefficients effect) of the k^{th} *explanatory variable* as derived from (9) is:

$$V_{\Delta\beta}^{k} = \frac{\left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)}{\left\{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)} \quad (13)$$

$$= \frac{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)} \quad (13.1)$$

The weight for the spatial factor, i.e., the weight due to the spatial autoregressive coefficient, ρ is:

$$V_{Coeff}^{Spatial} = \frac{\left\{\overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)}{\left\{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)} \quad (14)$$

$$= \frac{\overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)}{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)} \quad (14.1)$$

Again,

$$\sum_{k=1}^{K} V_{\Delta\beta}^{k} + V_{Coeff}^{Spatial} = 1 \tag{15}$$

Hence, the difference of poverty incidences between two regions can be written in an alternative fashion (i.e. in terms of contributions of each explanatory variable) using the relationship in (10.1, 11.1) and (13.1, 14.1) as:

H_A- H_B

$$= \left\{\sum_{k=1}^{K} V_{\Delta X}^{k} \left[\overline{\Phi\left(\widehat{\rho_A}^{*} W_A \left(\frac{Y}{Z}\right)_A^{*} + X_A \widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_A}^{*}\right)}\right] + \right.$$

$$\left. \sum_{k=1}^{K} V_{\Delta\beta}^{k} \left[\overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_B}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_B}^{*}\right)}\right]\right\}$$

$$+\left\{V_{Ch}^{Spatial} \left[\overline{\Phi\left(\widehat{\rho_A}^{*} W_A \left(\frac{Y}{Z}\right)_A^{*} + X_A \widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_A}^{*}\right)}\right]\right.$$

$$\left. + V_{Coeff}^{Spatial} \left[\overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_B}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B \widehat{\beta_B}^{*}\right)}\right]\right\} \tag{16}$$

Thus,

difference in the incidences of poverty between the two regions= [sum of the shares of characteristic effect of each explanatory variable +sum of the shares of coefficient effect of each explanatory variable]+[share of characteristic effect due to the spatially lagged dependent variable + share of coefficient effect due to the spatial autoregressive coefficient]

The last part in bracket constitutes the spatial effect. In case there is no spatial effect, the last term in bracket vanishes and we get the conventional decomposition methodology.

Now, from (7.1) and (7.2), we observe that both Characteristics effect, C and Coefficients effect, D are functions of the transformed autoregressive coefficient, ρ^* and the transformed coefficient vector, β^*.Thus for finding the statistical precision of C and D, we need to find out the variance-covariance structure associated with these transformed coefficients.

<u>The variance-covariance structure of (ρ^*, β^*)</u>

From the maximum likelihood estimation of the model as given in (2), we get the asymptotic variance-covariance structure of (ρ, β, σ^2) and from this, variance-covariance structure of (ρ^*, β^*) is obtained through the delta method.[17]

Now,

Asymptotic Variance (ρ, β, σ^2) is given by[18]

$$\begin{bmatrix} tr[W_A]^2 + tr\left[{W_A}^T W_A\right] + \frac{[W_A X\beta]^T [W_A X\beta]}{\sigma^2} & \frac{(X^T W_A X\beta)^T}{\sigma^2} & \frac{tr[W_A]}{\sigma^2} \\ \frac{X^T W_A X\beta}{\sigma^2} & \frac{X^T X}{\sigma^2} & 0 \\ \frac{tr[W_A]}{\sigma^2} & 0 & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \quad (17)$$

$$= \begin{bmatrix} A_{1\times1} & B_{1\times K} & C_{1\times1} \\ D_{K\times1} & E_{K\times K} & F_{K\times1} \\ G_{1\times1} & H_{1\times K} & I_{1\times1} \end{bmatrix}, \text{ (say)} \quad (18)$$

Here $W_A = W(\mathrm{I} - \rho W)^{-1}$

Now, $\rho^* = -\frac{\rho}{\sigma}$, $\beta^* = -\frac{\beta}{\sigma}$

Hence,

---

[17] See Powell, 2007; Seber, 1982.

[18] See Anselin et al., 1998.

Asymptotic Variance (ρ^*, β^*)

$$=\begin{bmatrix} \frac{\delta}{\delta\rho}\left(-\frac{\rho}{\sigma} \quad -\frac{\beta}{\sigma}\right) \\ \frac{\delta}{\delta\beta}\left(-\frac{\rho}{\sigma} \quad -\frac{\beta}{\sigma}\right) \end{bmatrix} \text{Asymptotic Variance } (\rho, \beta) \begin{bmatrix} \frac{\delta}{\delta\rho}\left(-\frac{\rho}{\sigma} \quad -\frac{\beta}{\sigma}\right) \\ \frac{\delta}{\delta\beta}\left(-\frac{\rho}{\sigma} \quad -\frac{\beta}{\sigma}\right) \end{bmatrix}^T$$

$$=\begin{bmatrix} -\frac{1}{\sigma} & \tilde{0}'_{1\times K} \\ \tilde{0}_{K\times 1} & -\frac{\tilde{1}}{\sigma}{}_{K\times K} \end{bmatrix} \begin{bmatrix} A_{1\times 1} & B_{1\times K} \\ D_{K\times 1} & E_{K\times K} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sigma} & \tilde{0}'_{1\times K} \\ \tilde{0}_{K\times 1} & -\frac{\tilde{1}}{\sigma}{}_{K\times K} \end{bmatrix}^T$$

$$= \begin{bmatrix} A^*{}_{1\times 1} & B^*{}_{1\times K} \\ D^*{}_{K\times 1} & E^*{}_{K\times K} \end{bmatrix}, \text{(say)} \qquad (19)$$

<u>The variance of characteristics and coefficients effects</u>

From the estimated variance –covariance structure of the coefficients of spatial regression model in (2), the variances of characteristics and coefficients effect are found using the delta method.

The characteristics effect is given by

$$C=\overline{\Phi\left(\widehat{\rho_A}^{*} W_A \left(\frac{Y}{Z}\right)_A^{*} + X_A\widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B\widehat{\beta_A}^{*}\right)} \quad \text{, [from (7.1)]}$$

$$=f\left(\widehat{\rho_A}^{*}, \widehat{\beta_A}^{*}\right)$$

Thus,

The asymptotic variance of characteristics effect, σ_C^2

$$= \left(\frac{\partial C}{\partial \widehat{\rho_A}^{*}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^{*}}\right) \text{Asymptotic Variance } (\rho_A{}^{*}, \beta_A{}^{*}) \left(\frac{\partial C}{\partial \widehat{\rho_A}^{*}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^{*}}\right)^T$$

$$or\ \sigma_C^2=\left(\frac{\partial C}{\partial \widehat{\rho_A}^{*}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^{*}}\right) \begin{bmatrix} A^{*}{}_{1\times 1} & B^{*}{}_{1\times K} \\ D^{*}{}_{K\times 1} & E^{*}{}_{K\times K} \end{bmatrix}^A \left(\frac{\partial C}{\partial \widehat{\rho_A}^{*}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^{*}}\right)^T, \qquad (20)$$

Here,

$\begin{bmatrix} A^{*}{}_{1\times 1} & B^{*}{}_{1\times K} \\ D^{*}{}_{K\times 1} & E^{*}{}_{K\times K} \end{bmatrix}^A$ =Asymptotic Variance $(\rho_A{}^{*}, \beta_A{}^{*})$; as derived from relationship (19).[19]

Again, the coefficients effect is given by

$$D=\overline{\Phi\left(\widehat{\rho_A}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B\widehat{\beta_A}^{*}\right)} - \overline{\Phi\left(\widehat{\rho_B}^{*} W_B \left(\frac{Y}{Z}\right)_B^{*} + X_B\widehat{\beta_B}^{*}\right)} \quad \text{[from (7.2)]}$$

[19] See Appendix- I for the exact form of σ_C^2.

$$=f\left(\widehat{\rho_A}^*, \widehat{\beta_A}^*; \widehat{\rho_B}^*, \widehat{\beta_B}^*\right);$$

Thus,

The asymptotic variance of coefficients effect, σ_D^2

$$= \left(\frac{\partial D}{\partial \widehat{\rho_A}^*} \quad \frac{\partial D}{\partial \widehat{\beta_A}^*}\right) \text{Asymptotic Variance } ({\rho_A}^*, {\beta_A}^*) \left(\frac{\partial D}{\partial \widehat{\rho_A}^*} \quad \frac{\partial D}{\partial \widehat{\beta_A}^*}\right)^T$$

$$+\left(\frac{\partial D}{\partial \widehat{\rho_B}^*} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right) \text{Asymptotic Variance } ({\rho_B}^*, {\beta_B}^*) \left(\frac{\partial D}{\partial \widehat{\rho_B}^*} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right)^T$$

$$=\left(\frac{\partial D}{\partial \widehat{\rho_A}^*} \quad \frac{\partial D}{\partial \widehat{\beta_A}^*}\right) \begin{bmatrix} {A^*}_{1\times 1} & {B^*}_{1\times K} \\ {D^*}_{K\times 1} & {E^*}_{K\times K} \end{bmatrix}^A \left(\frac{\partial D}{\partial \widehat{\rho_A}^*} \quad \frac{\partial D}{\partial \widehat{\beta_A}^*}\right)^T$$

$$+\left(\frac{\partial D}{\partial \widehat{\rho_B}^*} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right) \begin{bmatrix} {A^*}_{1\times 1} & {B^*}_{1\times K} \\ {D^*}_{K\times 1} & {E^*}_{K\times K} \end{bmatrix}^B \left(\frac{\partial D}{\partial \widehat{\rho_B}^*} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right)^T \qquad (21)$$

Here, $\begin{bmatrix} {A^*}_{1\times 1} & {B^*}_{1\times K} \\ {D^*}_{K\times 1} & {E^*}_{K\times K} \end{bmatrix}^B$ =Asymptotic Variance $({\rho_B}^*, {\beta_B}^*)$; as derived from relationship (19).[20]

Testing the significance of C and D

[20] See Appendix-II for the exact form of σ_D^2.

The characteristics effect and the coefficients effect are linear combinations of standard normal variables. So C and D also will follow normal distributions. The test statistic under the null hypothesis (C=0, D=0) are respectively given by [21]

$$t_c = \frac{C}{\sigma_C} \; and \;\; t_D = \frac{D}{\sigma_D} \qquad (22)$$

4. Data and Results

The data used in the analysis are the household level NSS 61st round (2004-05) employment-unemployment data for the rural sector. The logarithm of ratio of per-capita consumption expenditure to the poverty line, R $\left(= ln\left(\frac{y}{z}\right)\right)$ is our variable of study. Explanatory variables are broadly categorized as:

A. Demographic characteristics of the households

B. Educational status

C. Wealth status

D. Labour market characteristics

E. Government aid

[21] The significance test for the characteristic and coefficient effects at the individual variable level can be implemented in a similar fashion. This, however, has not been shown in this paper.

The variables under these broad categories are:

A. Demographic characteristics of the households:

1. 1-dependency ratio, (1-DEPRAT); where

   Dependency ratio

$$= \frac{\text{total number of children and old persons in the household}}{\text{household size}}$$

2. Dummy variable, D_FEMH , indicating whether the family is female headed or not;

   D_FEMH =1 if the family is female-headed

   =0 otherwise

B. Educational status of the households

1. The proportion of members having secondary education, PSECED.
2. The proportion of members having tertiary education, PTERTED.
3. The average general educational level, GENED.[22]

C. Wealth status of the households

[22] Educational levels considered are: not literate, literate without formal schooling, literate but below primary, primary, middle, secondary, higher secondary, diploma/certificate course, graduate, post graduate and above. The average educational level of each household is obtained as the average over codes assigned to different educational levels (in increasing order), starting from zero for the illiterate to the maximum for the category : post graduate and above.

1. Amount of land possessed (measured in Hectares), LAND.

D. <u>Labour Market Characteristics of the households</u>

1. The proportion of members engaged in own account work, POWNAC.
2. The proportion of members not attending school for supporting domestic income, PNSCH.
3. Proportion of members engaged in domestic and other duties , PDOMO.
4. Proportion of members engaged in domestic duties only, PDOM.
5. Proportion of members employed, PEMP.

E. <u>Government aid</u>

Dummy variable,

D_GOVAID = 1 if at least one member of the household is receiving social security benefit or is a beneficiary.

=0 otherwise

The most important step in the spatial regression analysis is the construction of a spatial weight matrix. The construction of the spatial weight matrix has been made on the basis of contiguity-based criterion. According to the adjacency criterion, spatial weight between two adjacent units will be one, while that between two non-adjacent units will be zero.

For incorporating the household specific spatial effects based on adjacency criterion[23] the following rule has been adopted.

1. For households living in the same district, spatial weight is equal to unity if the households are in the same second stage stratum[24], spatial weight being equal to zero, if otherwise.
2. For households living in different districts, spatial weight is equal to unity if both the following criteria are met:
   1) Districts are contiguous.
   2) Households are in the same second stage stratum.

Spatial weight is equal to zero if the above criteria are not satisfied.

The names of the districts and their adjacent districts for both North Bengal (Region A) and South Bengal (Region B) are reported in Tables 1 and 2.[25]

---

[23] Alternative weight functions can be defined subject to the availability of data. Since the household identification code does not provide information at sub-district level, this definition of weight has been adopted. It may be pointed out that recently Beck et al., 2006 constructed a spatial weight matrix based on *non-geographic notion of space*. They argued in favour of considering political economy notions of distance, such as relative trade or common dyad membership in situations of spatial analysis involving trade and democracy.

[24] Rural households are classified into three second stage stratum which are formed as:

SSS 1:	relatively affluent households
SSS 2:	from the remaining households, households having principal earning from non- agricultural activity
SSS 3:	other households

[25] Also see the map of West Bengal (drawn using Autocad) given in Figure 1.

The impact of these explanatory variables on consumption is studied separately for North and South Bengal. North Bengal consists of districts with codes 1-7 and South Bengal consists of districts with codes 8-18. The estimated coefficients and associated standard errors are reported in Table 3. Almost all the coefficients turn out to be highly significant. The poverty estimates for the two regions, thus found using the estimates from the regression analysis following Bhaumik et al., 2006 are reported in Table 4. The overall characteristics and coefficients effects of the difference in poverty estimates of the two regions as well as the contribution of each of the explanatory variable are reported in Table 5.[26]

Tests for detecting the presence of spatial effect are performed on the dependent variable, R (i.e., logarithm of the ratio of per-capita consumption to poverty line). Two measures of spatial autocorrelation have been analyzed in this paper: Moran's I (Moran, 1948) and Geary's C (Geary, 1954). Both measures suggest the presence of strong spatial autocorrelation in the data.[27] Table 6 and Table 7 summarize the results of these tests.

[26] See Bhaumik et al., 2006 for the conventional decomposition methodology.

[27] $\mathrm{I} > E(\mathrm{I}), C < 1$, which is indicative of significant spatial correlation , for both North and South Bengal.

Next the whole analysis is repeated in the framework of spatial regression.

The result of the spatial analysis as implemented using a spatial lag model is reported in Tables 8-12.

The important findings from the analysis are:

1. There is marked difference in the incidences of poverty between the two regions, viz., North and South Bengal in both spatial and non-spatial framework, the difference being more or less the same (approximately equal to 0.14) in both the cases (Table 4 and Table 9).
2. As found from Tables 6 and 7, within each region the two measures of spatial autocorrelation suggest that significant spatial effect is present in the incidence of poverty.
3. As found from Tables 10 and 11, significant lag coefficient (ρ) is found for both North and South Bengal. The coefficient (0.5945) is much higher in case of South Bengal as compared to that in North Bengal (0.4215) indicating that spatial effect is more pronounced in South Bengal.
4. In the ordinary regression analysis (Table 5), the share of aggregate characteristics effect is 40.9% (*column 3*) and that of aggregate coefficients effect is 59.1% *(column 5)*. In the case of spatial analysis

(Table 12), the share of aggregate characteristics effect is much higher (90.7%, see *column 3*) compared to that of the aggregate coefficients effect (9.3%, see *column 5*).

In fact the huge discrepancy in the share of the aggregate characteristics effect between the two cases may be attributed to the *effect of the spatial lagged dependent term*, $\left(\left(\overline{W_A\left(\frac{Y}{Z}\right)^*_A} - \overline{W_B\left(\frac{Y}{Z}\right)^*_B}\right)\widehat{\rho_A}^* \times \Phi^1(.)\right)$ as an additional factor (in the aggregate characteristics effect, see equation (8)). However *the shares of characteristic-wise effects* and hence the *combined shares of the aggregate characteristics effect due to all the explanatory variables,* viz., effect due to the term $\left((\bar{X}_A - \bar{X}_B)\,\beta^*_A \times \Phi^1(.)\right)$ remain more or less the same. The combined share (of all the explanatory variables) is 40.9% in case of non-spatial analysis (Table 5) and 37.1% in case of spatial analysis (Table 12)[28]. The small difference in the combined share of the aggregate characteristics effect due to the explanatory variables in the two (spatial and non spatial)

[28] Summing the characteristic-wise shares , i.e., summing the shares of each explanatory variable, we get the conventional effect (=37.1%).

cases[29] is due to the change in estimates of associated coefficients in the spatial analysis.

The large difference (49.8 %, See Table 13) between aggregate coefficients effects between the two cases is attributed to considerable change in both the conventional term, viz. *the effect due to the term* $\left(\overline{X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) \times \Phi^1(.)\right)$ and the additional term, viz. *the effect due to the term* $\left(\overline{W_B\left(\frac{Y}{Z}\right)_B}^* (\widehat{\rho_A}^* - \widehat{\rho_B}^*) \times \Phi^1(.)\right)$.[30]

Thus, the incorporation of the spatial effect has substantially changed the magnitude of the characteristics effect and the coefficients effect. While the contributions of the characteristics effect and coefficients effect are almost same in the general regression analysis, the characteristics effect has a much larger share in the spatial analysis. A comparative *summarization* of the results of the decomposition analysis has been given in Table 13.

It is observed from Table 13 that *educational status* has a major role to play in terms of characteristic effect and coefficient effect in both (spatial and non-spatial) cases.

---

[29] The difference of the share of aggregate characteristics effect (=40.9%) in Table 5 (*column 3*) and the share of conventional effect (=37.1%) in Table 12 (*column 3*).

[30] See equation 9.

Now, one can interpret the characteristics effect as a *resource effect*[31] for the explanatory variables which have positive marginal products.

From Table 12, we see that the estimates of the aggregate characteristics effect and aggregate coefficients effect are positive. Now a *positive characteristic effect* by any *specific variable,* x^k means that $\left(\overline{X_A}^k - \overline{X_B}^k\right) \beta_A{}^k < 0$.[32] With x^k having a positive marginal product, i.e., for $\beta_A{}^k > 0$, this implies that region A has a scarcity of the particular resource, x^k as compared to region B, i.e., $\overline{X_A}^k - \overline{X_B}^k < 0$.

On the other hand, the coefficients effect capturing the differential impact of the characteristics over the two regions can be interpreted as an *efficiency effect*[33]. *A positive coefficient effect* by any *specific variable* x^k means

[31] The regression relationship of consumption on a set of economic variables (*the income generating function*) can be thought of in the framework of frontier production functions (see Aigner et al., 1976). The variables with positive coefficients (i.e. with positive marginal products) can be considered as inputs (resources) of production. Thus, the characteristics effect can be thought of as a resource effect.

[32] See Appendix- III for the proof.

[33] The coefficients associated with each explanatory variable (with positive marginal product) signify the responsiveness of the output with respect to the inputs) are directly linked with productivity. The Coefficient effect which is also a function of estimated coefficients, showing the differential impact of the characteristics over the two regions, can thus be thought of as a productivity effect. Now Technical efficiency score of a household is the ratio of actual output to potential output and is thus a function of estimated coefficients. There should be some sort of correlation between average technical efficiency level and the aggregate coefficients effect. Coefficients effect can alternatively looked upon as an efficiency/productivity effect.

that ${\beta_A}^k - {\beta_B}^k < 0$.[34] When both ${\beta_A}^k and\ {\beta_B}^k$ are greater than zero, the above result signifies that as far as the utilization of resource (here x^k) is concerned region A is less efficient than region B.

The figures in Table 12 indicate that the difference in the incidences of poverty is attributable mainly to the characteristics effect, which in turn indicates that there is a deficiency of resources in North Bengal (which has a high poverty level) as compared to South Bengal. A positive coefficients effect indicates that households in the northern part of Bengal are less efficient in terms of utilization of resources, though the difference in the average level of efficiency between North and South Bengal is not quite significant at all.

5. Conclusion

The decomposition of the difference in the incidences of poverty between two regions into characteristics and coefficients effect is a very important issue in the perspective of (*anti-poverty*) policy formulation. Poverty as an economic problem can be formulated from the perspective of availability of resources and the utilization of resources. Characteristics effect and coefficients effect

[34] See Appendix- IV for the proof.

can alternatively be looked upon as a resource constraint and an efficiency constraint. It has been observed from the spatial analysis that the difference in the incidences of poverty in the two parts of Bengal is attributable much more to the characteristics effect than to the coefficients effect, while in the non-spatial case they were more or less equally important. This is an important finding in terms of policy perspective. This is because it clearly indicates that anti poverty schemes should focus on enhancing resource in North Bengal especially in terms of educational characteristics.

Acknowledgements:

The authors thank Prof. Babulal Seal of Burdwan University for some helpful discussions about the Delta method.

References

1. Aigner D J, Lovell C A K, Schmidt P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **6**: 21-37.
2. Anselin L .1980. *Estimation Methods for Spatial Autoregressive Structures*, Regional Science Dissertation and Monograph Series. Ithaca: NY.

3. Anselin, L .1988a. *Spatial Econometrics, Methods and Models* .Dordrecht: Kluwer Academic.

4. Anselin L. 1992. Space stat tutorial: a workbook for using space stat in the analysis of spatial data. University of Illinois: Urbana.

5. Anselin, L, Hudak S .1992. Spatial econometrics in practice: a review of software options. *Regional Science and Urban Economics* **22**: 509-36.

6. Anselin L. 1999. Spatial econometrics. University of Texas at Dallas.

7. Anselin L. 2000. Spatial econometrics. In *A Companion to Theoritical Econometrics,* Baltagi B (ed.). Basil Blackwell, Oxford: 310—330.

8. Anselin L. 2003. Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review* **26** (2): 152-166.

9. Anselin L, Bera A. 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In *Handbook of Applied Economic Statistics,* Ullah A and Giles D E (eds): 237– 289.

10. Beck N , Gleditsch K S. 2006. Space is more than geography: using spatial econometrics in the study of political economy. *International Studies Quarterly* **50**: 27–44.

11. Bhaumik S K, Gang I N, Yun M S. 2006. Note on decomposing differences in poverty incidence using regression estimates: algorithm and example. IZA Discussion Paper 2262.

12. Bigman D, Srinivasan P V. 2002. Geographical targeting of poverty alleviation programs: methodology and applications in rural India. *Journal of Policy Modeling* **24**(3): 237-255.

13. Bokosi F K. 2007. Household poverty dynamics in malawi: a bivariate probit analysis. *Journal of Applied Science* **7**(2): 258-262.

14. Chandrasiri G W J, Samarakoon L. 2008. Spatial patterns and geographic determinants of poverty in sri lanka: linking poverty mapping with geoinformatics. Asian Conference on Remote Sensing, Colombo: Sri Lanka.

15. Davis B. 2003. Choosing a method of poverty mapping. Food and Agriculture Organization of the United Nations.

16. Fairlie R W. 2005. An extension of the blinder-oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement* **30**: 305-316.

17. Gang I N , Sen K , Yun M S. 2008. Caste, ethnicity and poverty in rural India. *Review of Income and Wealth.* **54**(1): 50-70.

18. Geary R. 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician* **5**: 115–145.

19. Geda A , Jong N. Mwabu G , Kimenyi M. 2001. Determinants of poverty in kenya: household-level analysis. KIPPRA Discussion Paper 9. Kenya Institute for Public Policy Research and Analysis: Nairobi.

20. Lesage J P , Charles J S. 2008. Using home buyers' revealed preferences to define the urban–rural fringe . *J Geograph Syst* **10**: 1–21.

21. Moran P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society Series B* **10**: 243–251.

22. Nielsen H S .1998.Discrimination and detailed decomposition in a logit model. *Economics Letters* **61**: 115-120.

23. Oaxaca R . 1973.Male-female wage differentials in urban labor markets. *International Economic Review* **14**: 693-709.

24. Ord K. 1975. Estimation methods for models of spatial interaction, *Journal of the American Statistical Association* **70**: 120-26.

25. Patton M, Mcerlean S. 2003. Spatial Effects within the Agricultural Land Market in Northen Ireland. *Journal of Agricultural Economics* **54**(1): 35-54.

26. Petrucci A, Salvati N, Seghieri C. 2003. The application of a spatial regression model to the analysis and mapping of poverty. *Environment and Natural Series* **7**. Fao: Rome.

27. Pisati M. 2001. Tools for spatial data analysis. *Stata Technical Bulletin* Reprints **10**: 277–281.

28. Powell L A. 2007. Approximating variance of demographic parameters using the delta method: a reference for avian biologists. *The Condor* **109**: 949–954.

29. Seber G A F. 1982. *The estimation of animal abundance and related Parameters*, 2nd ed. Chapman: London and Macmillan: New York.

30. Voss P R, Long D D, Hammer R B, Friedman S. 2006. County child poverty rates in the U.S.: a spatial regression approach. *Population Research and Policy Review* **25**: 369-391.

31. Yun M S.2004. Decomposing differences in the first moment. *Economic Letters*. **82**(2): 275-280.

32. World Bank. 2003. Poverty reduction strategy sourcebook. (http://www.worldbank.org/poverty/strategies/sourcons.htm.)

Table 1: Districts of North Bengal (Region A)

District name	District code no.	Adjacent district code no.
Darjiling	1	2,4
Jalpaiguri	2	3,1
Kochbihar	3	2
Uttar dinajpur	4	5,6,1
Dakshin dinajpur	5	4,6
Maldah	6	4,5,7
Murshidabad	7	6

Table 2: Districts of South Bengal (Region B)[35]

District name	District code no.	Adjacent district code no.
Birbhum	8	9
Bardhaman	9	8,10,12,13,14
Nadia	10	9,11,12
North 24 Paraganas	11	10,18,12,16
Hugli	12	9,10,11,13,15,16
Bankura	13	9,12,14,15
Purulia	14	9,13,15
Medinipur	15	13,12,16,18,14
Howrah	16	12,15,18,11
South 24 Paraganas	18	16,11,15

[35] As there is no rural sector in the district, Kolkata, it has not been shown in Table 2.

Table 3: Factors Influencing Per-Capita Household Consumption: Analysis in the General Regression Framework (ML Estimation) (Dependent Variable : $ln\left(\frac{y}{z}\right)$)

		North Bengal (A)		South Bengal (B)	
(1)	(2)	(3)	(4)	(5)	(6)
Characteristics	Variables	Estimate (β_A)	t value	Estimate (β_B)	t value
Demographic characteristics of the households	1-DEPRAT	0.3206	7.2983*[36]	0.1935	8.6588*
	D_FEMH	0.0989	3.7174*	0.0970	4.5857*
Educational status of the household	PSECEDU	0.1162	1.4868	0.2235	2.2211*
	PTERTEDU	0.3946	5.2025*	0.3018	7.4137*
	GENEDU	0.0467	9.7449*	0.0646	12.3854*
Wealth status	LAND	0.0005	8.3675*	0.0004	10.8966*
Labour market characteristics	POWNAC	0.4322	7.2798*	0.3625	8.6374*
	PNSCH[37]	-0.3425	-4.8408*	-0.2628	-6.9602*
	PDOMO	0.2402	4.0309*	0.2729	5.0168*
	PDOM	0.4599	6.3807*	0.3285	8.3455*
	PEMP	0.1223	2.2204*	0.2339	2.8790*
Government aid	D_GOVAID	0.2394	6.7102*	0.2112	9.860*
Constant		-0.3750	-12.9295*	-0.2272	-14.9909*

Table 4: Poverty Incidences in North and South Bengal

(1)	(2)	(3)
	Sample size	Poverty incidence
North Bengal	1526	$H_A = 0.3108$
South Bengal	3407	$H_B = 0.1714$
Difference in poverty incidence: $(H_A\text{-}H_B) = 0.1394$		

[36] * indicates significance at 5% level.
[37] This variable is expected have a negative influence on household monthly per-capita consumption. A possible explanation is that the income earned by joining the labour market is smaller than the gain in income made through the increase in technical efficiency resulting from joining the educational institutions.

Table 5: Decomposing the Difference of Poverty Incidences: (H_A-H_B) between North Bengal (Region A) and South Bengal (Region B) Using Estimates of Table 3.

(1)		Characteristics effect (2) estimates	(3)[38] share (%)		Coefficients effect (4) estimates	(5) share(%)	
Aggregate effect		0.0570 (19.57)[39]	40.9		0.0824 (9.75)	59.1	
Demographic characteristics of the households	1-DEPRAT	0.0106	7.6	6.4	-0.0898	-64.4	-65.1
	D_FEMH	-0.0016	-1.2		-0.0010	-0.7	
Educational status of the household	PSECEDU	0.0015	1.1	27.1	0.0045	3.3	29.2
	PTERTEDU	0.0059	4.2		-0.0071	-5.1	
	GENEDU	0.0304	21.8		0.0432	31.0	
Wealth status	LAND	0.0010	0.7	0.7	-0.0107	-7.7	-7.7
Labour market characteristics	POWNAC	0.0001	0.1	4.5	-0.0153	-11.0	-6.6
	PNSCH	-0.0031	-2.2		0.0063	4.5	
	PDOMO	0.0045	3.2		0.0007	0.5	
	PDOM	0.0049	3.5		-0.0134	-9.6	
	PEMP	-0.0002	-0.1		0.0126	9.0	
Government aid	D_GOVAID	0.0031	2.3	2.3	-0.0040	-2.9	-2.9
Constant		-	-		0.1564	112.2	

[38] Column(3)=Column(2)/ Difference in poverty incidence(Table 4) x 100

Table 6: Tests for Spatial Autocorrelation of the Variable $\boldsymbol{ln\left(\frac{y}{z}\right)}$: Case of North Bengal (A)

(1)	(2)	(3)	(4)	(5)	(6)
Variable $\boldsymbol{ln\left(\frac{y}{z}\right)}$	I	$E(I)$	$sd(I)$	z	p-value (one tailed test)
Moran's I	0.169	-0.001	0.003	55.416	0.000
Geary's C	0.757	1.000	0.020	-12.146	0.000

Table 7: Tests for Spatial Autocorrelation of the Variable $\boldsymbol{ln\left(\frac{y}{z}\right)}$: Case of South Bengal (B)

(1)	(2)	(3)	(4)	(5)	(6)
Variable $\boldsymbol{ln\left(\frac{y}{z}\right)}$	I	$E(I)$	$sd(I)$	z	p-value (one tailed test)
Moran's I	0.075	-0.000	0.001	88.018	0.000
Geary's C	0.878	1.000	0.018	-6.595	0.000

[39] The figures in parentheses denote the t values. Both characteristics effect and coefficients effect are significant at 5% level.

Table 8: Factors Influencing Per-Capita Household Consumption: Analysis in the Spatial Regression Framework (ML Estimation)

		North Bengal (A)		South Bengal (B)	
(1)	(2)	(3)	(4)	(5)	(6)
Characteristics	Variables	Estimate (β_A)	t value	Estimate (β_B)	t value
Demographic characteristics of the households	1-DEPRAT	0.2875	6.7190*[40]	0.1958	5.4779*
	D_FEMH	0.1027	3.9743*	0.0912	4.3771*
Educational status of the household	PSECEDU	0.0778	1.0241	0.2135	4.2267*
	PTERTEDU	0.3440	4.6599*	0.2465	4.7856*
	GENEDU	0.0407	8.6780*	0.0592	16.1849*
Wealth status	LAND	0.0004	7.0275*	0.0003	7.5681*
Labour market characteristics	POWNAC	0.4575	7.9319*	0.4004	8.2792*
	PNSCH[41]	-0.3523	-5.1294*	-0.2741	-5.7701*
	PDOMO	0.2927	5.0374*	0.2747	5.9433*
	PDOM	0.4065	5.7912*	0.2897	5.4385*
	PEMP	0.1900	3.5234*	0.3021	7.3237*
Government aid	D_GOVAID	0.1875	5.3500*	0.1330	5.5521*
Constant		-0.4311	-14.9922*	-0.4600	-16.2479*

[40] * indicates significance at 5% level.

[41] As explained in footnote: 28, the influence of this variable is expected to be negative.

Table 9: Poverty Incidences in North and South Bengal (Spatial Framework)

(1)	(2)	(3)	(4)	(5)
	Sample size	Spatial rho (ρ)	Spatial sigma (σ)	Poverty incidence
North Bengal	1526	0.4215	0.3100	H_A = 0.3118
South Bengal	3407	0.5945	0.3400	H_B = 0.1711
	Difference in poverty incidence: H_A-H_B = 0.1392			

Table 10: Statistical Test for Spatial Rho (ρ): North Bengal

(1)	(2)	(3)	(4)	(5)
Coefficient (ρ)	Standard error	z	P>|z|	95% Confidence Interval
0.4216	0.0440	9.59	0.000	0.3354 - 0.5077
Wald test of rho=0:			chi2(1) = 91.968 (0.000)	
Likelihood ratio test of rho=0:			chi2(1) = 86.221 (0.000)	
Lagrange multiplier test of rho=0:			chi2(1) = 165.823 (0.000)	
Acceptable range for rho:			-15.000 < rho < 1.000	

Table 11: Statistical Test for Spatial Rho (ρ): South Bengal

(1)	(2)	(3)	(4)	(5)
Coefficient (ρ)	Standard error	z	P>|z|	95% Confidence Interval
0.5945	0.0377	15.76	0.000	0.5206 - 0 .6684
Wald test of rho=0:			chi2(1) = 248.452 (0.000)	
Likelihood ratio test of rho=0:			chi2(1) = 234.005 (0.000)	
Lagrange multiplier test of rho=0:			chi2(1) = 602.011 (0.000)	
Acceptable range for rho:			-1.898 < rho < 1.000	

Table 12: Decomposing the Difference of Poverty Incidences: (H_A-H_B) between North Bengal (Region A) and South Bengal (Region B) Using Estimates of Table 8.

(1)			Characteristics effect (2) Estimates	Characteristics effect (3)[42] Share (%)			Coefficients effect (4) Estimates	Coefficients effect (5) Share (%)		
Aggregate effect			**0.1276** (17.65)			**90.7**	**0.0131** (1.19) [43]			**9.3**
Conventional Effect	Demographic characteristics of the households	1-DEPRAT	0.0099	7.03	5.8	**37.1**	0.0389	27.6	28.3	**14.2**
		D_FEMH	-0.0018	-1.27			0.0010	0.7		
	Educational status of the household	PSECEDU	0.0011	0.76	24.2		-0.0036	-2.5	-19.6	
		PTERTEDU	0.0053	3.78			0.0040	2.8		
		GENEDU	0.0277	19.66			-0.0280	-19.9		
	Wealth status	LAND	0.0009	0.61	0.6		0.0049	3.5	3.5	
	Labour market characteristics	POWNAC	0.0001	0.06	4.7		0.0079	5.6	6.1	
		PNSCH	-0.0034	-2.39			-0.0035	-2.5		
		PDOMO	0.0057	4.05			0.0044	3.0		
		PDOM	0.0045	3.22			0.0068	4.8		
		PEMP	-0.0003	-0.21			-0.0068	-4.8		
	Government aid	D_GOVAID	0.0026	1.83	1.8		0.0031	2.2	2.2	
	Constant		-	-	-		-0.6872	-0.009	-6.4	
Spatial Effect			*0.0753*			***53.6***	*-0.0069*			***-4.9***

[42] Column(3)=Column(2)/ Difference in poverty incidence(Table 9) x 100

[43] The figures in parentheses denote the t values. The characteristic effect is significant at 5% level. The coefficients effect is not significant at 5% level.

Table 13: A Comparative Summarization of the Results.

		Ordinary Regression Analysis	**Spatial Regression Analysis**
(1)	(2)	(3)	(4)
Incidence of poverty	North Bengal	0.3108	0.3118
	South Bengal	0.1714	0.1711
Difference of poverty incidences		0.1394	0.1392
Oaxaca Decomposition of the difference of poverty incidences	Share[44] of Aggregate Characteristics effect	40.9	90.7
	Share of Aggregate Coefficients effect	59.1	9.3
Decomposition of Share of Aggregate Characteristics effect			
Demographic characteristics of the households		6.4	5.8
Educational status of the household		27.1	24.2
Wealth status		0.7	0.6
Labour market characteristics		4.5	4.8
Government aid		2.3	1.8
Constant		-	-
Spatial characteristics effect		-	***53.6***
Decomposition of Share of Aggregate Coefficients effect			
Demographic characteristics of the households		-65.1	28.3
Educational status of the household		29.2	-19.6
Wealth status		-7.7	3.5
Labour market characteristics		-6.6	6.1
Government aid		-2.9	2.2
Constant		112.2	-6.4
Spatial coefficients effect		-	***-4.9***

[44] Share is given in percentages.

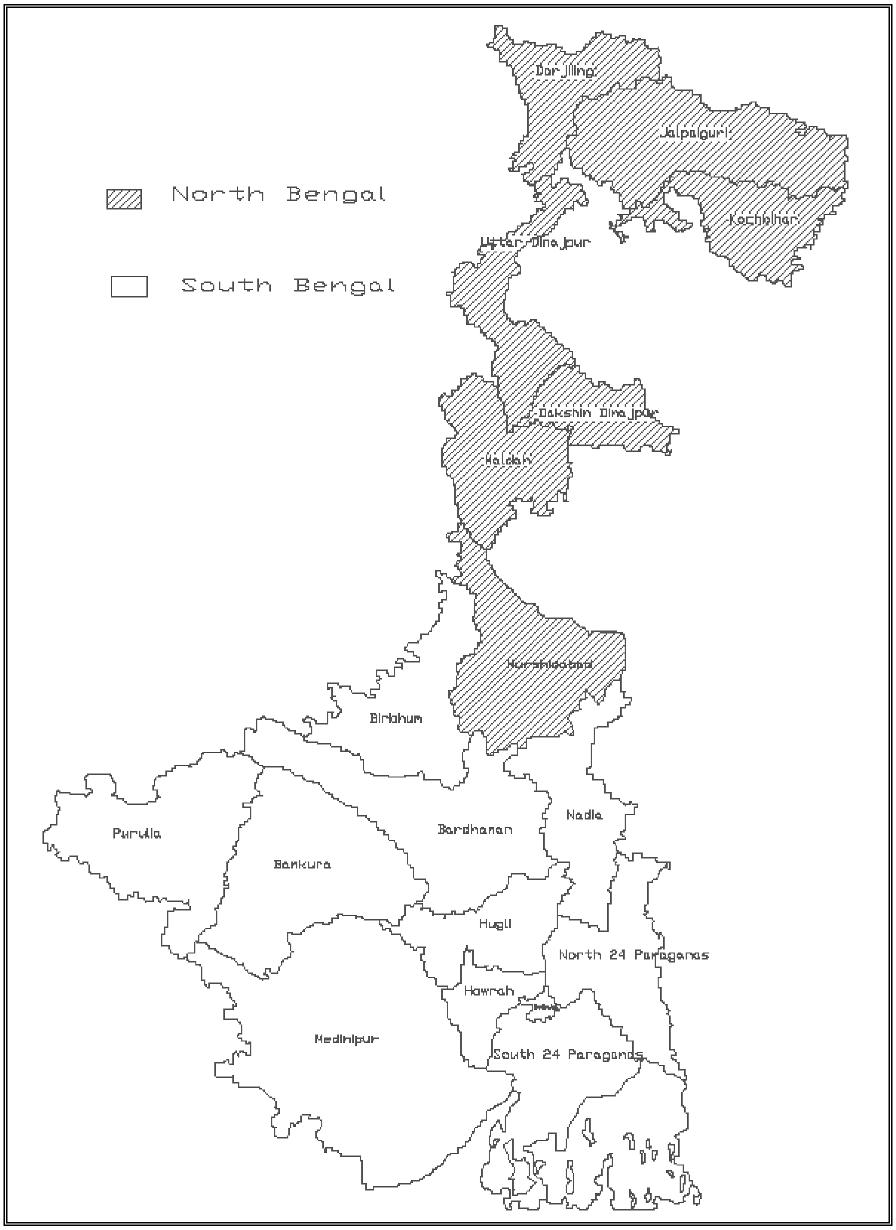

Figure 1: Map of West Bengal

Appendix- I

Writing (7.1) in terms of individual observations, we get:

$$C=\left[\frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(\widehat{\rho_A}^*\sum_{j=1}^{n^A}w_{ij_A}\left(\frac{y}{z}\right)_{j_A}^{*}+X_{i_A}\widehat{\beta_A}^*\right)\right]-$$

$$\left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi(\widehat{\rho_A}^*\sum_{j=1}^{n^B}w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^{*}+X_{i_B}\widehat{\beta_A}^*)\right]$$

$$\Rightarrow \frac{\partial C}{\partial \widehat{\rho_A}^*}$$

$$=\left[\frac{1}{n^A}\sum_{i=1}^{n^A}\sum_{j=1}^{n^A}w_{ij_A}\left(\frac{y}{z}\right)_{j_A}^{*}\phi\left(\widehat{\rho_A}^*\sum_{j=1}^{n^A}w_{ij_A}\left(\frac{y}{z}\right)_{j_A}^{*}+X_{i_A}\widehat{\beta_A}^*\right)\right]-$$

$$\left[\frac{1}{n^B}\sum_{i=1}^{n^B}\sum_{j=1}^{n^B}w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^{*}\phi\left(\widehat{\rho_A}^*\sum_{j=1}^{n^B}w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^{*}+X_{i_B}\widehat{\beta_A}^*\right)\right] \quad (A{:}\ 1)$$

Here ϕ is the PDF of standard normal distribution.

Again,

$$\frac{\partial C}{\partial \widehat{\beta_A}^*}=\left[\left(\frac{1}{n^A}\sum_{i=1}^{n^A}X_{i_A}{}^{k}\phi\left(\widehat{\rho_A}^*\sum_{i,j=1}^{n^A}w_{ij_A}\left(\frac{y}{z}\right)_{j_A}^{*}+X_{i_A}\widehat{\beta_A}^*\right)\right)\right]_{k=1(1)K}-$$

$$\left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B}X_{i_B}{}^{k}\phi\left(\widehat{\rho_A}^*\sum_{i,j=1}^{n^B}w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^{*}+X_{i_B}\widehat{\beta_A}^*\right)\right)\right]_{k=1(1)K} \quad (A{:}\ 2)$$

Appendix- II

Writing (7.2) in terms of individual observations , we get:

$$D = \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(\widehat{\rho_A}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_A}^*\right)\right] -$$

$$\left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(\widehat{\rho_B}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_B}^*\right)\right]$$

$$\frac{\partial D}{\partial \widehat{\rho_A}^*} = \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\sum_{j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* \phi\left(\widehat{\rho_A}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_A}^*\right)\right] \quad (A:3)$$

$$\frac{\partial D}{\partial \widehat{\beta_A}^*} = \left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B} X_{i_B}{}^k \phi\left(\widehat{\rho_A}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_A}^*\right)\right)\right]_{k=1(1)K} \quad (A:4)$$

$$\frac{\partial D}{\partial \widehat{\rho_B}^*} = -\left[\frac{1}{n^B}\sum_{i=1}^{n^B}\sum_{j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* \phi\left(\widehat{\rho_B}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_B}^*\right)\right]$$

$$(A:5)$$

$$\frac{\partial D}{\partial \widehat{\beta_B}^*} = -\left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B} X_{i_B}{}^k \phi\left(\widehat{\rho_B}^* \sum_{i,j=1}^{n^B} w_{ij_B}\left(\frac{y}{z}\right)_{j_B}^* + X_{i_B}\widehat{\beta_B}^*\right)\right)\right]_{k=1(1)K}$$

$$(A:6)$$

Appendix- III

The weight (the share of the particular variable in the aggregate characteristics effect) of the k^{th} *explanatory variable* is:

$$V_{\Delta X}^{k} = \frac{\left\{\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\beta^{*k}_{A}\right\}\times \Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right)}{\left\{\left(\bar{X}_{A-}\bar{X}_{B}\right)\beta^{*}_{A}+\left(\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}-\overline{W_B\left(\frac{Y}{Z}\right)^{*}_{B}}\right)\widehat{\rho_A}^{*}\right\}\times \Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right)}$$

[From (10)]

Hence,

The share of the k^{th} explanatory variable in the *difference in the incidence of poverty*

$=V_{\Delta X}^{k}\times$ Aggregate Characteristics Effect

$$=V_{\Delta X}^{k}\times\left[\left\{\left(\bar{X}_{A-}\bar{X}_{B}\right)\beta^{*}_{A}+\left(\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}-\overline{W_B\left(\frac{Y}{Z}\right)^{*}_{B}}\right)\widehat{\rho_A}^{*}\right\}\times \Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right)\right]$$

[From (8)]

$$=\left\{\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\beta^{*k}_{A}\right\}\times \Phi^{1}\left(\widehat{\rho_A}^{*}\overline{W_A\left(\frac{Y}{Z}\right)^{*}_{A}}+\overline{X_A}\widehat{\beta_A}^{*}\right) \qquad \text{(A: 7)}$$

Thus for positive estimate of the share of k^{th} explanatory variable in the difference in the incidences of poverty,

$$\left\{\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\beta^{*}{}_{A}^{k}\right\}\times\ \Phi^{1}\left(\widehat{\rho_{A}}^{*}\overline{\mathrm{W_{A}}\left(\frac{\mathrm{Y}}{\mathrm{Z}}\right)^{*}}_{\mathrm{A}}+\overline{X_{A}}\widehat{\beta_{A}}^{*}\right)>0\ ; \qquad \text{[From (A: 7)]}$$

$\Rightarrow\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\beta^{*}{}_{A}^{k}>0$, Φ^{1} being the pdf of standard normal distribution is greater than zero.

$$\Rightarrow\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\left(\frac{-\beta_{A}{}^{k}}{\sigma_{A}}\right)>0\ ,\left[since\ \beta^{*}{}_{A}^{k}\ =\ \frac{-\beta_{A}{}^{k}}{\sigma_{A}}\right]$$

$$\Rightarrow\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\beta_{A}{}^{k}<0$$

$$\Rightarrow\left(\bar{X}_{A-}^{k}\bar{X}_{B}^{k}\right)\ <0\quad\forall\ \ \beta_{A}{}^{k}\ \ >0 \qquad \text{(A: 8)}$$

Appendix- IV

The weight (the share of the particular variable in the aggregate coefficients effect) of the k^{th} *explanatory variable* is:

$$V_{\Delta\beta}^{k} = \frac{\left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)}{\left\{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)}$$

[From (13)]

Hence,

The share of the k^{th} explanatory variable in the *difference in the incidence of poverty*

$= V_{\Delta\beta}^{k} \times$ Aggregate Coefficients Effect

$$= V_{\Delta\beta}^{k} \times \left[\left\{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right) + \overline{W_B\left(\frac{Y}{Z}\right)_B}^*\left(\widehat{\rho^*}_A - \widehat{\rho^*}_B\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right)\right]$$

[From (9)]

$$= \left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B}^* + \overline{X_B}\widehat{\beta_B}^*\right) \qquad \text{(A: 9)}$$

Thus for positive estimate of the share of the k^{th} explanatory variable in the difference in the incidences of poverty,

$$\left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} \times \Phi^1\left(\widehat{\rho_B}^* \overline{W_B\left(\frac{Y}{Z}\right)_B^*} + \overline{X_B}\widehat{\beta_B}^*\right) > 0 \quad ; \qquad \text{[From (A: 9)]}$$

$\Rightarrow \left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} > 0$, Φ^1 being the pdf of standard normal distribution is greater than zero.

$$\Rightarrow \overline{X}_B^k\left(\frac{\widehat{\beta_A}^k}{\widehat{\sigma_A}} - \frac{\widehat{\beta_B}^k}{\widehat{\sigma_B}}\right) < 0 \quad \left[since\ \beta^{*k}_A = \frac{-\beta_A{}^k}{\sigma_A} \text{ and } \beta^{*k}_B = \frac{-\beta_B{}^k}{\sigma_B}\right]$$

$$\Rightarrow \frac{\widehat{\beta_A}^k}{\widehat{\sigma_A}} - \frac{\widehat{\beta_B}^k}{\widehat{\sigma_B}} < 0$$

$$\Rightarrow \widehat{\beta_A}^k - \widehat{\beta_B}^k < 0 \qquad \text{if} \qquad \widehat{\sigma_A} \leq \widehat{\sigma_B}$$

In our case $\widehat{\sigma_A} \cong \widehat{\sigma_B} = 0.3$ [See Table 9, Column (4)]

Hence,

$$\widehat{\beta_A}^k - \widehat{\beta_B}^k < 0 \qquad \text{(A: 10)}$$