ANALYSIS OF POVERTY IN RURAL WEST BENGAL:
A SPATIAL APPROACH

SOMNATH CHATTOPADHYAY

A DISSERTATION SUBMITTED TO THE
INDIAN STATISTICAL INSTITUTE
IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR
THE AWARD OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

INDIAN STATISTICAL INSTITUTE

KOLKATA

DECEMBER 2010

CONTENTS

LIST OF TABLES

LIST OF APPENDICES

PREFACE

The problem of poverty is one of the core issues concerning developing countries like India. The formulation of an adequate programme to combat poverty is the *sine qua non* of any meaningful development plan. The key features relevant in this connection are the construction of an appropriate index of poverty and proper estimation of the measure. The present thesis has come up with some theoretical as well as empirical contributions taking into consideration various aspects of poverty measurement in the context of rural West Bengal, an eastern state of India. It has proposed some simple methodologies for the estimation of poverty starting from the micro level and has tried to address the problem of poverty from the perspective of policy formulation by making use of the proposed methods alongside the existing econometric methods.

It is my great pleasure to submit this thesis at the end of five years of rigorous research at the Indian Statistical Institute. It is also an occasion to express my gratitude to the persons I had the privilege to be associated with during the course of my work. This thesis owes much to their unstinting help and generosity.

First and foremost is my supervisor Professor Amita Majumder who not only introduced me to the world of research but guided me throughout with a keen interest and active support at every stage of my work. She was kind enough to forgive my delays in submission of assignments and painstakingly to go through the bulk of drafts I produced at the final stages. I owe an immense debt of gratitude to her.

I am indebted to Professor Dipankor Coondoo for the help he has provided in so many ways. I have had the privilege of being his co-author in two papers together with my supervisor Professor Majumder.

To Professor Nityananda Sarkar I want to express my deep respect and gratitude not only for his erudition but also for his philanthropic ideals which has been a source of inspiration for me. I am grateful to Professor Sharmila Banerjee of Calcutta University who had given me certain valuable suggestions as the external examiner during one of my presentations on the work. I have a debt of gratitude to Dr. Samarjit Das for his valuable help on various issues. I have a debt of gratitude also to Dr.

Manisha Chakraborty of IIM, Joka who had introduced me to the first lessons in STATA, the software which later came to be the mainstay of my work.

I reserve my deepest gratitude for this Institute, the alma mater of my grown-up years, that has not only put me on my mettle with its exacting academic standards but has provided me with a wholesome nourishment as well, with its sacred scholastic ambience, exposure to outstanding seminars, diverse other academic activities and the illuminating company of the most brilliant teachers and students.

I must mention here the following names:

Prof. Manoranjan Pal, Prof. Satya Ranjan Chakravarty, Prof. Abhirup Sarkar, Dr. Snigdha Chakrabarti, Prof. Manash Ranjan Gupta, Prof. Pradip Maiti, Prof. Manabendu Chattopadhyay, Prof. Tarun Kabiraj, Dr. Brati Sankar Chakraborty, Dr. Manipushpak Mitra, Dr. Chiranjib Neogi of the Economic Research Unit and Prof. Arup Bose of Stat-Math Unit, Kolkata.

I deeply acknowledge the influence of the loving company of all my co-research fellows: Soumyananda Dinda, Debabrata Mukhopadhyay, Debashis Mandal, Bidisha Chakraborty, Rituparna Kar, Anup Bhandari, Sahana Roy Chowdhury, Pratyush Vershney, Sarbari Choudhury and Sattwik Santra, Trishita Ray Barman, Conan Mukhopadhyay, Debasmita Basu, Srikanta Kundu, Sandip Sarkar, Kushal Banik Chowdhury, Priyabrata Dutta, Mannu Dwivedi, Rajit Biswas.

I extend my sincere thanks to all the Faculty Members, Research Associates and Associate Scientists of the department. I take this opportunity to express my heartfelt love and best wishes for every one of the Office. I also gratefully remember that during the initial stage of my work Sri Abhijit Mandal of ASU had taken the trouble to write many codes in the software MATLAB although later on I switched over to STATA exclusively.

I must also mention here that two of my personal friends, Subhabrata Sarkar and Hasanur Jaman, offered valuable help and assistance relevant to my work.

I take this opportunity to extend my sincere thanks to the two anonymous referees for their insightful comments which have helped to enrich the thesis.

Having acknowledged the contributions of all, I solemnly state at this point that for all slips and mistakes that may exist and for any dispute that may arise, the responsibility is entirely mine.

Finally, I must say that in spite of all of them, I would still not be able to make it without the inspiration of my wife, Manjari; her self-sacrificing support, unfailing enthusiasm and loving care have seen me through. I must also mention here the other members of my family - my parents, my elder brother, my sister-in-law and my sweet little niece, Ritaja whose loving company rejuvenates me every day.

Kolkata, December 14, 2010 Somnath Chattopadhyay

INTRODUCTION

The problem of poverty is an issue of perennial concern for developing countries like India and the formulation of an adequate programme to combat poverty is central to any development programme. The key features relevant in this connection are the construction of an appropriate index of poverty and proper estimation of the measure. The main focus of this thesis is ‘estimation of poverty’ taking into consideration various aspects of poverty measurement in the context of rural West Bengal, an eastern state of India. The analysis is spatial in nature with only cross sectional comparisons across districts.

First, an attempt is made to address the problem of data inadequacy. For estimation of poverty in India at the national and state levels (for rural and urban sectors separately), the National Sample Survey (NSS) Organization, Government of India, is the single most important source of data. However, at sub-state levels like districts, until recently, not all districts had adequate sample size to permit reliable estimation of poverty owing to the sampling design. On the other hand, for successful monitoring and implementation of developmental programs, it is essential to have information on socio-economic aspects at geographically disaggregated levels of district or below. Chapter 1 proposes a procedure that combines NSS and Census data to overcome the problem of data inadequacy. The procedure can be regarded as a type of Small Area Estimation (SAE) technique ((Quintano, Castellano, & Punzo, 2007), (Albacea, 2009), (Molina & Rao, 2010), (Hentschel, Lanjouw, Lanjouw, & Poggi, 2000), (Demombynes, Elbers, Lanjouw, Lanjouw, Mistiaen, & OZler, 2002), (Elbers, Lanjouw, & Lanjouw, 2003)) in which the scanty district level observed data set obtained from the nation-wide survey is supplemented by much richer district level information available from census and other sources. The proposed procedure is illustrated using NSS 55th round (1999-2000) data, which has the problem of data inadequacy at the district level in some states.[1]

Next, the issue of *spatial aspect* of poverty has been addressed through various approaches in Chapters 2 – 5. The importance of this aspect lies in the fact that the targeting of spatial anti-poverty policies depends crucially on the ability to identify the characteristics of different areas. One source of spatial variation in poverty estimates is the spatial difference in prices. In the absence of district level official poverty lines and district level spatial price

[1] In the next round (61st round, 2004-2005), the latest one, this problem has been reduced to a large extent by a revision in the sampling scheme. The later chapters of this thesis are based on NSS 61st round data. See Appendix (at the end of this thesis) for a description of NSS data.

indices, conventionally, poverty at the district level is estimated using the state level poverty line provided by the Planning Commission, Government of India. To examine the extent to which this procedure masks the variation in poverty estimates compared to that using district level poverty lines, Chapter 2 proposes a method of estimating spatial price indices, using which district level price indices (with state as base) and the corresponding district level poverty lines are obtained. Estimates of district level poverty based on *district level poverty lines* and those using the *conventional state level poverty line* are compared for rural West Bengal using NSS 61st round data. The method does not require item-specific price or unit-value data and hence overcomes the problem of data inadequacy in the context of prices. More importantly, in calculating the price indices, it allows inclusion of items of expenditure for which separate data on price and quantity is usually not recorded.

An alternative source of spatial variation in estimates of poverty is the geographically segregated units characterized by their intrinsic nature of development status (level of living). Assessment of this variation is necessary for prioritization of policy measures with a view to lowering the disparities in the levels of economic well-being across the spatial units. Given the fact that there is considerable difference in the levels of economic well being in two parts of Bengal, viz., North and South Bengal, a traditional division of West Bengal with respect to the River Hooghly, the aim in Chapter 3 is to identify the sources and characteristics affecting the differential levels of economic well being (poverty) in the two parts. The difference in the incidences of poverty is decomposed using the Oaxaca decomposition method (Oaxaca, 1973) into a *characteristics effect*, showing the effect of the regional characteristics and a *coefficients effect* showing the effect of the differential impact of the characteristics over the two regions.

Chapter 4 introduces the earnings frontier approach in explaining monthly consumption expenditure (a proxy for income) in terms of *human capital* and *endowments* of a household. Individuals who translate their potential earnings into actual earnings enjoy a fully efficient position. In contrast, individuals who earn less than their potential earnings suffer from some kind of *earnings inefficiency*. This chapter estimates an earnings frontier using the parametric stochastic frontier approach (SFA) (Jensen, Gartner, & Rassler, 2006) and classifies households in terms of *efficiency* scores. Splitting the sample into an *efficient* and *inefficient* part based on the estimated frontier, the status of poverty in the two groups is studied using the Oaxaca decomposition of the poverty gap. It thus tries to establish a link between the notion of *efficiency* and the *coefficients effect* discussed in Chapter 3. The result obtained is interpreted in light of the *poor but efficient hypothesis* (Chong, Lizarondo, Cruz, Guerrero, & Smith, 1984).

Chapter 5 is a spatial reformulation of Chapter 3 through introduction of spatial autoregressive dependence in the monthly consumption expenditure values within North and South Bengal. This is based on the notion, known as *Tobler's First Law of Geography* (Tobler, 1970), that nearby entities often share more similarities than entities which are far apart. Here the proximity between 'neighbours' has been defined in terms of 'economic' distance. The spirit of the model is that in addition to the overall differences between North and South Bengal characteristics, the determinants of the 'neighbouring households' within the two parts have a role to play in the difference in poverty estimates. A comparison of the results with those of Chapter 3 shows that there is marked difference in the shares and magnitudes of *aggregate characteristics effect* and *aggregate coefficients effect* from those obtained in the non-spatial analysis in Chapter 3, where the *aggregate characteristics effect* and the aggregate coefficients effect had a more or less balanced share.

Each chapter has Appendices, which mainly present detailed derivations of some of the results used in the chapter and some additional Tables.

Chapter 6 summarizes the contents of previous chapters and gives concluding remarks.

A description of the NSS data, used throughout the thesis, is provided at the end of the thesis in the form of Appendix.

The Bibliography has been prepared using Microsoft Word 2007 and is based mainly on the APA style.

The thesis consists of theoretical as well as empirical contributions. The empirical work has been done using the software STATA (Versions 8 & 9). Starting from the estimation of poverty at the micro level, it proposes some methodologies and attempts to address the problem of poverty from the perspective of policy formulation using the proposed methods and existing econometric methods.

CHAPTER 1

DISTRICT LEVEL POVERTY ESTIMATION

1.1 Introduction

Nation-wide socio-economic surveys are usually designed using *large area* as the *domain of estimation*. The sample design and/or the sample size of these surveys are thus such that fairly reliable estimates of the basic parameters of interest can be obtained at the national level (and also at the state level, in case of large countries like India), but not at lower (sub-state) levels like district/county etc. However, socio-economic information at geographically disaggregated levels of district/county or below is often required nowadays for successful monitoring and implementation of developmental programs at such levels. For example, while examining the efficiency gains from *targeting* in anti-poverty program in Mexico, (Baker & Grosh, 1994) observed that only a small improvement over *uniform transfer* of money would be achieved, if such a program was designed at the state level. The improvement would, however, be considerable, if the program was designed at district or neighbourhood level and that would require reliable estimates of poverty at district or neighbourhood level. Household level data obtained from a nation-wide survey based on *large area* as the *domain of estimation* may not give reliable district level poverty estimates because the number of sample households of a district/ neighbourhood may be smaller than that required to get a reliable estimate and/or the set of sample households of a district/neighbourhood may not constitute a *representative sample* for the district/neighbourhood.

The problem, in principle, may be resolved by substantially increasing the total sample size, which may increase the number of sample households observed in districts. But that may not be feasible due to resource constraints, apart from the possibility of substantial increase in non-sampling errors.[1]

---

[1] For NSS surveys, there is a provision of centre-state participation. For every state, NSSO and the state statistical office survey equal number of sample units. The samples covered by NSSO and state statistical office are known as the *central* and the *state* sample, respectively. NSSO processes only the central sample data and publishes reports based on these. Formally, pooling the central and state sample data sets would double the sample size at every stage of sampling and hence might ease the problem of inadequate sample size at the district level. However, pooling may be undertaken only if difference between the central and state estimates at district level is within 30 per cent of pooled estimates. The other necessary condition for obtaining pooled estimates is that data entry layout for both state and central samples are identical, or at least compatible. Otherwise pooling of the estimates is not advisable as it may worsen the situation (Sastry, 2003).

An alternative is to use the *Small Area Estimation* (SAE) technique ((Quintano, Castellano, & Punzo, 2007), (Albacea, 2009); (Molina & Rao, 2010) ; (Hentschel, Lanjouw, Lanjouw, & Poggi, 2000); (Demombynes, Elbers, Lanjouw, Lanjouw, Mistiaen, & OZler, 2002); (Elbers, Lanjouw, & Lanjouw, 2003). In SAE methods, the scanty district level observed data set obtained from the nation-wide survey is supplemented by much richer district level information available from census and other sources. Such information augmentation may help getting reliable district level statistics without any increase of the survey cost and the non-sampling error. The SAE models are broadly categorized into two groups, viz., (i) *the traditional indirect techniques* including the synthetic and composite methods of estimation and (ii) the *model based methods* including the regression-synthetic, empirical best linear unbiased prediction (EBLUP), empirical Bayes (EB) and the hierarchical Bayes (HB) techniques. Another model based approach developed of late by the World Bank is the (Elbers, Lanjouw, & Lanjouw, 2003) (ELL) method of estimation.

So far as the *indirect methods* are concerned, *Broad Area Ratio Estimator* (BARE) is one simple SAE model. By applying the rate obtained from a broad area using the survey data to the small area populations (obtained from the census), estimates are found for the small area. The crucial assumption underlying BARE is that the broad area should be large enough to allow for a reliable direct survey estimate but should be homogenous with respect to the characteristic of interest (See (McEwin & Elazar, 2006)). The *indirect synthetic estimation* technique, described by (Purcell & Kish, 1979), is a procedure that first uses sample data to estimate the variable of interest for different subclasses of the population at some higher level of aggregation. The estimates are then scaled down by adjusting it for compositional differences at the small area level. Like the BARE, the underlying assumption is still quite restrictive in the sense that the small area is assumed to exactly represent the larger area structurally with respect to the variable of interest. For correction of the bias of the synthetic estimator against the potential instability of a design-based direct estimator, a composite estimator, which is the weighted average of the above two estimators is formed. The optimal weights are obtained as the function of the mean square errors of the estimators and their covariance and can be estimated from the data.

The *model based* SAE techniques are broadly classified as *Area Level Random Effect Models* (Fay & Herriot, 1979), used when auxiliary information is available only at area level and *Nested Error Unit Level Regression Model* (Battese, Harter, & Fuller, 1988), when specific covariates are available at unit level. The *regression-synthetic model estimation* is a two-stage procedure which utilizes the linear regression model in predicting the poverty incidence. The predicted values using a two stage weighted least squares regression estimates serve as the regression-synthetic estimates (Albacea, 2009). The *empirical best linear unbiased prediction* (EBLUP) estimator is a model based estimator and it is similar to a composite estimator in the sense that it combines the direct or design-based unbiased estimator with the regression-synthetic estimator. In the *empirical Bayes* (EB) approach, the posterior distribution of the parameters of interest given the data is first obtained, assuming that the model parameters are known. The model parameters are estimated from the marginal distribution of the data, and inferences are then based on the estimated posterior distribution. In the *hierarchical Bayes* (HB) approach, a prior distribution on the model parameters is specified and the posterior distribution of the parameters of interest is then obtained (Ghosh & Rao, 1994). While in the regression based models the mean of the variable of interest is modeled, a more complete picture is obtained in the *M-quantile regression* methods by modeling the different quantile values of the variable of interest along with the mean. The central idea behind using M-quantiles to measure area effects is that area effects can be described by estimating a quantile value for each area (group) of a hierarchical data set (Chambers & Tzavidis, 2006). Extensions to deal with nonlinearities in the relationship between the variable of interest and the covariates have been proposed for linear mixed models (Opsomer, Claeskens, Ranalli, Kauermann, & Breidt, 2008) and for M-quantile models in the context of small area estimation (Pratesi & Salvati, 2008). In the M-quantile model, a specific quantile of the variable of interest, given the covariates, is described as an additive model in which some covariates enter the model parametrically and some others non parametrically. The relationship is left unspecified and learnt from the data through penalized splines (Pratesi, Ranalli, & Salvati, 2008) in the nonparametric case.

The ELL method has two stages, the first and second stages involving analysis with survey data and census data, respectively. Briefly, in the first stage a regression relationship explaining variation in per capita household total consumer expenditure in terms of a vector of household characteristics is estimated taking care of the various econometric issues involved. The model of logarithm of per-capita expenditure is estimated using Feasible Generalized Least Squares (FGLS) method.[2] In the second stage, this estimated relationship is used to generate a simulated value of per capita household total consumer expenditure,

[2] See (White, 1980), (Greene W. H., 2003).

based on which the district level poverty is estimated. Repeating the process of simulated data generation and the corresponding poverty estimation several times and then averaging the district level poverty estimate over simulation, the final district level poverty estimate is obtained. However, as pointed out by (Tarozzi & Deaton, 2009), in the ELL method useful matching of survey and census data requires a degree of homogeneity in terms of definition of explanatory variables.

In this chapter, an alternative method of estimation of poverty at sub-state level (district) is proposed for situations where the number of sample households at the sub-state level of interest is not always abundant. This method belongs to the category of *synthetic indirect methods* and uses minimal auxiliary information in terms of population. Using the subgroup decomposable property of the Foster, Greer, Thorbecke (FGT) measure of poverty (Foster, Greer, & Thorbecke, 1984), poverty estimates for the sub-state level are obtained by solving a system of linear equations. The merit of this procedure is contingent upon the assumption that, given a reliable state level poverty estimate, the estimate excluding any one of the districts (that is, the estimate based on all other districts pooled together) is reliable. The proposed method can be applicable to other economic indicators measuring proportions, where there is serious scarcity of data at the required level.[3]

It is expected that the proposed method will yield reliable district level poverty estimates essentially because of the more intensive use of the available data.

The plan of the chapter is as follows. Section 1.2 proposes the estimation method; Section 1.3 describes the data and results; Section 1.4 presents the conclusions. Appendix A1.1 – A1.4 at the end of this chapter present derivation of results and additional Tables.

1.2 A Proposed Estimator of District Level Poverty Measure

Suppose a state has *K* districts with population $(\boldsymbol{N_k}; \boldsymbol{k = 1, 2, \ldots\ldots, K})$. Denote the district poverty measures required to be estimated by $(\boldsymbol{\pi_k}; \mathbf{k = 1, 2, \ldots\ldots, K})$. Let $(\boldsymbol{\Pi_{-k}}; \mathbf{k = 1,2,\ldots\ldots,K}$ be the poverty measure for the pooled population of households belonging to districts $(\mathbf{1, 2, \ldots\ldots, k-1, k+1, \ldots K})$ and $\boldsymbol{P_{-k}}$ be an estimate of $\boldsymbol{\Pi_{-k}}$ based on the pooled data set for all the (K-1) districts except district *k*.

[3] A paper titled 'District-level Poverty Estimation: A Proposed Method' (Coondoo, Majumder, & Chattopadhyay), based on this chapter, is forthcoming in *Journal of Applied Statistics*.

Now, for any *subgroup decomposable*[4] poverty measure, we may write

$$\Pi_{-k} = a_{k1}\pi_1 + a_{k2}\pi_2 + \cdots + a_{k,k-1}\pi_{k-1} + 0.\pi_k + a_{k,k+1}\pi_{k+1} + \cdots + a_{kK}\pi_K,$$

$$k = 1,2,\ldots,K; \quad (1.1)$$

where $a_{kj}\left(= \frac{N_j}{\sum_{i\neq k}^{K} N_i}\right)$ is the share of population of district *j* in the pooled population of districts $1,2,..k-1,k+1\ldots K$ (i.e., known functions of $(N_k; k = 1,2,..,K)$. Note that (1.1) constitutes a system of *K* linear equations in *K* unknown district poverty parameters $(\pi_k; k = 1,2,\ldots,K)$, given $\Pi_{-k}{}'$ s and $a_{kj}{}'$ s. Let us write this equation system in vector-matrix notation as:

$$\Pi_{-} = A\pi_{\sim} \quad (1.2)$$

where

$\boldsymbol{\Pi_{-}} = \{\boldsymbol{\Pi_{-k}}; \mathbf{k} = \mathbf{1},\mathbf{2},\ldots.,\mathbf{K}\}$, $\boldsymbol{\pi_{\sim}} = \{\boldsymbol{\pi_k}; \mathbf{k} = \mathbf{1},\mathbf{2},\ldots.,\mathbf{K}\}$ are (*K*×1) vectors and $\boldsymbol{A} = (\boldsymbol{a_{kj}})$ is a $(\boldsymbol{K}\times\boldsymbol{K})$ non-singular matrix of population shares[5]. Solving (1.2) for $\boldsymbol{\pi_{\sim}}$ we get

$$\boldsymbol{\pi_{\sim}} = \mathbf{A^{-1}\Pi_{-}} \quad (1.3)$$

Thus, (1.3) suggests the following estimator of $\boldsymbol{\pi_{\sim}}$

$$\boldsymbol{p_{\sim}} = \boldsymbol{A^{-1}}\mathbf{P_{-}} \quad (1.4)$$

where $\boldsymbol{p_{\sim}} = \{\boldsymbol{p_k}; \boldsymbol{k} = \mathbf{1},\mathbf{2},\ldots.,\boldsymbol{K}\}$, $\boldsymbol{P_{-}} = \{\boldsymbol{P_{-k}}; \boldsymbol{k} = \mathbf{1},\mathbf{2},\ldots.,\boldsymbol{K}\}$, $\boldsymbol{p_k}$ and $\boldsymbol{P_{-k}}$ being the estimator of $\boldsymbol{\pi_k}$ and $\mathbf{\Pi_{-k}}$, *k=1,2,…,K*, respectively. Given the variance-covariance matrix of $\boldsymbol{P_{-}} = \left[\left(\mathbf{Covariance}(\boldsymbol{P_{-j}},\boldsymbol{P_{-k}})\right)\right] = \boldsymbol{\Sigma}$, say, the corresponding variance- covariance matrix of $\boldsymbol{p_{\sim}}$ is given by

$$\boldsymbol{D}(\boldsymbol{p_{\sim}}) = \boldsymbol{A^{-1}}\mathbf{\Sigma}(\boldsymbol{A^{-1}})' \quad (1.5)$$

Note that, given the available data for districts, $\boldsymbol{P_{-k}}$'s may be estimated by pooling the data for all but one districts, in turn, and *A* may be calculated using data available from an extraneous source like population census. The estimated district level poverty measures $\boldsymbol{p_{\sim}}$will then be given by (1.4). In this context, it may be mentioned that even if the sample size for some districts is not large, $\boldsymbol{P_{-k}}$'s, being based on the pooled data for all but one districts, are expected to be fairly reliable estimates of the corresponding population poverty

[4] A subgroup decomposable poverty measure is the one for which the overall measure can be written as the population share- weighted sum of poverty measures of the individual subgroups (see (Foster, Greer, & Thorbecke, 1984) and (Bishop, Chow, & Zheng, 1995)).

[5] See Appendix A1.1 for proof of non-singularity of *A*.

measures.[6] Hence, $\boldsymbol{p_k}$'s, being based on $\boldsymbol{P_{-k}}$'s, are likely to be reasonably reliable estimates of the district level poverty measures. Thus, even when a district does not have enough sample observations to warrant a reliable estimate of the district level poverty measure, the proposed method may provide reasonably reliable estimates of district level poverty.[7]

To calculate the standard errors of the estimated $\boldsymbol{p_k}$'s using (1.5), an estimate of $\boldsymbol{\Sigma}$, the variance-covariance matrix of $\boldsymbol{P_{-k}}$'s, is required. This may be estimated by one of the following two proposed procedures.

(i) **Estimation of Σ Based on Sub-Sample Divergence of $p_{\sim}$:**

In situations where the sample design of the available survey data is based on an *interpenetrating network of samples* (IPNS), the survey data are available in the form of two or more sub-samples drawn *independently* from the same universe following the same sampling scheme. From each of these sub-sample data sets, estimate of the population parameter(s) of interest is (are) obtained and the weighted arithmetic mean of such *sub-sample estimates* gives the corresponding *combined sample estimate*. The sampling variance of the combined sample estimate is given by the variance of the sub-sample estimates[8]. For NSS data two independent subsamples are drawn. Thus, if p_k^m denotes estimated value of π_k from the *m*-th sub-sample, *m* = 1, 2 the *combined sample* estimate of π_k is

$p_k^c = \frac{1}{\sum_{m=1}^{2} n_m^s} \sum_{m=1}^{2} n_m^s p_k^m$; n_m^s denoting the number of sample households in subsample *m*.

The sampling variance of p_k^c is

$$V(p_k^c) = \frac{1}{2}\sum_{m=1}^{2}(p_k^m - p_k^c)^2$$

$$= \frac{1}{2}\left[\left(p_k^1 - \frac{1}{n_1^s+n_2^s}(n_1^s p_k^1 + n_2^s p_k^2)\right)^2 + \left(p_k^2 - \frac{1}{n_1^s+n_2^s}(n_1^s p_k^1 + n_2^s p_k^2)\right)^2\right]$$

$$= \frac{1}{2}\left\{\frac{(n_1^s)^2+(n_2^s)^2}{(n_1^s+n_2^s)^2}\right\}(p_k^1 - p_k^2)^2 = \frac{(p_k^1-p_k^2)^2}{4}; \text{ [when } n_1^s = n_2^s]$$

The standard error of p_k^c, therefore, is

---

[6] The method is applicable to situations where there are a large number of districts in a state so that leaving out one district will not cause much change in the state level poverty estimate. The implicit assumption here is that the characteristics of the districts are not extremely divergent.

[7] The method is different from the leave-one-out Jacknife method for testing robustness (see (Jiang, Lahiri, Wan, & Wu, 2001); (Larse, 2003); (Rao J. N., Small Area Estimation, 2003); (Sinha & Rao, 2008)).

[8] For surveys based on a complicated sample design, analytical derivation of the formula for sampling variance of the estimator of a parameter of interest is often difficult. The technique of IPNS eases the problem of estimation of standard errors of survey estimates in such cases. See (Cochran, 1953), (Som, 1965), (Murthy, 1967), (Levy & Lameshow, 1991) for a description of the IPNS technique.

$s.e.(p_k^c)=\left|\frac{(p_k^1-p_k^2)}{2}\right|$; [when $n_1^s = n_2^s$] (1.6)

(ii) **Bootstrap Estimation of Σ**:

We have P_{-k}'s estimated from the available data based on the method described above. Now, suppose X is the original data set consisting of data from all the K districts. For *each district*, a bootstrap[9] sample is generated independently. This yields observations from *all* districts (with their original sample size) comprising a state level sample (thus the districts are essentially treated as strata when the re-sampling is done at the state level from *X*). Let X^1 be the first set of re-sampled state level data and let P_{-k}^1 be the estimate of Π_{-k} ($k=1,2,\ldots K$) from this re-sampled data set. Repeating the process of re-sampling R times, R values of P_{-k}, i.e., P_{-k}^1, $P_{-k}^2, \ldots, P_{-k}^R$ are obtained for every k. Based on these, an estimate of the variance-covariance matrix Σ of P_{-k}'s is obtained, using which in (1.5) the required variance-covariance matrix $D(p_{\sim})$ of $p_{\sim}$ is obtained. The positive square roots of the diagonal elements of $D(p_{\sim})$ are then taken as the standard errors of corresponding elements of $p_{\sim}$.

1.3 Data and Results

The method of estimation proposed in the previous section has been applied to the household level data on consumer expenditure collected through the employment-unemployment enquiry in the NSS 55th round (July 1999 – June 2000) survey[10]. The estimation exercise has been done for the rural sector of West Bengal. For illustration of the methodology estimates have also been presented for rural Madhya Pradesh, which has the problem of data

[9] See (Efron & Tibshirani, 1986) for a description of bootstrap method. See (Heinrich, 1988), (Rongve, 1995), (Osberg & Xu, 2000), (Davidson & Flachaire, 2007) for application of bootstrap method in the analysis of poverty.

[10] The empirical exercise has been repeated with the 61st round (2004-2005) NSS employment-unemployment data for rural West Bengal, results of which have been presented in the appendix. The state level rural poverty line of Rs.382.82 per capita per month has been used for the 61st round data (Source: http://www.cbhidghs.nic.in/writereaddata/mainlinkFile/Socio-Economic%20Indicators.pdf). Since the proposed methodology is particularly relevant for scanty data, results relating to the 55th round, which has this feature for some districts, have been reported here.

inadequacy. The required district level population data for these states have been taken from the Indian 2001 population census.[11]

A class of sub-group decomposable measures of poverty proposed by (Foster, Greer, & Thorbecke, 1984), has been used in the present empirical exercise. In its continuous form, the measure is given by

$$F_\alpha = \int_0^z \left(\frac{z-y}{z}\right)^\alpha dy \qquad (1.7)$$

y and z being the individual income level and the state-level poverty line, respectively. Depending on the value of the parameter α, three different poverty measures are obtained; viz., $\alpha = 0$, $\alpha = 1$ and $\alpha = 2$ give the head count ratio, the poverty gap measure and the squared poverty gap measure, respectively. Henceforth, we shall refer to these measures as FGT0, FGT1 and FGT2, respectively. The measure (1.7) in discrete form is written as

$$F_\alpha = \frac{1}{N}\sum_{y_i \le z} \left(\frac{z-y_i}{z}\right)^\alpha \qquad (1.8)$$

y_i and N being the income of the ith person and the number of persons in the society, respectively. To use these poverty measures for estimation of poverty from unit level household survey data on per capita income/consumer expenditure, the sample design of which is not *self-weighting*, the following *multiplier-adjusted* discrete form of the measure has been used

$$P_\alpha = F_\alpha = \frac{1}{\sum_{j=1}^{n} m_j}\sum_{y_i \le z} m_i \left(\frac{z-y_i}{z}\right)^\alpha \qquad (1.9)$$

where n is the sample size and m_i is the *multiplier* associated with the ith sample household. Here, we first estimate FGT0, FGT1 and FGT2 for each district using the conventional method. The corresponding standard errors have been found using (i) sub-sample divergence, and (ii) bootstrap method. Next, we estimate the FGT measures using the proposed methodology and compute the corresponding standard errors using (i) sub-sample divergence and (ii) the variance-covariance matrix $D(p_\sim)$.

For both methods, the state level rural poverty lines of Rs.350.17 per capita per month for West Bengal and Rs. 311.34 for Madhya Pradesh (at 1999-2000 prices) have been used [12]. Tables 1.1 - 1.3 present the estimates of the measures and the corresponding standard errors for FGT0 ($\alpha = 0$), FGT1 ($\alpha = 1$) and FGT2 ($\alpha = 2$), respectively, for West Bengal. Tables

[11] Since population is the only auxiliary variable that is required in this methodology, empirical comparison with other SAE methods have not been done, as the data requirement for these methods is much higher.

[12] Source: http://planningcommission.nic.in/reports/articles/ncsxna/ar_pvrty.htm, Planning Commission, Government of India.

1.4 - 1.6 present the estimates of the measures and the corresponding standard errors for FGT0 ($\alpha = 0$), FGT1 ($\alpha = 1$) and FGT2 ($\alpha = 2$), respectively, for Madhya Pradesh. The important observations that emerge from Tables 1.1 - 1.6 are as follows: The poverty estimates from the proposed method are quite close to the usual estimates, in particular, for FGT1 and FGT2, for most of the districts where sample sizes are reasonably large. For small sample size (e.g., for Panna: Table 1.4, columns 3 and 4) considerable difference is observed in the FGT0 measure. However, for Panna, which has only 12 observations, standard errors from subsample divergence could not be estimated because all the observations belong to one subsample only. It is interesting to note that in this case the proposed method yields a reliable estimate of poverty.

(i) For all districts and all the three measures the bootstrapped standard errors estimated using the proposed variance-covariance structure (column 8) are smaller than the standard errors estimated for each district separately (columns 6) for a large number of cases. Comparison based on subsample divergence (columns 5 and 7), however, does not show any clear pattern.[13]

(ii) A major discrepancy is observed for district Haora in West Bengal. In this case, for FGT2, while the estimate based on the individual district (column 3) is quite low (0.001), the one estimated using the proposed method (column 4) turns out to be negative (-0.0001), which, however, is small in magnitude and statistically non significant. As the method is all about solving a system of linear equations, one cannot possibly guarantee positive solutions always. A source of such discrepancy could be that the sample from this district is possibly not a representative sample.[14]

Tables 1.7 and 1.8 present a summary of the results obtained above for West Bengal and Tables 1.9 and 1.10 present similar results for Madhya Pradesh. It is observed that, in general, for all categories of 'Percentage discrepancy of the proposed estimates compared to conventional estimates', in majority of the cases the Relative Standard Error (RSE)[15] for the proposed method is less than the corresponding RSE of the conventional method. This is

[13] The bootstrap method is likely to be much superior for estimation of covariance matrices compared to the method of subsample divergence.

[14] The possibility of such a case arises due to the presence of negative eigen value of the weight matrix A.

[15] RSE is computed as the ratio of the standard error and the point estimate.

more clearly observed for cases with higher discrepancy between the two estimates. This indicates that for such districts the proposed method yields better estimates.

1.4 Conclusion

The problem of inadequacy of data at the required level of aggregation has been tackled by gathering information from a higher level of aggregation and applying it for some specific poverty measures. The proposed method generally yields more reliable estimates at the district level, because here the district level estimate is based on a much larger sample size obtained by pooling several district level data. This simple method is expected to be useful at any level of aggregation, given that reliable estimates are available at next higher level of aggregation. The method has been illustrated using poverty measures satisfying the property of sub-group decomposability, where the weights are population weights. Other weight functions like district-share of Domestic Product in a state may also be used if data on such shares are available. The proposed method has enormous potential in the sense that, it can be applied to other socio-economic indicators, where population share weighted pooled estimates are meaningful and there is serious scarcity of data at the required level. Some examples are: Human Development Index (HDI), employment/unemployment rate, literacy rate and additively decomposable measures of income inequality and occupational segregation.

TABLES

Table 1.1 Estimates of FGT0 for Districts of West Bengal (Rural: NSS 55th Round)

District	No. of Obs.	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Kochbihar	216	0.433	0.432	0.0017	0.0379	0.0092	0.0388
Jalpaiguri	204	0.566	0.518	0.0315	0.0404	0.0148	0.0320
Darjiling	96	0.217	0.250	0.1481	0.0493	0.1151	0.0420
West Dinajpur	204	0.351	0.350	0.0881	0.0363	0.0885	0.0362
Maldah	204	0.599	0.753	0.0364	0.0731	0.1790	0.1430
Murshidabad	324	0.681	0.687	0.1164	0.0303	0.1119	0.0365
Nadia	216	0.304	0.307	0.1857	0.0371	0.1853	0.0362
North 24 Paraganas	384	0.162	0.147	0.0231	0.0226	0.0158	0.0246
South 24 Paraganas	468	0.276	0.285	0.0026	0.0258	0.0034	0.0246
Haora	192	0.053	0.041	0.0039	0.0182	0.0102	0.0250
Hugli	287	0.193	0.198	0.0670	0.0287	0.0694	0.0279
Medinipur	816	0.303	0.305	0.0368	0.0188	0.0341	0.0182
Bankura	196	0.562	0.542	0.0223	0.0421	0.0183	0.0410
Puruliya	192	0.656	0.669	0.0680	0.0374	0.0954	0.0432
Bardhaman	360	0.204	0.212	0.0375	0.0247	0.0536	0.0243
Birbhum	192	0.572	0.559	0.0446	0.0421	0.0517	0.0400

Table 1.2 Estimates of FGT1 for Districts of West Bengal (Rural: NSS 55th Round)

District	No. of Obs.	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Kochbihar	216	0.079	0.078	0.0054	0.0091	0.0067	0.0090
Jalpaiguri	204	0.094	0.089	0.0060	0.0098	0.0025	0.0073
Darjiling	96	0.043	0.049	0.0320	0.0124	0.0247	0.0104
West Dinajpur	204	0.072	0.071	0.0294	0.0101	0.0299	0.0101
Maldah	204	0.149	0.201	0.0344	0.0314	0.0932	0.0612
Murshidabad	324	0.170	0.172	0.0542	0.0121	0.0532	0.0135
Nadia	216	0.050	0.051	0.0318	0.0079	0.0330	0.0078
North 24 Paraganas	384	0.027	0.023	0.0107	0.0045	0.0097	0.0051
South 24 Paraganas	468	0.045	0.047	0.0024	0.0056	0.0021	0.0053
Haora	192	0.005	0.002	0.0013	0.0020	0.0023	0.0039
Hugli	287	0.020	0.022	0.0039	0.0039	0.0049	0.0044
Medinipur	816	0.044	0.045	0.0104	0.0035	0.0094	0.0035
Bankura	196	0.121	0.116	0.0121	0.0113	0.0113	0.0108
Puruliya	192	0.144	0.147	0.0110	0.0120	0.0178	0.0135
Bardhaman	360	0.033	0.035	0.0015	0.0054	0.0028	0.0055
Birbhum	192	0.115	0.113	0.0098	0.0112	0.0114	0.0105

Table 1.3 Estimates of FGT2 for Districts of West Bengal (Rural: NSS 55th Round)

District	No. of Obs.	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Kochbihar	216	0.021	0.021	0.0028	0.0033	0.0028	0.0032
Jalpaiguri	204	0.023	0.023	0.0007	0.0036	0.0003	0.0027
Darjiling	96	0.012	0.013	0.0093	0.0042	0.0073	0.0035
West Dinajpur	204	0.023	0.023	0.0099	0.0045	0.0102	0.0045
Maldah	204	0.051	0.071	0.0182	0.0138	0.0426	0.0264
Murshidabad	324	0.058	0.059	0.0238	0.0062	0.0234	0.0066
Nadia	216	0.014	0.014	0.0084	0.0031	0.0089	0.0029
North 24 Paraganas	384	0.006	0.005	0.0032	0.0013	0.0030	0.0016
South 24 Paraganas	468	0.012	0.013	0.0014	0.0022	0.0012	0.0021
Haora	192	0.001	-0.0001	0.0003	0.0004	0.0005	0.0011
Hugli	287	0.003	0.004	0.0002	0.0009	0.0006	0.0012
Medinipur	816	0.010	0.010	0.0033	0.0011	0.0028	0.0011
Bankura	196	0.034	0.032	0.0054	0.0041	0.0049	0.0039
Puruliya	192	0.046	0.047	0.0008	0.0061	0.0031	0.0067
Bardhaman	360	0.009	0.010	0.0023	0.0020	0.0009	0.0020
Birbhum	192	0.033	0.032	0.0038	0.0045	0.0041	0.0041

Table 1.4 Estimates of FGT0 for Districts of Madhya Pradesh (Rural: NSS 55th Round)

District	No. of Obs.	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Morena	96	0.1663	0.1397	0.1453	0.0466	0.1883	0.0526
Bhind	96	0.1098	0.1469	0.0395	0.0357	0.0068	0.0351
Gwallior	59	0.3422	0.3428	0.0467	0.0690	0.0574	0.0682
Datia	36	0.0452	0.1826	0.0533	0.0415	0.0618	0.0290
Shivpuri	108	0.2643	0.2954	0.1070	0.0452	0.0928	0.0349
Guna	96	0.2184	0.2511	0.0947	0.0473	0.0804	0.0368
Tikamgarh	96	0.3707	0.3752	0.0704	0.0551	0.0545	0.0421
Chhatarpur	84	0.4662	0.4876	0.0533	0.0617	0.0747	0.0808
Panna	12	0.6917	0.4657		0.1416	0.0705	0.0387
Sagar	96	0.4872	0.4618	0.1677	0.0563	0.1260	0.0416
Damoh	95	0.6846	0.5974	0.0412	0.0537	0.0293	0.0385
Satna	102	0.2599	0.2818	0.0050	0.0477	0.0035	0.0403
Rewa	108	0.3075	0.3316	0.1355	0.0489	0.0951	0.0352
Shahdol	95	0.4968	0.5121	0.0294	0.0565	0.0373	0.0659
Sidhi	107	0.3194	0.3254	0.0584	0.0517	0.0615	0.0473
Mandsaur	96	0.1992	0.1692	0.0049	0.0445	0.0026	0.0522
Ratlam	96	0.2771	0.3050	0.0226	0.0488	0.0154	0.0367
Ujjain	96	0.1737	0.2295	0.0198	0.0395	0.0099	0.0313
Shajapur	96	0.2463	0.2777	0.0248	0.0475	0.0192	0.0391
Dewas	108	0.3913	0.3915	0.1097	0.0539	0.0901	0.0431
Jhabua	90	0.4314	0.4329	0.1714	0.1186	0.1671	0.1209
Dhar	108	0.4645	0.4544	0.0198	0.0594	0.0211	0.0510
Indore	48	0.3659	0.3667	0.0095	0.0781	0.0179	0.0791
Khargone	192	0.4070	0.4111	0.0892	0.0390	0.1181	0.0499
East Nimar - Khandwa	96	0.5223	0.5010	0.0513	0.0579	0.0351	0.0483
Rajgarh	47	0.2884	0.3384	0.0255	0.0694	0.0390	0.0372
Vidisha	96	0.4549	0.4425	0.0092	0.0522	0.0076	0.0439
Sehore	96	0.2904	0.2773	0.0027	0.0513	0.0232	0.0589
Raisen	96	0.5486	0.5236	0.0624	0.0540	0.0499	0.0459
Betul	96	0.5471	0.5229	0.0428	0.0582	0.0324	0.0521
Hoshangabad	96	0.3977	0.3984	0.0961	0.0529	0.1132	0.0620
Jabalpur	192	0.2187	0.1245	0.0477	0.0328	0.0556	0.0527
Narsimhapur	96	0.4497	0.4328	0.0490	0.0547	0.0327	0.0364
Mandla	96	0.6023	0.6983	0.0092	0.0513	0.0106	0.0824
Chhindwara	103	0.6231	0.5909	0.0664	0.0537	0.0513	0.0469
Seoni	97	0.5788	0.5275	0.0598	0.0533	0.0521	0.0387
Balaghat	96	0.6669	0.6192	0.0177	0.0504	0.0133	0.0439
Surguja	192	0.3345	0.3321	0.0554	0.0365	0.0539	0.0376
Bilaspur	384	0.4901	0.6010	0.0026	0.0297	0.0066	0.0602
Raigarh	192	0.3180	0.2695	0.0638	0.0368	0.1049	0.0587
Rajnandgaon	96	0.4166	0.4181	0.0221	0.0524	0.0244	0.0562
Durg	180	0.3551	0.3639	0.0173	0.0406	0.0111	0.0309
Raipur	384	0.3025	0.2568	0.0148	0.0257	0.0152	0.0377
Bastar	192	0.4936	0.5662	0.0002	0.0402	0.0007	0.0661

Table 1.5 Estimates of FGT1 for Districts of Madhya Pradesh (Rural: NSS 55th Round)

District	No. of Obs.	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Morena	96	0.0330	0.0267	0.0318	0.0124	0.0417	0.0133
Bhind	96	0.0081	0.0185	0.0060	0.0042	0.0024	0.0055
Gwallior	59	0.0729	0.0731	0.0025	0.0201	0.0023	0.0200
Datia	36	0.0176	0.0451	0.0208	0.0162	0.0184	0.0103
Shivpuri	108	0.0471	0.0568	0.0202	0.0093	0.0189	0.0073
Guna	96	0.0366	0.0461	0.0157	0.0096	0.0136	0.0075
Tikamgarh	96	0.0614	0.0668	0.0191	0.0123	0.0146	0.0098
Chhatarpur	84	0.1004	0.1043	0.0372	0.0200	0.0491	0.0240
Panna	12	0.0834	0.0863		0.0256	0.0013	0.0065
Sagar	96	0.1241	0.1142	0.0468	0.0175	0.0358	0.0128
Damoh	95	0.2186	0.1794	0.0273	0.0231	0.0197	0.0169
Satna	102	0.0425	0.0498	0.0016	0.0095	0.0010	0.0080
Rewa	108	0.0772	0.0800	0.0445	0.0161	0.0316	0.0114
Shahdol	95	0.0945	0.0956	0.0009	0.0153	0.0013	0.0173
Sidhi	107	0.0436	0.0472	0.0115	0.0099	0.0154	0.0094
Mandsaur	96	0.0387	0.0312	0.0044	0.0110	0.0058	0.0123
Ratlam	96	0.0598	0.0664	0.0125	0.0136	0.0100	0.0105
Ujjain	96	0.0383	0.0508	0.0081	0.0107	0.0048	0.0081
Shajapur	96	0.0489	0.0571	0.0064	0.0127	0.0048	0.0099
Dewas	108	0.0895	0.0891	0.0447	0.0166	0.0367	0.0134
Jhabua	90	0.1081	0.1089	0.0281	0.0330	0.0248	0.0314
Dhar	108	0.1731	0.1611	0.0232	0.0315	0.0144	0.0267
Indore	48	0.0674	0.0681	0.0011	0.0160	0.0085	0.0164
Khargone	192	0.0816	0.0799	0.0160	0.0104	0.0206	0.0127
East Nimar - Khandwa	96	0.1158	0.1111	0.0216	0.0163	0.0163	0.0137
Rajgarh	47	0.0523	0.0690	0.0251	0.0169	0.0193	0.0092
Vidisha	96	0.1149	0.1094	0.0139	0.0194	0.0108	0.0155
Sehore	96	0.0618	0.0586	0.0100	0.0164	0.0060	0.0186
Raisen	96	0.1146	0.1102	0.0303	0.0170	0.0250	0.0145
Betul	96	0.1496	0.1398	0.0295	0.0226	0.0232	0.0200
Hoshangabad	96	0.0981	0.1001	0.0353	0.0173	0.0415	0.0197
Jabalpur	192	0.0543	0.0365	0.0197	0.0106	0.0334	0.0162
Narsimhapur	96	0.1061	0.1005	0.0130	0.0161	0.0095	0.0108
Mandla	96	0.1656	0.2016	0.0024	0.0212	0.0028	0.0327
Chhindwara	103	0.1430	0.1352	0.0047	0.0211	0.0055	0.0180
Seoni	97	0.1701	0.1472	0.0062	0.0220	0.0088	0.0158
Balaghat	96	0.1399	0.1307	0.0095	0.0170	0.0076	0.0138
Surguja	192	0.0762	0.0758	0.0086	0.0102	0.0083	0.0103
Bilaspur	384	0.1191	0.1555	0.0076	0.0100	0.0175	0.0211
Raigarh	192	0.0578	0.0387	0.0229	0.0085	0.0371	0.0140
Rajnandgaon	96	0.0828	0.0825	0.0046	0.0143	0.0048	0.0151
Durg	180	0.0498	0.0586	0.0079	0.0074	0.0032	0.0060
Raipur	384	0.0477	0.0278	0.0074	0.0057	0.0086	0.0084
Bastar	192	0.1089	0.1245	0.0077	0.0122	0.0127	0.0203

Table 1.6 Estimates of FGT2 for Districts of Madhya Pradesh (Rural: NSS 55th Round)

District	No. of Obs	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional Method		Proposed Method	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Bootstrapped (using $D(p_{\sim})$)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Morena	96	0.0089	0.0064	0.0091	0.0054	0.0127	0.0049
Bhind	96	0.0012	0.0049	0.0010	0.0007	0.0018	0.0016
Gwallior	59	0.0216	0.0217	0.0047	0.0089	0.0016	0.0080
Datia	36	0.0069	0.0159	0.0081	0.0050	0.0067	0.0041
Shivpuri	108	0.0107	0.0153	0.0047	0.0035	0.0054	0.0020
Guna	96	0.0081	0.0122	0.0030	0.0027	0.0029	0.0022
Tikamgarh	96	0.0149	0.0180	0.0061	0.0027	0.0045	0.0032
Chhatarpur	84	0.0335	0.0346	0.0188	0.0109	0.0245	0.0129
Panna	12	0.0138	0.0258		0.0062	0.0042	0.0019
Sagar	96	0.0460	0.0416	0.0182	0.0101	0.0140	0.0064
Damoh	95	0.0889	0.0712	0.0191	0.0218	0.0135	0.0083
Satna	102	0.0093	0.0126	0.0009	0.0021	0.0006	0.0022
Rewa	108	0.0264	0.0274	0.0172	0.0124	0.0123	0.0045
Shahdol	95	0.0266	0.0261	0.0011	0.0061	0.0015	0.0065
Sidhi	107	0.0097	0.0113	0.0043	0.0035	0.0062	0.0029
Mandsaur	96	0.0105	0.0075	0.0014	0.0031	0.0019	0.0043
Ratlam	96	0.0196	0.0220	0.0075	0.0106	0.0058	0.0048
Ujjain	96	0.0114	0.0160	0.0027	0.0043	0.0017	0.0033
Shajapur	96	0.0144	0.0177	0.0050	0.0063	0.0039	0.0046
Dewas	108	0.0304	0.0302	0.0168	0.0118	0.0137	0.0059
Jhabua	90	0.0407	0.0412	0.0009	0.0174	0.0010	0.0142
Dhar	108	0.0917	0.0830	0.0125	0.0384	0.0067	0.0236
Indore	48	0.0173	0.0178	0.0026	0.0054	0.0026	0.0054
Khargone	192	0.0237	0.0219	0.0072	0.0095	0.0091	0.0051
East Nimar - Khandwa	96	0.0353	0.0343	0.0047	0.0127	0.0035	0.0057
Rajgarh	47	0.0158	0.0224	0.0111	0.0081	0.0082	0.0054
Vidisha	96	0.0433	0.0406	0.0104	0.0094	0.0083	0.0079
Sehore	96	0.0224	0.0215	0.0079	0.0065	0.0072	0.0094
Raisen	96	0.0363	0.0353	0.0085	0.0087	0.0070	0.0062
Betul	96	0.0591	0.0545	0.0174	0.0171	0.0140	0.0097
Hoshangabad	96	0.0329	0.0335	0.0133	0.0124	0.0156	0.0084
Jabalpur	192	0.0220	0.0179	0.0149	0.0169	0.0238	0.0082
Narsimhapur	96	0.0383	0.0357	0.0128	0.0132	0.0092	0.0058
Mandla	96	0.0619	0.0767	0.0034	0.0220	0.0053	0.0154
Chhindwara	103	0.0536	0.0503	0.0021	0.0140	0.0012	0.0136
Seoni	97	0.0722	0.0605	0.0017	0.0223	0.0034	0.0095
Balaghat	96	0.0442	0.0417	0.0008	0.0098	0.0007	0.0057
Surguja	192	0.0232	0.0229	0.0019	0.0077	0.0016	0.0042
Bilaspur	384	0.0443	0.0611	0.0052	0.0107	0.0114	0.0105
Raigarh	192	0.0154	0.0062	0.0085	0.0048	0.0139	0.0037
Rajnandgaon	96	0.0228	0.0223	0.0026	0.0087	0.0029	0.0062
Durg	180	0.0126	0.0166	0.0034	0.0026	0.0014	0.0020
Raipur	384	0.0123	0.0035	0.0023	0.0028	0.0023	0.0031
Bastar	192	0.0344	0.0378	0.0028	0.0075	0.0050	0.0092

Table 1.7 Comparison of the Magnitudes of the Poverty Estimates and the Corresponding RSE between the Proposed and Conventional Methods: Case of Bootstrapped Standard Error West Bengal (Rural: NSS 55th Round)

	FGT0		FGT1		FGT2	
Percentage discrepancy of the proposed estimates compared to conventional estimates	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventio nal)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-5	68.75	63.63	50.00	50.00	56.25	55.56
5-10	12.50	50.00	25.00	75.00	12.50	100.00
10-15	0.00		12.50	50.00	6.25	0.00
15-20	6.25	100.00	0.00		12.50	50.00
20-25	6.25	0.00	0.00		0.00	
25-30	6.25	0.00	0.00		0.00	
30-35	0.00		6.25	0.00	0.00	
35-40	0.00		0.00		6.25	0.00
40-45	0.00		0.00		0.00	
>45	0.00		6.25	0.00	6.25	100.00
Total	100.00		100.00		100.00	

Table 1.8 Comparison of the Magnitudes of the Poverty Estimates and the Corresponding RSE between the Proposed and Conventional Methods: Case for Sub-sample Divergence West Bengal (Rural: NSS 55th Round)

	FGT0		FGT1		FGT2	
Percentage discrepancy of the proposed estimates compared to conventional estimates	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (convention al)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-5	68.75	36.36	50.00	37.50	56.25	44.44
5-10	12.50	100.00	25.00	50.00	12.50	100.00
10-15	0.00		12.50	50.00	6.25	0
15-20	6.25	100.00	0.00		12.50	50.00
20-25	6.25	0	0.00		0.00	
25-30	6.25	0	0.00		0.00	
30-35	0.00		6.25	0	0.00	
35-40	0.00		0.00		6.25	0
40-45	0.00		0.00		0.00	
>45	0.00		6.25	0	6.25	100.00
Total	100.00		100.00		100.00	

Table 1.9 Comparison of the Magnitudes of the Poverty Estimates and the Corresponding RSE between the Proposed and Conventional Methods: Case of Bootstrapped Standard Error Madhya Pradesh (Rural: NSS 55th Round)

	FGT0		FGT1		FGT2	
Percentage discrepancy of the proposed estimates compared to conventional estimates	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-5	45.45	55.00	34.09	53.33	29.55	38.46
5-10	13.64	100.00	20.45	88.89	20.45	77.78
10-15	13.64	83.33	6.82	66.67	2.27	100.00
15-20	13.64	16.67	13.64	66.67	9.09	75.00
20-25	2.27	0	4.55	50.00	6.82	66.67
25-30	0		2.27	100.00	4.55	0
30-35	6.82	100	11.36	40.00	2.27	100.00
35 -40	0		0		4.55	50.00
40-45	2.27	0	2.27	0.00	6.82	100.00
>45	2.27	100	4.55	100	13.64	66.67
Total	100.00		100.00		100.00	

Table 1.10 Comparison of the Magnitudes of the Poverty Estimates and the Corresponding RSE between the Proposed and Conventional Methods: Case for Sub-sample Divergence Madhya Pradesh (Rural: NSS 55th Round)

	FGT0		FGT1		FGT2	
Percentage discrepancy of the proposed estimates compared to conventional estimates	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)	% of cases in each category	% of cases (within each category) where RSE (proposed) < RSE (conventional)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-5	45.45	50.00	34.09	60.00	29.55	61.54
5-10	13.64	100.00	20.45	77.78	20.45	66.67
10-15	13.64	83.33	6.82	33.33	2.27	100.00
15-20	13.64	33.33	13.64	66.67	9.09	25.00
20-25	2.27	0	4.55	100.00	6.82	66.67
25-30	0		2.27	100.00	4.55	0
30-35	6.82	100.00	11.36	40.00	2.27	100.00
35 -40	0		0		4.55	50.00
40-45	2.27	0	2.27	0	6.82	100.00
>45	2.27	100.00	4.55	100.00	13.64	66.67
Total	100.00		100.00		100.00	

APPENDICES

Appendix A1.1: Proof of Non-Singularity of the Weight Matrix

$|A| = |(a_{kj})|$

$=\left|\left(\frac{n_j}{\sum_{s\neq k} n_s}\right)\right|$

$$=\prod_{j=1}^{K} n_j \times \frac{1}{\prod_{k=1}^{K}\sum_{s\neq k} n_s}\begin{vmatrix} 0\ 1\ 1\ \dots\ 1\ 1\ 1 \\ 1\ 0\ 1\ \dots\ 1\ 1\ 1 \\ 1\ 1\ 0\ \dots\ 1\ 1\ 1 \\ \dots\dots\dots\dots\dots\dots. \\ \dots\dots\dots\dots\dots\dots. \\ \dots\dots\dots\dots\dots\dots. \\ 1\ 1\ 1\ \dots\ 0\ 1\ 1 \\ 1\ 1\ 1\ \dots\ 1\ 0\ 1 \\ 1\ 1\ 1\ \dots\ 1\ 1\ 0 \end{vmatrix}$$

[As multiplication of a determinant by a scalar is equivalent to a determinant with a row (column) multiplied by the same scalar]

Let $D_K = (d_{ij})_{K\times K}$; d_{ij}=1$\forall\ i,j$.

i = 1,2,…,K ; j=1,2,….,K

Now, the characteristic polynomial of D is

$|D_K - \lambda I_K|$=$C_0\lambda^K + C_1\lambda^{K-1} + C_2\lambda^{K-2} +$……..+$C_K$

where $C_0 = (-1)^K$, λ being variable.

$C_r = (-1)^{K-r} \times$ [Sum of principal minors of $D_K$of order r] ;

r = 1,2,…,K

Here C_r= 0 for r = 2,3,…..,K as all principal minors of D of order $r \geq 2$ are zero.

Hence $|D_K - \lambda I_K|$=$(-1)^K\lambda^K + (-1)^{K-1}K\lambda^{K-1}$

So the characteristic roots are λ =0, K

So for λ =+1, $|D_K - \lambda I_K| \neq 0$

$$\text{So}\begin{vmatrix} 0\ 1\ 1\ \dots\ 1\ 1\ 1 \\ 1\ 0\ 1\ \dots\ 1\ 1\ 1 \\ 1\ 1\ 0\ \dots\ 1\ 1\ 1 \\ \dots\dots\dots\dots\dots\dots. \\ \dots\dots\dots\dots\dots\dots. \\ \dots\dots\dots\dots\dots\dots. \\ 1\ 1\ 1\ \dots\ 0\ 1\ 1 \\ 1\ 1\ 1\ \dots\ 1\ 0\ 1 \\ 1\ 1\ 1\ \dots\ 1\ 1\ 0 \end{vmatrix} \neq 0$$

Also, $\prod_{j=1}^{K} n_j \neq 0$, $\prod_{k=1}^{K} \sum_{s \neq k} n_s \neq 0$

$\Longrightarrow |A| \neq 0$

i.e., A is non-singular.

Hence A^{-1} exists.

Appendix A1.2 Table A1.1: Estimates of FGT0 for Districts of West Bengal (Rural: NSS 61st Round)

District	No of observations	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional		Proposed	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Boot-strapped
Darjiling	80	0.2464	0.2464	0.0174	0.0597	0.0137	0.0531
Jalpaiguri	240	0.3231	0.3148	0.0030	0.0356	0.0031	0.0330
Kochbihar	200	0.1290	0.1435	0.0374	0.0305	0.0346	0.0259
Uttar dinajpur	200	0.5158	0.5117	0.0292	0.0461	0.0322	0.0490
Dakshin dinajpur	120	0.2304	0.2311	0.0988	0.0495	0.0962	0.0461
Maldah	270	0.4175	0.4097	0.0406	0.0397	0.0399	0.0388
Murshidabad	440	0.4877	0.4982	0.0037	0.0311	0.0125	0.0326
Birbhum	240	0.3583	0.3558	0.0573	0.0394	0.0561	0.0365
Bardhaman	400	0.1162	0.1165	0.0196	0.0191	0.0185	0.0207
Nadia	320	0.1876	0.1869	0.0073	0.0266	0.0059	0.0270
North 24 Paraganas	360	0.1371	0.1262	0.0033	0.0239	0.0010	0.0261
Hugli	280	0.1324	0.1338	0.0099	0.0259	0.0080	0.0251
Bankura	280	0.2495	0.2493	0.0644	0.0344	0.0619	0.0313
Purulia	200	0.5074	0.5003	0.0806	0.0442	0.0706	0.0481
Medinipur	638	0.1493	0.1496	0.0003	0.0180	0.0042	0.0172
Howrah	200	0.1295	0.1197	0.0415	0.0300	0.0502	0.0327
South 24 Paraganas	520	0.1595	0.1528	0.0355	0.0204	0.0365	0.0218

Appendix A1.3 Table A1.2: Estimates of FGT1 for Districts of West Bengal (Rural: NSS 61st Round)

District	No of observations	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional		Proposed	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Boot-strapped
Darjiling	80	0.0439	0.0438	0.0043	0.0147	0.0043	0.0133
Jalpaiguri	240	0.0647	0.0623	0.0012	0.0092	0.0012	0.0084
Kochbihar	200	0.0199	0.0227	0.0018	0.0067	0.0021	0.0057
Uttar dinajpur	200	0.1012	0.1003	0.0064	0.0118	0.0070	0.0127
Dakshin dinajpur	120	0.0406	0.0407	0.0197	0.0100	0.0191	0.0097
Maldah	270	0.0772	0.0757	0.0145	0.0097	0.0141	0.0096
Murshidabad	440	0.0904	0.0925	0.0048	0.0080	0.0032	0.0083
Birbhum	240	0.0630	0.0625	0.0201	0.0091	0.0197	0.0087
Bardhaman	400	0.0170	0.0171	0.0042	0.0034	0.0039	0.0037
Nadia	320	0.0265	0.0263	0.0053	0.0047	0.0049	0.0048
North 24 Paraganas	360	0.0201	0.0179	0.0037	0.0047	0.0034	0.0052
Hugli	280	0.0242	0.0244	0.0072	0.0060	0.0068	0.0059
Bankura	280	0.0364	0.0367	0.0106	0.0073	0.0104	0.0068
Purulia	200	0.1035	0.1018	0.0153	0.0125	0.0130	0.0134
Medinipur	638	0.0243	0.0243	0.0046	0.0040	0.0038	0.0039
Howrah	200	0.0142	0.0119	0.0049	0.0043	0.0066	0.0049
South 24 Paraganas	520	0.0191	0.0173	0.0071	0.0033	0.0072	0.0038

Appendix A1.4 Table A1.3: Estimates of FGT2 for Districts of West Bengal (Rural: NSS 61st Round)

District	No of observations	Estimates from Individual Districts	Estimates from Proposed Method	Standard Errors			
				Conventional		Proposed	
				Subsample Divergence	Boot-strapped	Subsample Divergence	Boot-strapped
Darjiling	80	0.0118	0.0118	0.0016	0.0051	0.0015	0.0046
Jalpaiguri	240	0.0185	0.0178	0.0026	0.0034	0.0025	0.0031
Kochbihar	200	0.0050	0.0058	0.0006	0.0022	0.0004	0.0018
Uttar dinajpur	200	0.0276	0.0274	0.0017	0.0040	0.0019	0.0042
Dakshin dinajpur	120	0.0104	0.0105	0.0042	0.0032	0.0041	0.0032
Maldah	270	0.0210	0.0206	0.0034	0.0035	0.0033	0.0035
Murshidabad	440	0.0249	0.0255	0.0009	0.0032	0.0004	0.0032
Birbhum	240	0.0166	0.0165	0.0076	0.0033	0.0074	0.0032
Bardhaman	400	0.0039	0.0039	0.0002	0.0011	0.0001	0.0012
Nadia	320	0.0056	0.0055	0.0013	0.0014	0.0011	0.0014
North 24 Paraganas	360	0.0048	0.0041	0.0019	0.0016	0.0019	0.0017
Hugli	280	0.0073	0.0073	0.0041	0.0027	0.0040	0.0026
Bankura	280	0.0099	0.0099	0.0028	0.0029	0.0027	0.0028
Purulia	200	0.0290	0.0285	0.0039	0.0047	0.0032	0.0050
Medinipur	638	0.0074	0.0074	0.0027	0.0019	0.0026	0.0018
Howrah	200	0.0024	0.0016	0.0012	0.0010	0.0017	0.0012
South 24 Paraganas	520	0.0043	0.0038	0.0020	0.0011	0.0020	0.0012

CHAPTER 2

DISTRICT LEVEL POVERTY ESTIMATION: A SPATIAL APPROACH

2.1 Introduction

As mentioned in the 'Introduction', the targeting of spatial anti-poverty policies depends crucially on the ability to identify the characteristics of different areas. Several questions that arise in this context are: do some regions within the states have exceptional rates of poverty compared to other regions? To what extent do state-level poverty studies mask intra-regional variance in levels of poverty within a state? How does a regional focus on poverty affect our current measurement and understanding of economic well-being? Clearly, the choice and size of spatial units fundamentally alters the measurement of poverty. At high levels of *spatial disaggregation*, disparities in levels of income per capita increase. Similarly, at high levels of *aggregation* differences between areas are averaged out. Now, since there is also the possibility of variations in standards for defining poverty across the regions/districts of a state, another important question that arises is: does the poverty line vary spatially according to the costs of the goods in the market basket?

In the context of India, official estimate of state level poverty lines are provided by the Planning Commission. Given that there are ample variations in the cost of living conditions across different districts within a state, the official state level poverty line used for estimating poverty at the district level may not be appropriate and requires to be properly adjusted with district price index numbers (with reference to the state) to get the true picture of poverty at the district level. However, non availability of item wise price data at the sub-state level has so far prevented estimation of district specific price index numbers and hence of district specific poverty lines.

This chapter proposes a procedure for estimating regional consumer price index numbers based on the estimation of item-specific region wise Engel curves.[1] Given the problem of data inadequacy in developing countries, the basic question that it tries to answer here is "is it possible to find a method of estimation of a set of spatial consumer price index

[1] This Chapter is based on (Coondoo, Majumder, & Chattopadhyay, Estimating Spatial Consumer Price Indices Through Engel Curve Analysis, DOI: 10.1111/j.1475-4991.2010.00382.x, Published Online : March 25, 2010).

numbers using Engel curves that does not require data (i) at the household level, or (ii) on prices/unit values of goods and services consumed?" The district/region level price index numbers (with the state as base) are estimated using this method. The district/region level poverty lines are estimated by multiplying the official state level poverty line by the relevant price indices.

The chapter is organized as follows. Section 2.2 presents the background literature; Section 2.3 discusses the proposed methodology; Section 2.4 describes the data and results and finally Section 2.5 concludes. Appendix A2.1 – A2.10 at the end of this chapter present derivation of results, justifications and additional Tables.

2.2 The Background Literature

Appropriate consumer price index numbers are essential for comparison of real income levels or consumption patterns over time, across regions or across population groups. When more than two (regions/countries/population) groups are involved in a comparison of price or real income levels, the price index number problem is resolved in one of two major ways. The simpler and straightforward approach is to use a set of binary price index numbers and make pair-wise comparisons. Examples of this approach are (Sen, 1976), (Bhattacharyya, Joshi, & Roychowdhury, 1980), (Bhattacharya, Chatterjee, & Pal, 1988), (Coondoo & Saha, 1990), (Deaton, 2003) and (Deaton & Tarozzi, 2005). Use of binary comparisons approach, however, does not guarantee transitivity of price level comparisons except under unrealistic assumptions.

A second approach is to have a multilateral price level comparison, whereby a set of internally consistent price index numbers, popularly known as Purchasing Power Parities (PPP), are constructed on the basis of a set of group-specific price and quantity data for a common set of commodities (see (Geary, 1958); (Khamis, 1972); (Kravis, Heston, & Summers, 1978); (Balk, 1996); (Rao D. S., 1997); (Hill, 1997); (Diewert, 1999); (Neary, 2004)). As in the case of binary price index numbers, computation of a set of multilateral price index numbers requires price and quantity data of uniform quality, which is often rather difficult to obtain. To resolve the data problems arising from quality variation of items across groups and from gaps in the available price data, the Country Product Dummy (CPD) methodology was proposed (Summers, 1973). The CPD procedure, which is essentially a hedonic approach, offers a regression analysis-based methodology for constructing multilateral price index numbers (see (Kokoski, Moulton, & Zeischang, 1999), (Rao D. S., 2001)). A large part of the literature on multilateral price index numbers today is concerned with the construction of PPP's from commodity-specific price and quantity/expenditure share data using the CPD methodology. Application of CPD and similar methods to household level data has been proposed in recent works of (Aten & Menezes, 2002) and (Coondoo, Majumder, & Ray, 2004).

(Costa D. L., 2001) and (Hamilton, 2001) pioneered the use of Engel curves in the context of Consumer Price Indices (CPI). The basic idea underlying these studies is that if *a given* CPI is an accurate measure of cost of living, the CPI-deflated Engel curves (log-linear/ log-quadratic food share equations) estimated at different time points should coincide and temporal drift in CPI-deflated Engel curves will reflect systematic bias in measurement of CPI ((Barrett & Brzozowski, 2010). There have been various extensions of this approach in terms of specification (introduction of household demographics: (Logan, 2008); flexible semi parametric Engel curves: (Beatty & Larsen, 2005), (Larsen, 2007); flexible Almost Ideal Demand System (AIDS): (Barrett & Brzozowski, 2010)) as well as application (in the context of regional price index: (Papalia, 2006); Purchasing Power parity (PPP): (Almas, 2008)). All these studies, however, are based on pooled time series of *household level* cross section data. Also, a basic data requirement for these kinds of exercises is availability of estimates of relative price changes over time /region.

2.3 The Proposed Method

The data requirement for the procedure proposed here is minimal. It does not require region-specific data on prices of individual items. Formally, given a system of demand functions derived from an underlying cost (expenditure) function, it may be possible to derive estimates of the parameters appearing in the cost (expenditure) function from the estimated demand functions. One should then, in principle, be able to estimate the True Cost of Living Index (TCLI) number corresponding to a specified utility level. When a consumer expenditure data set covers regions facing different price situations, the region-specific Engel curves for individual items estimated from such a data set contain information about regional price level differentials, which if retrieved, can be used to construct regional TCLI's. This kind of procedure has already been suggested by (Fry & Pashardes, 1989). They investigate the conditions under which the Tornqvist price index number can be a reasonable approximation to the TCLI underlying a Price Independent Generalized Log linear (PIGLOG) demand system. Using the decomposition of the TCLI under PIGLOG as the sum of a basic index (the cost of living index at some minimum level of consumer expenditure) and a marginal index (Deaton & Muellbauer, 1980a), they apply the Tornqvist method to estimate the TCLI in a *systems* framework. It makes explicit reference to expenditure levels, commodity prices and household characteristics in the context of the AIDS of (Deaton & Muellbauer, 1980b) and the Translog model of demand, both members of the PIGLOG class.

The method proposed here is a more general one based on a two-component decomposition (a basic index and a marginal index) of the TCLI underlying a *Quadratic* PIGLOG system. The proposed procedure has several useful features:

(i) It overcomes the already mentioned problem of data inadequacy: it does not require item-specific price or unit-value data, and more importantly, allows inclusion of items of expenditure for which separate data on price and quantity are usually not recorded (e.g., meals away from home, expenditure on recreation, educational and health services etc.).

(ii) The method is essentially based on (single equation) Engel curve analysis and hence it is computationally simpler and no explicit algebraic form for the coefficients of the Engel curves (which are functions of prices) is required. An underlying assumption here is that the Engel curve is quadratic logarithmic (in budget share form)[2] and the form is the same for all the regions being compared.

(iii) It is not necessary that all items must be consumed in all regions[3]. Finally,

(iv) The procedure does not require household level expenditure data and can be applied to consumer expenditure data grouped by per capita income/total consumer expenditure class.

Estimation of TCLI using this method involves three steps. In the first step, a set of item-specific Engel curves, relating item-specific budget shares to the logarithm of per capita income/total consumer expenditure, are estimated for each region. The first component of the TCLI (the basic index) is estimated in the second step based on a pooled regression over items and regions. In the third step, the marginal index and the TCLI are estimated.

The cost function underlying Quadratic Logarithmic (QL) systems, (e.g., the Quadratic Almost Ideal Demand System (QUAIDS) of (Banks, Blundell, & Lewbel, 1997) and the Generalized Almost Ideal Demand System (GAIDS) of (Lancaster & Ray, 1998) is of the form:

---

[2] This is the most popular and commonly used form of budget share equation in the literature.

[3] A requirement, however, is that the set of items of the base region must be the union of all items consumed in different regions, as will be seen in the estimation procedure.

$$C(u,p) = a(p).exp\left(\frac{b(p)}{\frac{1}{ln\,u}-\lambda(p)}\right) \tag{2.1}$$

where p is the price vector, $a(p)$ is a homogeneous function of degree one in prices, $b(p)$ and $\lambda(p)$ are homogeneous functions of degree zero in prices, and u denotes the level of utility. By Shephard's lemma, the budget share functions corresponding to the cost function (2.1) are of the form

$$w_i = \frac{\partial\, ln\, a(p)}{\partial\, ln\, p_i} + \frac{\partial\, ln\; b(p)}{\partial\, ln\, p_i}\; ln\frac{y}{a(p)} + \frac{1}{b(p)}\frac{\partial\, \lambda(p)}{\partial\, ln\, p_i}\left(ln\frac{y}{a(p)}\right)^2 \quad ; i = 1, 2\ldots, n$$

or,

$$w_i = a_i(p) + b_i(p) ln\frac{y}{a(p)} + \frac{\lambda_i(p)}{b(p)}\left(ln\frac{y}{a(p)}\right)^2 \tag{2.2}$$

where y denotes nominal per capita income and i denotes item of expenditure.

The corresponding TCLI in logarithmic form comparing price situation p^1 with price situation p^0 is given by

$$ln\,P(p^1,p^0,u^*)=[ln\,a(p^1) - ln\,a(p^0)] + \left[\frac{b(p^1)}{\frac{1}{ln\,u^*}-\lambda(p^1)} - \frac{b(p^0)}{\frac{1}{ln\,u^*}-\lambda(p^0)}\right] \tag{2.3}$$

where u^* is the reference utility level. The first term of the R.H.S. of (2.3) is the logarithm of the basic index (measuring the cost of living index at some minimum benchmark utility level) and the second term is the logarithm of the marginal index. Note that for $p^1 = \theta p^0, \theta > 0$, $a(p^1) = \theta a(p^0)$, so that the basic index takes a value θ and hence, may be interpreted as that component of TCLI that captures the effect of uniform or average inflation on the cost of living. On the other hand, for $p^1 = \theta p^0, \theta > 0$; $b(p^1) = b(p^0)$ and $\lambda(p^1) = \lambda(p^0)$, the marginal index takes a value of unity. Hence, the marginal index may be interpreted as the other component of TCLI that captures the effect of a change in the relative price structure. If prices are normalized such that $b(p^0)$= 1 and $\lambda(p^0) = 1$ the TCLI for a reference utility level u^* becomes

$$P(p^1,p^0,u^*) = \frac{a(p^1)}{a(p^0)} exp\left[\frac{b(p^1)}{\frac{1}{ln\,u^*} - \lambda(p^1)} - \frac{1}{\frac{1}{ln\,u^*} -1}\right] \tag{2.4}$$

To build up our procedure, look at the region specific Engel curves of the form (2.2). Let p^r denote the price vector of region r, r = 0, 1, 2…R. Then, from (2.2), the budget share equations for region r can be written as

$$w_{ir} = a_i(p^r) + b_i(p^r)\, ln\frac{y_r}{a(p^r)} + \frac{\lambda_i(p^r)}{b(p^r)}\left(ln\frac{y_r}{a(p^r)}\right)^2$$

or,

$$w_{ir} = \alpha_{ir} + \beta_{ir}\, ln\left(\frac{y_r}{P_r}\right) + \gamma_{ir}\left(ln\left(\frac{y_r}{P_r}\right)\right)^2 \quad ; i = 1, 2, \ldots\ldots, n, \tag{2.5}$$

where $\alpha_{ir} = a_i(p^r); \beta_{ir} = b_i(p^r); \gamma_{ir} = \frac{\lambda_i(p^r)}{b(p^r)}; P_r = a(p^r)$.

In (2.5), P_r denotes the price level for region *r*, homogeneous of degree one in prices of the region.[4] The parameters $\alpha_{ir}, \beta_{ir}, \gamma_{ir}, P_r$ are functions of the price vector and are parameters for a given cross-sectional data situation where prices are fixed.

The budget share curves in (2.5), which are quadratic in logarithm of income, correspond to those of QUAIDS and GAIDS having underlying *cost functions* of the forms

$$C(u, p^r) = a(p^r)\, exp\left(\frac{b(p^r)}{\left(\frac{1}{ln\, u}\right) - \lambda(p^r)}\right); \tag{2.6}$$

$b(p^r) = \prod_{i=1}^{n} p_{ir}^{\beta_i}$, $\lambda(p^r) = \sum_{i=1}^{n} \lambda_i \ln p_{ir}$ for both QUAIDS and GAIDS, and

$\ln a(p^r) = \alpha_0^* + \sum_{i=1}^{n} \alpha_i^*\, ln\, p_{ir} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln p_{ir} \ln p_{jr}$, for QUAIDS and

$\ln a(p^r) = \alpha_0^* + \frac{1}{1-\sigma} ln\left(\sum_{i=1}^{n} \alpha_0^*\, ln\, p_{ir}^{(1-\sigma)}\right) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln p_{ir} \ln p_{jr}$, for GAIDS.

The parameters of the *budget share functions* derived from (2.6) using the above expressions relate to the parameters in (2.5) as follows (see (Banks, Blundell, & Lewbel, 1997)) and (Lancaster & Ray, 1998)):

$\alpha_{ir} = \alpha_i^* + \sum_{k=1}^{n} \delta_{ik}\, ln\, p_{kr}$ for QUAIDS,

$\alpha_{ir} = \frac{\alpha_i^* p_{ir}^{1-\sigma}}{\sum_{k=1}^{n} \alpha_k^* p_{kr}^{1-\sigma}} + \sum_{k=1}^{n} \delta_{ik} ln\, p_{kr}$ for GAIDS and

$\beta_{ir} = \beta_i, \gamma_{ir} = \frac{\lambda_i}{b(p^r)}, P_r = a(p^r)$ for both QUAIDS and GAIDS.

If the QUAIDS or the GAIDS system is estimated using an appropriate panel data set that contains adequate price variation for individual regions, then reliable estimates of all the parameters of the system will be obtained. Using these and the formula given in (2.4), TCLI's measuring cross-sectional regional price level differentials may then be easily calculated.

The objective here, however, is to explore whether or not such regional price level differences can be estimated using data from a single cross-section, without item level information on prices and *without specifying the algebraic forms* of $a(p^r)$, $b(p^r)$ and $\lambda(p^r)$ explicitly. Thus, the method is essentially based on (*single equation*) Engel curve analysis as opposed to the *systems* approach.

To estimate $\boldsymbol{a(p^r)}$, *i.e.*, $\boldsymbol{P_r}$, rewrite the budget share functions (2.5) as

$$w_{ir} = (\alpha_{ir} - \beta_i \pi_r - \gamma_{ir}\pi_r^2) + (\beta_i - 2\gamma_{ir}\, \pi_r) y_r^* + \gamma_{ir} y_r^{*2} \tag{2.7}$$

[4] Ratios of P_r's will measure the basic price index number of a region with the other region taken as base.

where $y_r^* = ln(y_r)$, $\pi_r = ln(P_r)$.[5] As mentioned earlier, for a single cross-sectional data set corresponding to a given price situation, $\alpha_{ir}, \gamma_{ir}, \beta_i$ and π_r are parameters to be estimated from the given data. Now, (2.7) is a set of $(R + 1)$ complete systems of *n-commodity* Engel curves, which are nonlinear in parameters. With an appropriate stochastic specification, (2.7) will be a *large* system of *nonlinear* SUR equations, which can, in principle, be estimated from a given set of data. However, such a simultaneous estimation of all the equations of (2.7) under parametric restrictions would be extremely difficult, if not impossible. We have, therefore, used the following alternative indirect estimation route.

As mentioned earlier, the suggested procedure for estimating TCLI's in (2.4) involves three stages.

In the first stage, a set of item-specific Engel curves relating budget shares to the logarithm of income are estimated (using equation (2.8) defined later) for each region.

In the second stage a(p^r), r = 0, 1, 2…,R is estimated.

In the third stage b(p^r) and $\lambda(p^r)$, r = 1, 2…,R are estimated using the normalisation $b(p^0) = \lambda(p^0) = 1$ (where p^0 denotes the price vector of the base region). Using these, the TCLI's are estimated for a given reference level of utility of the base region. It may be emphasized that $a(p^r)$, $b(p^r)$ and $\lambda(p^r)$ are estimated as composite variables and no explicit algebraic forms for these functions are assumed. The three stages are described below in detail.

Stage 1: Estimate the following log-quadratic budget share function, which is in the form of a linear regression equation:

$$w_{irj} = a_{ir} + b_{ir}y_{rj}^* + c_{ir}y_{rj}^{*2} + \varepsilon_{irj} \quad (2.8)$$

where the subscript $j (j = 1,2, \dots, H_r)$ denotes the per capita income/total consumer expenditure (PCE) class of a region, ε_{irj} is a random disturbance term and a_{ir}, b_{ir}, c_{ir} are the parameters[6].

Stage 2: Let $\hat{a}_{ir}$, $\hat{b}_{ir}$ and $\hat{c}_{ir}$ be the estimates of a_{ir}, b_{ir} and c_{ir} respectively. Given the estimates $\hat{a}_{ir}$, $\hat{b}_{ir}$ and $\hat{c}_{ir}$ from (2.7) and (2.8), we have

[5] Note that the $\beta_{ir}'s$ have been replaced by $\beta_i's$. That is, they do not have any *region effect*, or to put it differently, they are independent of prices. This is in line with the specifications in QUAIDS and GAIDS. Also, a justification for this form of budget share can be found in (Banks, Blundell, & Lewbel, 1997).

[6] The subscript *j* would denote the *j*-th sample household of region *r*, when household level expenditure data are used to estimate these Engel curves. As one of the objectives of proposing the method is to be able to estimate price indices in the absence of household level data, the procedure has been described with respect to grouped data. Note that as compared to household level data, for grouped data H_r will be small leading to smaller degrees of freedom.

$\hat{c}_{ir} = \gamma_{ir} + e^c_{ir}$, say (2.9.1)

$\hat{b}_{ir} = (\beta_i - 2\gamma_{ir}\pi_r) + e^b_{ir}$, say (2.9.2)

$\hat{a}_{ir} = (\alpha_{ir} - \beta_i\pi_r - \gamma_{ir}\pi_r^2) + e^a_{ir}$, say (2.9.3)

where $e^b_{ir,}$ $e^c_{ir,}$ e^b_{i0} are the errors in estimation of the parameters. The π_r's are then estimated as follows: equation (2.9.2) implies

$$\hat{b}_{ir} - \hat{b}_{i0} = \pi_0(2\hat{c}_{i0}) - \pi_r(2\hat{c}_{ir}) + e_{ir}\ ; \qquad i = 1, 2, \ldots, n; \qquad r = 1, 2, \ldots, R; \qquad (2.10)$$

where e_{ir} is a composite error term, which is a linear combination of the individual errors $e^b_{ir,}$ $e^c_{ir,}$ e^b_{i0} and e^c_{i0}.[7] Thus, the regression error is assumed to be present only because of estimation errors in the first stage and since the first stage parameters are consistently estimated, asymptotically equation (2.10) would hold exactly. However, typically, the error e_{ir} and the explanatory variables will be correlated. As an approximation, if we treat this as a multivariate errors-in-variables set up, consistent estimates of the π_r's can be obtained as

$$\hat{\pi} = \frac{1}{2}\left(\hat{C}'\Sigma^{-1}\hat{C} - NE_C'\Sigma^{-1}E_C\right)^{-1}\left(\hat{C}'\Sigma^{-1}\hat{B} - NE_C'\Sigma^{-1}E_B\right); \qquad (2.11)$$ [8]

where $\hat{C}$ is the matrix of explanatory variables in (2.10), $\hat{B}$ is the vector of dependent variables in (2.10), E_C is the matrix of estimation errors in *C*'s, E_B is the vector of estimation errors in *B* , *N* (= *nR*) is the sample size and Σ is the variance-covariance matrix of the error terms in (2.10). From the consistency property of OLS estimates it can be shown that asymptotically $E_C'\Sigma^{-1}E_C$ will converge to a null matrix and $E_C'\Sigma^{-1}E_B$ will converge to a null vector[9], thus yielding the usual GLS estimates.

Three points are noteworthy so far as this estimation procedure is concerned.

First, the procedure does not necessarily require that the number of items of expenditure and the composition of the set of items of expenditure be same for all regions.

Second, whereas in the literature π_0 is not estimated and is fixed exogenously ((Banks, Blundell, & Lewbel, 1997), (Deaton & Muellbauer, 1980a)), here an estimate of π_0 is obtained from the estimation process itself.

Finally, as already pointed out, the estimates of π_r's obtained by the above procedure are conditional upon the fact that the β_i's in equation (2.7) do not have any region effect.

---

[7] The matrix forms of the elements of the equation (2.10) are given in the Appendix A2.1.

[8] See (Deaton, 1997), (Coondoo, Majumder, & Chattopadhyay, DOI: 10.1111/j.1475-4991.2010.00382.x, Published Online : March 25, 2010).

[9] See Appendix A2.2.

Stage 3: Once the estimates of π_r's (and hence of $\ln a(p^r)$'s) are obtained, the next step involves the estimation of $b(p^r)$ and $\lambda(p^r)$ for every r. However, this set of estimated parameters would not suffice for calculating TCLI's reflecting regional price level differentials by equation (2.4). This is because without detailed price information $b(p^r)$ and $\lambda(p^r)$ cannot be calculated, even though estimates of $\gamma_{ir}\left(=\frac{\lambda_i}{b(p^r)}\right)$ are available.

The problem is resolved as follows:

Treat region r = 0 as the base region and take the utility levels of the base region as reference utility levels. Using equation (2.6) and the normalisation $(p^0) = \lambda(p^0) = 1$, the money metric utility u_0^h of a household of the base region that has nominal per capita income $y_{0h}\left(= C\left(u_{0,}^h p^0\right)\right)$ is given by

$$\frac{1}{\ln u_0^h} = \frac{1}{\ln \frac{y_{0h}}{a(p^0)}} + 1 \tag{2.12}$$

Now, combine equations (2.6) and (2.12) to obtain for region r

$$\frac{1}{ln\left(\frac{y_{rh}}{a(p^r)}\right)} = \frac{1}{b(p^r)}\left(\frac{1}{ln\frac{y_{0h}}{a(p^0)}} + 1\right) - \frac{\lambda(p^r)}{b(p^r)}, \tag{2.13}$$

where y_{rh} denotes the per capita nominal income required by a household of region r to have u_0^h utility level[10].

Using the relationship (2.13) estimation of $b(p^r)$ and $\lambda(p^r)$ is proposed from the following regression equation[11]

$$\left(\frac{1}{\ln \overline{y}_{rq} - \hat{\pi}_r}\right) = \frac{1}{b(p^r)}\left(\frac{1}{\ln \overline{y}_{oq} - \hat{\pi}_0} + 1\right) - \frac{\lambda(p^r)}{b(p^r)} + error \tag{2.14}$$

or, $\left(\frac{1}{\ln \overline{y}_{rq} - \hat{\pi}_r}\right) = \psi_r \left(\frac{1}{\ln \overline{y}_{oq} - \hat{\pi}_0} + 1\right) + \tau_r + error$, say, (2.14.1)

where $\hat{\pi}_r$ is the estimate of $ln\, a(p^r)$, which has already been obtained in the second stage of estimation and $\overline{y}_{rq}$ (q = 1, 2,..,Q) is the q-th quantile of per capita income/total consumer expenditure (PCE). The regression equations are estimated using region-specific data on PCE by quantiles (so that data from different regions are comparable), viz., $\left(\overline{y}_{rq}, q = 1,2,..,Q\right)$.[12]

[10] See Appendix A2.3 for derivation of equation (2.13).

[11] See Appendix A2.4 for a justification of the regression set up.

[12] This may produce noisy estimates if Q is not large. Thus, while for household level data there may not be any problem, for grouped data this may typically be the case. In the present empirical exercise, twenty quantile groups of PCE have been considered.

It is assumed that all the *q-th* quantile households of a given region have comparable utility levels.

Here again, it may be noted that both the regressor and the regressand contain estimated values of π's and hence are measured with error. However, under some mild conditions, the use of OLS can be justified[13]. Once estimates of $\pi_r, b(p^r), \lambda(p^r)$ are obtained this way, consistent estimates of the TCLI's for regions corresponding to given quantile levels, may be calculated using equation (2.4).

Recalling equation (2.4), the TCLI for any region is a function of the estimated coefficients at different stages. i.e., $\hat{P}(p^r, p^0, u^*) = f(\hat{\pi}_r, \hat{\pi}_0, \hat{\psi}_r, \hat{\tau_r})$.[14] The variance of the estimated index can be derived using the *delta method* ((Powell L. A., 2007), (Seber, 1982), (Oehlert, 1992), (Xu & Long, 2005))[15] as:

$$V(\hat{P}) = \left(\frac{\partial f}{\partial \hat{\pi}_r}\right)^2 var(\hat{\pi}_r) + \left(\frac{\partial f}{\partial \hat{\pi}_0}\right)^2 var(\hat{\pi}_0) + \left(\frac{\partial f}{\partial \hat{\psi}_r}\right)^2 var(\hat{\psi}_r) + \left(\frac{\partial f}{\partial \hat{\tau_r}}\right)^2 var(\hat{\tau_r})$$
$$+ 2\, cov[(\hat{\pi}_r, \hat{\pi}_0)] \left(\frac{\partial f}{\partial \hat{\pi}_r}\right)\left(\frac{\partial f}{\partial \hat{\pi}_0}\right) + 2\, cov[(\hat{\psi}_r, \hat{\tau_r})] \left(\frac{\partial f}{\partial \hat{\psi}_r}\right)\left(\frac{\partial f}{\partial \hat{\tau_r}}\right) \qquad (2.15)$$

The estimates of $var(.)'s$ and $cov(.)'s$ on the R.H.S of equation (2.15) can be obtained from the respective stages of regression. The covariance terms including second stage and third stage parameters have been dropped since regressions at respective stages are carried out independently.[16]

2.4 Data and Results

The data used here is 61st round NSS employment-unemployment data for rural West Bengal and the base region has been taken to be the state as a whole, which relates to the data for all districts combined. The data set covers 20 item expenditure categories, a number of which contain non-food and service items, jointly comprising total consumer expenditure.[17] To illustrate the method, the data have been grouped into twenty quantile classes of total consumption expenditure. This has been done as one of the objectives of proposing the

[13] See Appendix A2.4 for a justification of the use of OLS.

[14] $\hat{\psi}_r = \frac{1}{b(p^r)}$; $\tau_r = -\frac{\lambda(p^r)}{b(p^r)}$ [see equation (2.14.1)].

[15] The variance of any parameter which is a function of random variables can be approximated using the Delta method. The description of the delta method is given in Appendix A2.5.

[16] This part is based on (Chattopadhyay, 2010).

[17] See Appendix A2.6 for the list of items.

method is to be able to estimate price indices in the absence of household level data.[18] Price indices have been computed using the proposed method first taking the 'districts' as regions and then considering the 'NSS regions' as specified in the NSS data structure and treating the corresponding district price indices within a region to be the same.[19] The reference utility level has been taken to be the utility value (obtained from equation (2.12)) at the median level of expenditure for the reference region. Using the estimated price index numbers two sets of district specific poverty lines have been estimated from the official state level poverty line.

Estimated price indices have been reported in Table 2.1.[20] It is noted that there is ample variation in values of the estimated indices across the districts as well as across definitions of 'region'. It is also observed that in the rural sector, taking 'West Bengal' as numeraire, the Northern districts have lower price levels and the Southern districts have higher price levels. The size of the standard error in Table (2.1) suggests that there may be no significant differences in price levels across regions. Estimating price indices in North and South of the river, it is found that price index for the northern part is 0.87 and that for the southern part is 1.07. It is interesting to note that except for few districts (Kochbihar, Birbhum and Purulia) this segregation largely coincides with the traditional (geographical) division of North and South Bengal with respect to the River Hooghly[21]. The district specific poverty lines (= district price index × state poverty line), adjusted for the district price indices, and the corresponding adjusted FGT poverty measures have been reported in Table 2.2 and Table 2.3. A comparison of the adjusted FGT estimates with the unadjusted FGT estimates (derived in Chapter 1)[22] shows that, largely, poverty is overestimated in the districts of North Bengal and underestimated in the districts of South Bengal when the conventional state level poverty line is used for estimating poverty at the district level.

---

[18] The item-specific Engel curves for regions have been estimated by single-equation weighted least squares, using the estimated population proportion of individual PCE classes as weights. This should take care of the heteroscedasticity arising out of grouping of data. The heteroscedasticity problem due to dependence of the error variance on y^*, if any, should be largely taken care of by the use of Engel curve formulation in budget share form in our case due to the grouped nature of expenditure data used. However, for household level data, the issue of heteroscedasticity needs to be addressed appropriately (see (Deaton, 1997)). The analysis has also been done at the household level with the incorporation of household demographics. The parameter estimates in both the cases are found to be quite close.

[19] See Appendix A2.7 for the listing of Districts and NSS Regions.

[20] The estimates of the parameters of equation (2.14) are presented in Appendix A2.8.

[21] See Appendix A2.9 for the map of West Bengal at the end of this chapter. The river flows through the middle of Mursidabad district. The map in Appendix A2.10 shows the spatial variation in the cost of living indices across districts.

[22] See Appendix A1.2-A1.4.

2.5 Conclusion

This chapter has developed a method of estimating district specific price indices, which are the TCLI's of a quadratic PIGLOG demand system. The procedure is based on estimating region-specific Engel curves for a set of expenditure categories. The most important contribution of the procedure is that it addresses the issue of data inadequacy, a major problem in the context of developing countries. In other words, the method works even in a situation where unit level data are not available. This procedure does not require the expenditure categories to exhaust the consumer's budget. However, if the set of expenditure categories considered is exhaustive, the estimated consumer prices index numbers will be more accurate measures of the underlying true price level differentials.

The other notable features are as follows:

First, as it is intimately related to the quadratic PIGLOG demand system, it has a well-defined theoretical underpinning. Second, the data requirement is minimal in the sense that it can be implemented even on a set of grouped consumer expenditure data covering several regions. More importantly, region-specific separate data on quantity and price of individual consumer goods are not required for this procedure. Therefore, items of expenditure like "*services consumed*", "*medical expenses*", for which only expenditure data are available and separate quantity and price are often not well defined, can also be included. Third, no explicit algebraic form for the coefficients of the Engel curves (which are functions of prices) is required. Finally, as the results presented here would suggest, the empirical performance of the proposed procedure is satisfactory (this is also evident from Table A2.3 in the Appendix A2.8, as each component of the TCLI turns out to be highly significant).

Few issues, however, need to be addressed for future application of the methodology. Some outstanding questions, which remain unresolved in the present exercise, regarding the statistical properties of the estimates at different stages, need to be explored, possibly using Monte Carlo studies. We leave this as a future exercise.

Estimates of district level price indices reveal that districts in rural North Bengal have lower price levels compared to those in rural South Bengal. As a consequence, it is also observed that poverty is overestimated in the districts of North Bengal and underestimated in the districts of South Bengal when the conventional state level poverty line is used in place of district level poverty lines for estimating poverty at the district level. The fact that the demarcation of West Bengal by price levels is corroborated by the natural geographical demarcation provides a basis for exploration of the poverty situation in the two parts of West Bengal separately. This issue is taken up in the subsequent chapters of the thesis.

TABLES

Table 2.1 Estimates of Price Indices (Rural West Bengal: 2004-2005)

District	Price indices using the proposed method			
	Districts taken as Regions		NSS Regions taken as Regions	
	Estimate	Standard Error	Estimate	Standard Error
(1)	(2)	(3)	(4)	(5)
Darjiling	0.9802	0.1199	0.9392	0.3445
Jalpaiguri	0.9081	0.2039	0.9392	0.3445
Kochbihar	1.0276	0.2319	0.9392	0.3445
Uttar dinajpur	0.7759	0.2169	0.8763	0.2568
Dakshin dinajpur	0.9370	0.2032	0.8763	0.2568
Maldah	0.8425	0.1753	0.8763	0.2568
Murshidabad	0.7775	0.1385	0.8763	0.2568
Birbhum	0.8394	0.0876	0.8763	0.2568
Bardhaman	1.0974	0.1660	1.1047	0.3075
Nadia	1.0682	0.1868	0.8763	0.2568
North 24 Paraganas	1.0910	0.2465	1.1047	0.3075
Hugli	1.3329	0.3767	1.1047	0.3075
Bankura	1.0081	0.1777	1.0363	0.2600
Purulia	0.7547	0.1668	1.0363	0.2600
Medinipur	1.1239	0.1767	1.0363	0.2600
Howrah	1.0522	0.1656	1.1047	0.3075
South 24 Paraganas	1.0492	0.1747	1.1047	0.3075
West Bengal	1.0000		1.0000	

Table 2.2 Poverty Estimates Based on State and District Level Poverty Lines (Rural West Bengal: 2004-2005)

District	Estimated poverty line (Rs.) (Using Col (2) of Table 2.1)	Poverty estimate					
		FGT0		FGT1		FGT2	
		Using Col (2)	Using WB Poverty line (Rs. 382.82)	Using Col (2)	Using WB Poverty line (Rs. 382.82)	Using Col (2)	Using WB Poverty line (Rs. 382.82)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Darjiling	375.2450	0.2177	0.2464	0.0399	0.0439	0.0106	0.0118
Jalpaiguri	347.6227	0.2400	0.3231	0.0426	0.0647	0.0111	0.0185
Kochbihar	393.3720	0.1537	0.1290	0.0234	0.0199	0.0059	0.0050
Uttar dinajpur	297.0313	0.2135	0.5158	0.0261	0.1012	0.0046	0.0276
Dakshin dinajpur	358.7115	0.1740	0.2304	0.0296	0.0406	0.0070	0.0104
Maldah	322.5208	0.2088	0.4175	0.0328	0.0772	0.0075	0.0210
Murshidabad	297.6502	0.1726	0.4877	0.0234	0.0904	0.0055	0.0249
Birbhum	321.3266	0.1661	0.3583	0.0245	0.0630	0.0057	0.0166
Bardhaman	420.0930	0.1999	0.1162	0.0299	0.0170	0.0072	0.0039
Nadia	408.9210	0.2355	0.1876	0.0385	0.0265	0.0089	0.0056
North 24 Paraganas	417.6744	0.2052	0.1371	0.0330	0.0201	0.0082	0.0048
Hugli	510.2699	0.3986	0.1324	0.0790	0.0242	0.0255	0.0073
Bankura	385.9318	0.2519	0.2495	0.0381	0.0364	0.0103	0.0099
Purulia	288.9282	0.1627	0.5074	0.0251	0.1035	0.0048	0.0290
Medinipur	430.2520	0.2286	0.1493	0.0430	0.0243	0.0128	0.0074
Howrah	402.7885	0.1642	0.1295	0.0206	0.0142	0.0038	0.0024
South 24 Paraganas	401.6521	0.2038	0.1595	0.0267	0.0191	0.0060	0.0043

Table 2.3 Poverty Estimates Based on State and Region Level Poverty Lines (Rural West Bengal: 2004-2005)

District	Estimated poverty line (Rs.) (Using Col (4.) of Table 2.1)	Poverty estimate					
		FGT0		FGT1		FGT2	
		Using Col (2)	Using WB Poverty line (Rs. 382.82)	Using Col (2)	Using WB Poverty line (Rs. 382.82)	Using Col (2)	Using WB Poverty line (Rs. 382.82)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Darjiling	359.5470	0.2008	0.2464	0.0326	0.0439	0.0083	0.0118
Jalpaiguri	359.5470	0.2761	0.3231	0.0495	0.0647	0.0134	0.0185
Kochbihar	359.5470	0.0865	0.1290	0.0145	0.0199	0.0034	0.0050
Uttar dinajpur	335.4723	0.3267	0.5158	0.0560	0.1012	0.0124	0.0276
Dakshin dinajpur	335.4723	0.1343	0.2304	0.0203	0.0406	0.0044	0.0104
Maldah	335.4723	0.2444	0.4175	0.0405	0.0772	0.0097	0.0210
Murshidabad	335.4723	0.3207	0.4877	0.0474	0.0904	0.0117	0.0249
Birbhum	335.4723	0.2214	0.3583	0.0317	0.0630	0.0075	0.0166
Bardhaman	422.8859	0.2056	0.1162	0.0311	0.0170	0.0075	0.0039
Nadia	335.4723	0.1034	0.1876	0.0096	0.0265	0.0019	0.0056
North 24 Paraganas	422.8859	0.2205	0.1371	0.0353	0.0201	0.0089	0.0048
Hugli	422.8859	0.1935	0.1324	0.0366	0.0242	0.0114	0.0073
Bankura	396.7284	0.2718	0.2495	0.0442	0.0364	0.0119	0.0099
Purulia	396.7284	0.5494	0.5074	0.1183	0.1035	0.0346	0.0290
Medinipur	396.7284	0.1778	0.1493	0.0292	0.0243	0.0087	0.0074
Howrah	422.8859	0.1970	0.1295	0.0280	0.0142	0.0057	0.0024
South 24 Paraganas	422.8859	0.2608	0.1595	0.0370	0.0191	0.0085	0.0043

APPENDICES

Appendix A2.1 Expansion in Vector – matrix form of equation (2.10)

Writing (2.10) for each item, $i = 1(1)n$ and Region $r = 1(1)R$, one gets the following:

$$
\begin{pmatrix}
\hat{b}_{11}-\hat{b}_{10}\\
\hat{b}_{21}-\hat{b}_{10}\\
\hat{b}_{31}-\hat{b}_{10}\\
\cdot\\
\cdot\\
\cdot\\
\hat{b}_{n1}-\hat{b}_{10}\\
\hat{b}_{12}-\hat{b}_{10}\\
\hat{b}_{22}-\hat{b}_{20}\\
\hat{b}_{32}-\hat{b}_{30}\\
\cdot\\
\cdot\\
\cdot\\
\hat{b}_{n2}-\hat{b}_{n0}\\
\cdot\\
\vdots\\
\cdot\\
\cdot\\
\cdot\\
\hat{b}_{1R}-\hat{b}_{10}\\
\hat{b}_{2R}-\hat{b}_{20}\\
\hat{b}_{3R}-\hat{b}_{30}\\
\cdot\\
\cdot\\
\cdot\\
\hat{b}_{nR}-\hat{b}_{n0}
\end{pmatrix}
= 2 *
\begin{bmatrix}
\hat{c}_{10} & -\hat{c}_{11} & 0 & 0 & \cdots & & 0\\
\hat{c}_{20} & -\hat{c}_{21} & 0 & 0 & \cdots & & 0\\
\hat{c}_{30} & -\hat{c}_{31} & 0 & 0 & \cdots & & 0\\
. & . & 0 & 0 & \cdots & & 0\\
. & . & 0 & 0 & \cdots & & 0\\
. & . & 0 & 0 & \cdots & & 0\\
\hat{c}_{n0} & -\hat{c}_{n1} & 0 & 0 & \cdots & & 0\\
\hat{c}_{10} & 0 & -\hat{c}_{12} & 0 & \cdots & & 0\\
\hat{c}_{20} & 0 & -\hat{c}_{22} & 0 & \cdots & & 0\\
\hat{c}_{30} & 0 & -\hat{c}_{32} & 0 & \cdots & & 0\\
. & 0 & . & 0 & \cdots & & 0\\
. & 0 & . & 0 & \cdots & & 0\\
. & 0 & . & 0 & \cdots & & 0\\
\hat{c}_{n0} & 0 & -\hat{c}_{n2} & 0 & \cdots & & 0\\
. & 0 & 0 & . & \cdots & & .\\
. & . & . & . & \cdots & & .\\
. & . & . & . & \cdots & & .\\
. & . & . & . & \cdots & & .\\
. & . & . & 0 & \cdots & & .\\
. & . & . & 0 & \cdots & & .\\
\hat{c}_{10} & 0 & 0 & 0 & \cdots & 0 & -\hat{c}_{1R}\\
\hat{c}_{20} & 0 & 0 & 0 & \cdots & 0 & -\hat{c}_{2R}\\
\hat{c}_{30} & 0 & 0 & 0 & \cdots & 0 & -\hat{c}_{3R}\\
 & 0 & 0 & 0 & \cdots & & .\\
 & 0 & 0 & 0 & \cdots & & .\\
 & 0 & 0 & 0 & \cdots & & .\\
\hat{c}_{n0} & 0 & 0 & 0 & \cdots & 0 & -\hat{c}_{nR}
\end{bmatrix}
\begin{pmatrix}
\pi_0\\
\pi_1\\
\pi_2\\
\cdot\\
\cdot\\
\cdot\\
\cdot\\
\cdot\\
\cdot\\
\pi_R
\end{pmatrix}
+
\begin{pmatrix}
e_{11}\\
e_{21}\\
e_{31}\\
\cdot\\
\cdot\\
\cdot\\
e_{n1}\\
e_{12}\\
e_{22}\\
e_{32}\\
\cdot\\
\cdot\\
\cdot\\
e_{n2}\\
\cdot\\
\vdots\\
\cdot\\
\cdot\\
\cdot\\
e_{1R}\\
e_{2R}\\
e_{3R}\\
\cdot\\
\cdot\\
\cdot\\
e_{nR}
\end{pmatrix}
$$

$$\Rightarrow B_{nR\times 1} = 2 * C_{nR\times(R+1)}\,\pi_{(R+1)\times 1} + \boldsymbol{e}_{nR\times 1} \qquad \text{(A2.1.1)}$$

$$E_B = \begin{pmatrix} (b_{11} - \hat{b}_{11}) - (b_{10} - \hat{b}_{10}) \\ (b_{21} - \hat{b}_{21}) - (b_{20} - \hat{b}_{20}) \\ (b_{31} - \hat{b}_{31}) - (b_{30} - \hat{b}_{30}) \\ \vdots \\ (b_{n1} - \hat{b}_{n1}) - (b_{n0} - \hat{b}_{n0}) \\ (b_{12} - \hat{b}_{12}) - (b_{10} - \hat{b}_{10}) \\ (b_{22} - \hat{b}_{22}) - (b_{20} - \hat{b}_{20}) \\ (b_{32} - \hat{b}_{32}) - (b_{30} - \hat{b}_{30}) \\ \vdots \\ (b_{n2} - \hat{b}_{n2}) - (b_{n0} - \hat{b}_{n0}) \\ \vdots \\ (b_{1R} - \hat{b}_{1R}) - (b_{10} - \hat{b}_{10}) \\ (b_{2R} - \hat{b}_{2R}) - (b_{20} - \hat{b}_{20}) \\ (b_{3R} - \hat{b}_{3R}) - (b_{30} - \hat{b}_{30}) \\ \vdots \\ (b_{nR} - \hat{b}_{nR}) - (b_{n0} - \hat{b}_{n0}) \end{pmatrix}$$

E_B consists of elements which are of the form: $(b_{ij} - \hat{b}_{ij}) - (b_{i0} - \hat{b}_{i0})$. Now, each term in bracket contains two terms, the actual value of the unknown parameter and the estimated value obtained from the first stage regression. Since the parameters in the first stage regression are consistently estimated, E_B will tend to a null vector. Again, the first column of E_C (given below) will contain elements which are of the form$(c_{i0} - \hat{c}_{i0})$. The non zero entries in the j^{th} column ($j \neq 1$) will contain elements which are of the form $\{-(c_{ij} - \hat{c}_{ij})\}$. By similar argument as above, E_C will tend to a null matrix.

$$
\mathrm{E_C} = \begin{bmatrix}
(c_{10} - \hat{c}_{10}) & -(c_{11} - \hat{c}_{11}) & 0 & 0 & . & . & . & . & 0 \\
(c_{20} - \hat{c}_{20}) & -(c_{21} - \hat{c}_{21}) & 0 & 0 & . & . & . & . & 0 \\
(c_{30} - \hat{c}_{30}) & -(c_{31} - \hat{c}_{31}) & 0 & 0 & . & . & . & . & 0 \\
. & . & 0 & 0 & . & . & . & . & 0 \\
. & . & 0 & 0 & . & . & . & . & 0 \\
. & . & 0 & 0 & . & . & . & . & 0 \\
(c_{\mathrm{n}0} - \hat{c}_{\mathrm{n}0}) & -(c_{\mathrm{n}1} - \hat{c}_{\mathrm{n}1}) & 0 & 0 & . & . & . & . & 0 \\
(c_{10} - \hat{c}_{10}) & 0 & -(c_{12} - \hat{c}_{12}) & 0 & . & . & . & . & 0 \\
(c_{20} - \hat{c}_{20}) & 0 & -(c_{22} - \hat{c}_{22}) & 0 & . & . & . & . & 0 \\
(c_{30} - \hat{c}_{30}) & 0 & -(c_{32} - \hat{c}_{32}) & 0 & . & . & . & . & 0 \\
. & 0 & . & 0 & . & . & . & . & 0 \\
. & 0 & . & 0 & . & . & . & . & 0 \\
. & 0 & . & 0 & . & . & . & . & 0 \\
(c_{\mathrm{n}0} - \hat{c}_{\mathrm{n}0}) & 0 & -(c_{\mathrm{n}2} - \hat{c}_{\mathrm{n}2}) & 0 & . & . & . & . & 0 \\
. & 0 & 0 & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
. & . & . & 0 & . & . & . & . & . \\
. & . & . & 0 & . & . & . & . & . \\
(c_{10} - \hat{c}_{10}) & 0 & 0 & 0 & . & . & . & 0 & -(c_{1\mathrm{R}} - \hat{c}_{1\mathrm{R}}) \\
(c_{20} - \hat{c}_{20}) & 0 & 0 & 0 & . & . & . & 0 & -(c_{2\mathrm{R}} - \hat{c}_{2\mathrm{R}}) \\
(c_{30} - \hat{c}_{30}) & 0 & 0 & 0 & . & . & . & 0 & -(c_{3\mathrm{R}} - \hat{c}_{3\mathrm{R}}) \\
 & 0 & 0 & 0 & . & . & . & . & . \\
 & 0 & 0 & 0 & . & . & . & . & . \\
 & 0 & 0 & 0 & . & . & . & . & . \\
(c_{\mathrm{n}0} - \hat{c}_{\mathrm{n}0}) & 0 & 0 & 0 & . & . & . & 0 & -(c_{\mathrm{nR}} - \hat{c}_{\mathrm{nR}})
\end{bmatrix}
$$

Thus,

$E_C^{'}\Sigma^{-1}E_C$ will tend to a null matrix and $E_C^{'}\Sigma^{-1}E_B$ will tend to a null vector.

For the reference utility level u_0^h equation (2.6) can be written as

$$C(u_0^h, p^r) = a(p^r)\exp\left(\frac{b(p^r)}{\frac{1}{\ln u_0^h} - \lambda(p^r)}\right)$$

or,

$$y_{rh} = a(p^r)\exp\left(\frac{b(p^r)}{\frac{1}{\ln u_0^h} - \lambda(p^r)}\right),$$

where y_{rh} denotes the per capita nominal income required by a household of region *r* to have u_0^h utility level.

We thus have

$$ln\frac{y_{rh}}{a(p^r)} = \frac{b(p^r)}{\left(\frac{1}{\ln u_0^h}\right) - \lambda(p^r)}$$

or

$$\frac{1}{ln\frac{y_{rh}}{a(p^r)}} = \frac{1}{b(p^r)}\left(\frac{1}{lnu_0^h}\right) - \frac{\lambda(p^r)}{b(p^r)}$$

Now, substituting for $\frac{1}{lnu_0^h}$ from equation (2.12) in the above equation we get equation (2.13).

Let $\hat{\pi}_r = \pi_r + \delta_r$, say, for *r =0,1,….,R,* where δ_r is the error of estimation.

Then,

$$\frac{1}{\ln \overline{y}_{rq} - \hat{\pi}_r} = \frac{1}{\ln \overline{y}_{rq} - \pi_r - \delta_r}$$

$$= \left(\frac{1}{\ln \overline{y}_{rq} - \pi_r}\right)\left(\frac{\ln \overline{y}_{rq} - \pi_r}{\ln \overline{y}_{rq} - \pi_r - \delta_r}\right) = \frac{1}{\ln \overline{y}_{rq} - \pi_r} + \frac{\delta_r}{\left(\ln \overline{y}_{rq} - \pi_r\right)\left(\ln \overline{y}_{rq} - \pi_r - \delta_r\right)}$$

$$= \frac{1}{\ln \overline{y}_{rq} - \pi_r} + \delta_r^*, \text{ say,}$$

Therefore, equation (2.13) becomes

$$\frac{1}{\ln \overline{y}_{rq} - \pi_r} + \delta_r^* = \frac{1}{b(p^r)}\left(\frac{1}{\ln \overline{y}_{0q} - \hat{\pi}_0} + \delta_0^* + 1\right) - \frac{\lambda(p^r)}{b(p^r)}$$

or,

$$\frac{1}{\ln \overline{y}_{rq} - \hat{\pi}_r} = \frac{1}{b(p^r)}\left(\frac{1}{\ln \overline{y}_{0q} - \hat{\pi}_0} + 1\right) - \frac{\lambda(p^r)}{b(p^r)} + \left(\frac{\delta_0^*}{b(p^r)} - \delta_r^*\right),$$

which can be written in the form of equation (2.14) as

$$\frac{1}{\ln \overline{y}_{rq} - \hat{\pi}_r} = \frac{1}{b(p^r)}\left(\frac{1}{\ln \overline{y}_{0q} - \hat{\pi}_0} + 1\right) - \frac{\lambda(p^r)}{b(p^r)} + \text{error}$$

This again gives rise to the issue that regression error is present only because of estimation errors in the second stage, where the *error* term involves δ_0^* and δ_r^*. However, in the absence of a linear association between the error term and the regressor, OLS has been used to estimate $\frac{1}{b(p^r)}$ and $\frac{\lambda(p^r)}{b(p^r)}$.

This issue needs to be further explored possibly using a Monte Carlo study.

Delta method approximates the variance of any parameter that is a function of one or more random variables, each with its own estimate of variance.

Suppose *G* is the parameter, which is a function of the random variables: $(X_1, X_2, \ldots., X_N)$; i.e., $= f(X_1, X_2, \ldots., X_N)$, the variance of *G* is given by

$$var(G) = \sum_{i=1}^{n} var(X_i) \left[\frac{\partial f}{\partial X_i}\right]^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_i, X_j)\left[\left(\frac{\partial f}{\partial X_i}\right)\left(\frac{\partial f}{\partial X_j}\right)\right],$$

where $\frac{\partial f}{\partial X_i}$ is the partial derivative of *G* with respect to X_i.

Appendix A2.6 Table A2.1 Showing List of Items

1. Cereals	9. Sugar	15. Rent etc
2. Pulses and products	10. Salt and spices	16. Medical expense
3. Milk	11. Beverages etc	17. Educational Expense
4. Milk products	12. Betel leaf, tobacco, intoxicants	18. Clothing
5. Edible oils	13. Fuel and light	19. Footwear
6. Vegetables	14. Miscellaneous goods and services	20. Durable goods
7. Fruits and nuts		
8. Meat, egg and fish		

Appendix A2.7 Table A2.2 NSS Regions and Districts of West Bengal (2004-2005)

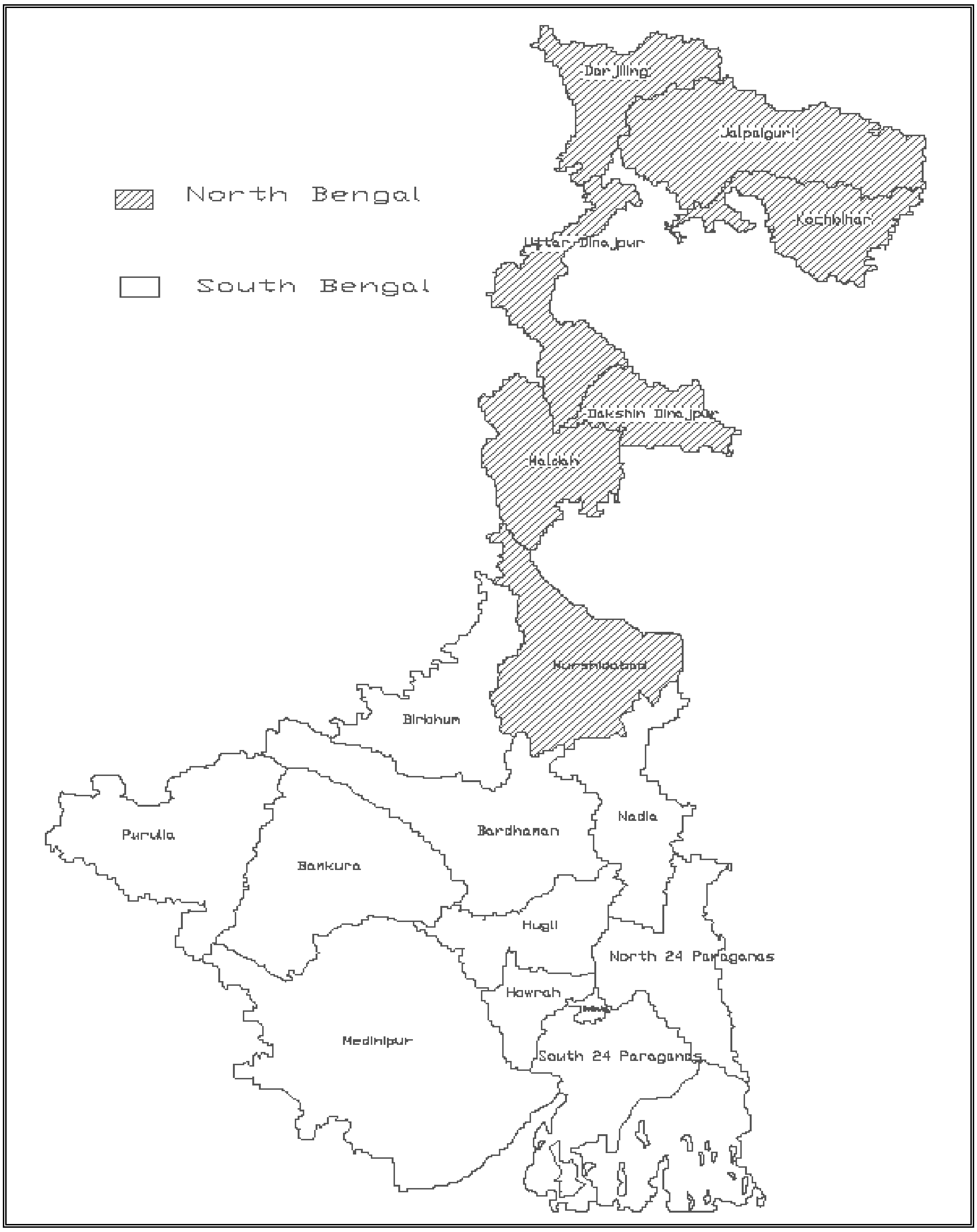
Districts	District Code	NSS Regions of West Bengal	
		Region Code	Name of Region
Darjiling	1	1	Himalayan
Jalpaiguri	2	1	
Kochbihar	3	1	
Uttar dinajpur	4	2	Eastern Plains
Dakshin dinajpur	5	2	
Maldah	6	2	
Murshidabad	7	2	
Birbhum	8	2	
Bardhaman	9	3	Central Plains
Nadia	10	2	Eastern Plains
North 24 Paraganas	11	3	Central Plains
Hugli	12	3	
Bankura	13	4	Western Plains
Purulia	14	4	
Medinipur	15	4	
Howrah	16	3	Central Plains
South 24 Paraganas	18	3	

Appendix A2.8 Table A2.3 Estimates of Parameters of Equation (2.14)

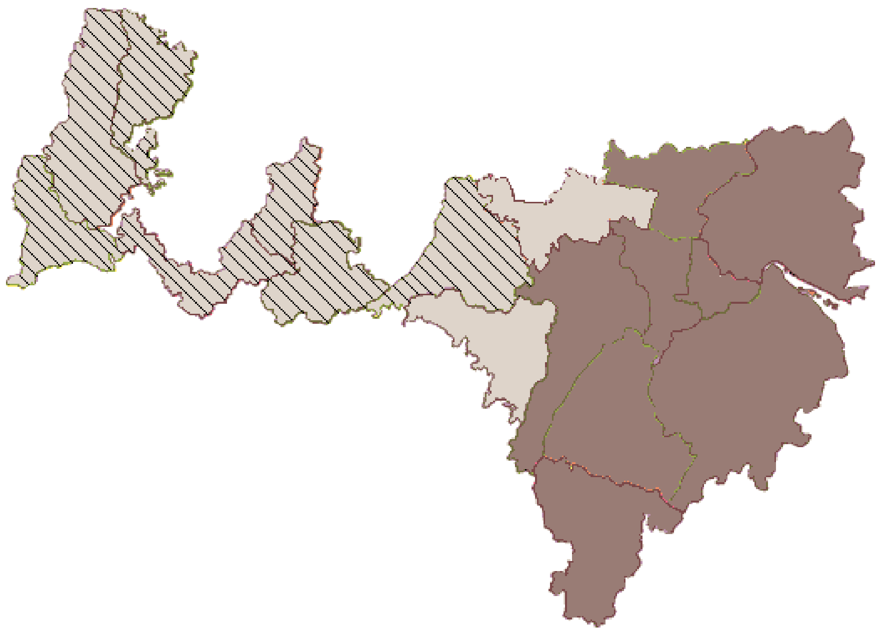
	Districts Taken as Regions			NSS Regions Taken as Regions		
District	π_r	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$	π_r	$\frac{1}{b(p^r)}$	$\frac{\lambda(p^r)}{b(p^r)}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Darjiling	7.64 (0.04)**	1.33 (0.03)	1.12 (0.02)	8.42 (0.13)	0.33 (0.02)	0.59 (0.01)
Jalpaiguri	7.61 (0.10)	0.60 (0.03)	0.88 (0.02)	8.42 (0.13)	0.33 (0.02)	0.59 (0.01)
Kochbihar	7.76 (0.10)	0.35 (0.03)	0.79 (0.02)	8.42 (0.13)	0.33 (0.02)	0.59 (0.01)
Uttar dinajpur	7.58 (0.13)	0.67 (0.01)	0.84 (0.01)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
Dakshin dinajpur	7.47 (0.06)	-11.29 (1.08)	-2.64 (0.64)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
Maldah	7.61 (0.10)	0.73 (0.01)	0.88 (0.01)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
Murshidabad	7.63 (0.07)	1.24 (0.05)	0.99 (0.03)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
Birbhum	7.60 (0.03)	0.35 (0.03)	0.77 (0.02)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
Bardhaman	7.74 (0.07)	0.76 (0.02)	0.95 (0.01)	7.98 (0.09)	1.06 (0.01)	1.05 (0.002)
Nadia	7.67 (0.08)	0.99 (0.02)	1.04 (0.01)	7.83 (0.08)	1.04 (0.01)	1.02 (0.004)
North 24 Paraganas	7.70 (0.11)	1.22 (0.01)	1.11 (0.01)	7.98 (0.09)	1.06 (0.01)	1.05 (0.002)
Hugli	7.69 (0.14)	3.36 (0.11)	1.90 (0.06)	7.98 (0.09)	1.06 (0.01)	1.05 (0.002)
Bankura	7.80 (0.08)	0.96 (0.01)	0.95 (0.01)	8.07 (0.07)	0.83 (0.01)	0.90 (0.005)
Purulia	7.41 (0.10)	0.42 (0.03)	0.83 (0.01)	8.07 (0.07)	0.83 (0.01)	0.90 (0.005)
Medinipur	7.85 (0.07)	0.72 (0.02)	0.90 (0.01)	8.07 (0.07)	0.83 (0.01)	0.90 (0.005)
Howrah	7.86 (0.07)	0.48 (0.02)	0.79 (0.01)	7.98 (0.09)	1.06 (0.01)	1.05 (0.002)
South 24 Paraganas	7.82 (0.07)	1.06 (0.02)	0.98 (0.01)	7.98 (0.09)	1.06 (0.01)	1.05 (0.002)

** Figures in parentheses are the standard errors.

Appendix A2.9 Figure A2.1 Map of West Bengal



Appendix A2.10 Cost of Living Variations Across Districts of West Bengal



FigureA2.2
Estimated Price Indices Constructed on the Basis of NSS Regions

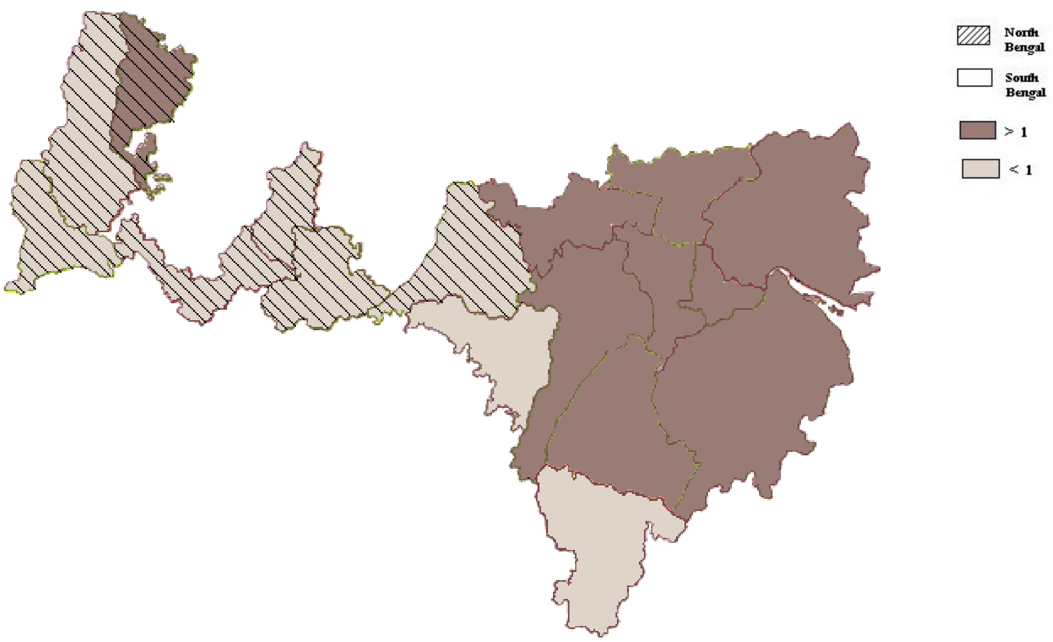
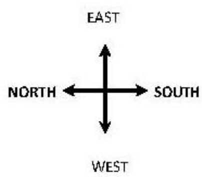


Figure A2.3
Estimated Price indices Constructed Taking Districts as Regions.



CHAPTER 3

COMPARISON OF POVERTY BETWEEN NORTH BENGAL AND SOUTH BENGAL

3.1 Introduction

There is considerable difference in the levels of economic well being in the two parts of Bengal, viz., North and South Bengal, and there is a debate in terms of the inequitable distribution of aid and welfare measures between these two parts[1]. Proper assessment of regional disparities in the levels of economic well-being is essential for designing welfare measures and for prioritization of policy measures with a view to lowering the disparities. For this, one needs to identify the sources and characteristics affecting economic well-being (deprivation).

This chapter aims to explore the causes of the differential levels of economic well being in the two parts of Bengal in terms of incidences of poverty and various socio economic explanatory variables. It also attempts to interpret the results in terms of policy implication. The analysis has been done in terms of nominal income (total expenditure) at the household level.[2]

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, Jong, Mwabu, & Kimenyi, 2001); (Bokosi, 2007)) as well as with reference to India (see e.g., (Gang, Sen, & Yun, 2008), (Bigman & Srinivasan, 2002)). Using a slightly different approach, a World Bank study (World Bank, 2003) considers the logarithm of the ratio of income to poverty line as the dependent variable (which is a common way of allowing for the lognormality of the variable) instead of a binary dependent variable in the logit/probit regression.[3] 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.

---

[1] See (Barma, 2007) ; (Ganguly, 2005).

[2] For comparability of results across districts, the issue of spatial price variation discussed in Chapter 2 has been omitted here and also in the subsequent chapters. Taking into account differential price levels may, however, alter the estimates of poverty incidence.

[3] See (Coudouel, Hentschel, & Wodon, 2002) for a discussion about the (World Bank) method *using linear regression in analyzing the determinants of poverty*. This method has been used also in (Bhaumik, Gang, & Yun, 2006) for studying difference in poverty incidences between Serbians and Albanians in Kosovo using Living Standard Measurement Survey. (Gang, Sen, & Yun, 2008) used this methodology to analyze the

In this chapter the incidence of poverty is analyzed, using the World Bank approach mentioned above, in the two parts of West Bengal. The disparity in poverty estimates (in particular, the Head Count Ratio (HCR) or (FGT0) between rural North and South Bengal is studied. The difference between the poverty estimates is then decomposed into a *characteristics effect*, showing the effect of the regional characteristics and a *coefficients effect*, showing the effects of the differential impact of the characteristics over the regions using the Oaxaca decomposition method (Oaxaca, 1973).

The plan of the chapter is as follows: Section 3.2 describes the regression based analysis of poverty and the Oaxaca decomposition methodology; Section 3.3 presents the data and results and finally Section 3.4 concludes. Appendices A3.1-A3.6 at the end of this chapter present detailed derivation of some results.

3.2 Regression Based Estimation of Poverty and Oaxaca Decomposition Methodology

The logarithm of the ratio of income to poverty line is regressed on a set of socio-economic factors (*poverty correlates*) and from the parameter estimates the probability of poverty incidence is obtained for each household. Poverty incidence for a region is then obtained as the sample average of household level probabilities of poverty incidence.

Following (Bhaumik, Gang, & Yun, 2006), the model is specified as:

$$\left(\frac{y}{z}\right)_i^* = X_i\beta + \varepsilon_i \,; \quad i = 1, 2, \ldots\ldots, n \tag{3.1}$$

where $\left(\frac{y}{z}\right)_i^* = ln\left(\frac{y}{z}\right)_i$; y is the household per capita total consumption expenditure, z is the poverty line. The subscript i denotes the i^{th} household. X is a vector of socio-economic variables influencing consumption.

The i^{th} household will be poor if its per capita total consumption expenditure is less than the poverty line. That is, the probability of the incidence of poverty of the i^{th} household,

$$p_i = \text{prob}\left(\left(\frac{y}{z}\right)_i < 1\right) = \text{prob}\left(\left(\frac{y}{z}\right)_i^* < 0\right) = \text{prob}\left(X_i\beta + \varepsilon_i < 0\right) \,; \qquad \text{[from (3.1)]}$$

$$= \text{prob}\left(\varepsilon_i < -X_i\beta\right).$$

That is, $p_i = \Phi\left(\frac{-X_i\beta - E(\varepsilon_i)}{\sqrt{var(\varepsilon_i)}}\right)$; [Φ is the C.D.F of standard normal distribution]

---

determinants of rural poverty in India, contrasting the situation of scheduled caste (SC) and scheduled tribe (ST) households with the general population.

$= \Phi\left(\frac{-X_i\beta}{\sigma}\right)$; [Assuming $var(\varepsilon_i) = \sigma^2$]

$= \Phi(X_i\beta^*)$. ; $\left[\beta^* = -\frac{\beta}{\sigma}\right]$ (3.2)

Now consider two regions, A and B. The FGT0 measure, i.e., the proportion of households below the poverty line, for any region is asymptotically equal to the sample average of the poverty incidences[4]. Therefore, the FGT0 measure for a region A will be

$H_A = \frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(X_{i_A}\widehat{\beta_A}^*\right)$; n^A being the number of households in region A. (3.3)

The FGT0 measure for region B will be

$H_B = \frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}\widehat{\beta_B}^*\right)$; n^B being the number of households in region B. (3.4)

The difference of poverty estimates between the regions A and B may thus be written as:

$H_A - H_B$

$= \left[\frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(X_{i_A}\widehat{\beta_A}^*\right)\right] - \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}\widehat{\beta_B}^*\right)\right]$

$= \overline{\Phi\left(X_A\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}$; (the over bar denotes sample average)

$= \left\{\overline{\Phi\left(X_A\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_A}^*\right)}\right\} + \left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}$ (3.5)

The first part in bracket is the *aggregate characteristics effect*, *C*, say, and the other part in bracket is the *aggregate coefficients effect*, *D*, say. In other words, in the difference of poverty, *C* explains the portion that is due to the difference in the *characteristics* (*X*'s), given the coefficients (β's) and *D* explains the portion that is due to the difference in the *coefficients*, given the characteristics.

The decomposition is done from the viewpoint of Region B in the sense that $\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} = \frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}\widehat{\beta_A}^*\right)$ is actually the counter factual poverty in Region B, i.e., the poverty level that would prevail in Region B if it would have the same Coefficient vector (β^*) as in A.

The *aggregate characteristics effect*, *C*, is thus the difference of the actual level of poverty at Region A (= H_A) and the counter factual poverty of Region B (= H_B^C) with Region A's Coefficient vector. The *aggregate coefficients effect*, D, is the difference of the counter factual poverty of Region B (= H_B^C) with Region A's Coefficient vector and the actual level of poverty in Region B (= H_B).

[4] See (Bhaumik, Gang, & Yun, 2006).

Now, evaluating the function $\overline{\Phi(.)}$'s at the mean values, the first part in (3.5), i.e., C, can be written as:

$$C = \left\{\overline{\Phi(X_A\widehat{\beta_A}^*)} - \overline{\Phi(X_B\widehat{\beta_A}^*)}\right\} \quad (3.5.1)$$

$=\left[\left\{\Phi(\overline{X_A}\widehat{\beta_A}^*) + R_{M_1}\right\} - \left\{\Phi(\overline{X_B}\widehat{\beta_A}^*) + R_{M_2}\right\}\right]$, R_{M_i}'s being approximation residuals.

$$=\left\{\overline{(X_A} - \overline{X_B)}\,\widehat{\beta_A}^*\right\}\phi(\overline{X_A}\widehat{\beta_A}^*) + \left(R_{T_1} + R_{M_1} - R_{M_2}\right) \quad (3.6)$$

[R_{T_1} is the approximation residual resulting from evaluating the difference of the function $\Phi(\cdot)$'s by using the first order Taylor expansion around $\overline{X_A}\widehat{\beta_A}^*$, ϕ being the first derivative of the function Φ, and hence the P.D.F of standard normal distribution (Yun, 2004)].[5]

Thus,

$$C = \left\{(\overline{X}_A - \overline{X}_B)\,\widehat{\beta_A}^*\right\} \times \phi(\overline{X_A}\widehat{\beta_A}^*) + \left(R_{T_1} + R_{M_1} - R_{M_2}\right) \quad (3.7)$$

$\overline{X_A}$ and $\overline{X_B}$ are vectors representing average values of the explanatory variables in regions A and B, respectively.

$$= C^1 + C_R \text{ , say,} \quad (3.7.1)$$

where $C^1 = \left\{\overline{(X_A} - \overline{X_B)}\,\widehat{\beta_A}^*\right\} \times \phi(\overline{X_A}\widehat{\beta_A}^*)$ and $C_R = \left(R_{T_1} + R_{M_1} - R_{M_2}\right)$.

Again, evaluating the function $\overline{\Phi(.)}$'s at the mean values, the *coefficients effect*, D can be written as:

$$D = \left\{\overline{\Phi(X_B\widehat{\beta_A}^*)} - \overline{\Phi(X_B\widehat{\beta_B}^*)}\right\} \quad (3.5.2)$$

$=\left[\left\{\Phi(\overline{X_B}\widehat{\beta_A}^*) + R'_{M_1}\right\} - \left\{\Phi(\overline{X_B}\widehat{\beta_B}^*) + R'_{M_2}\right\}\right]$, R'_{M_i}'s being approximation residuals.

$$= \overline{X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) \times \phi\left(\overline{X_B}\widehat{\beta_B}^*\right) + \left(R_{T_2} + R'_{M_1} - R'_{M_2}\right) \quad (3.8)$$

[R_{T_2} is the approximation residual resulting from evaluating the difference of the function $\Phi(\cdot)$'s by using the first order Taylor expansion around $\overline{X_B}\widehat{\beta_B}^*$ (Yun, 2004)][6]

$$= \overline{X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) \times \phi\left(\overline{X_B}\widehat{\beta_B}^*\right) + \left(R_{T_2} + R'_{M_1} - R'_{M_2}\right) \quad (3.9)$$

$$= D^1 + D_R \text{ , say;} \quad (3.9.1)$$

where $D^1 = \overline{X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) \times \phi\left(\overline{X_B}\widehat{\beta_B}^*\right)$; $D_R = \left(R_{T_2} + R'_{M_1} - R'_{M_2}\right)$.

[5] $R_{M_1} = \overline{\Phi\left(X_A\widehat{\beta_A}^*\right)} - \Phi\left(\overline{X_A}\widehat{\beta_A}^*\right)$; $R_{M_2} = \overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \Phi\left(\overline{X_B}\widehat{\beta_A}^*\right)$;
$R_{T_1} = \left\{\Phi(\overline{X_A}\widehat{\beta_A}^*) - \Phi(\overline{X_B}\widehat{\beta_A}^*)\right\} - \left\{\overline{(X_A} - \overline{X_B)}\,\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A}\widehat{\beta_A}^*\right)$

[6] $R'_{M_1} = \overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \Phi\left(\overline{X_B}\widehat{\beta_A}^*\right)$; $R'_{M_2} = \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)} - \Phi\left(\overline{X_B}\widehat{\beta_B}^*\right)$
$R_{T_2}{}' = \left\{\Phi\left(\overline{X_B}\widehat{\beta_A}^*\right) - \Phi(\overline{X_B}\widehat{\beta_B}^*)\right\} - \overline{X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)$

Detailed Decomposition Analysis:

The *aggregate characteristics effect, C* and the *aggregate coefficients effect, D* contain the effects of all the explanatory variables. The contribution of specific factors can be factored out from the overall contribution as follows.

Following (Yun, 2004), the weight (the share of the particular variable in the *aggregate characteristics effect*) of the k^{th} explanatory variable as derived from (3.7) is:

$$V_{\Delta X}^{k} = \frac{\left\{(\bar{X}_A^k - \bar{X}_B^k)\, {\beta^*}_A^k\right\} \times \phi(\overline{X_A \widehat{\beta_A}^*})}{\left\{(\bar{X}_A - \bar{X}_B)\, {\beta^*}_A\right\} \times \phi(\overline{X_A \widehat{\beta_A}^*})} = \frac{\left\{(\bar{X}_A^k - \bar{X}_B^k)\, {\beta^*}_A^k\right\} \times \phi(\overline{X_A \widehat{\beta_A}^*})}{C^1} \qquad (3.10)$$

The characteristic effect due to the k^{th} explanatory variable is thus,

$$C_k = V_{\Delta X}^{k} \times C, \quad \text{i.e., } C = \textstyle\sum_{k=1}^{K} C_k \qquad (3.11)$$

Again, the weight (the share of the particular variable in the *aggregate coefficients effect*) of the k^{th} explanatory variable as derived from (3.9) is:

$$\overset{k}{V}_{\Delta\beta} = \frac{\left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\} \times \phi(\overline{X_B \widehat{\beta_B}^*})}{\left\{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right\} \times \phi(\overline{X_B \widehat{\beta_B}^*})} = \frac{\left\{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right\}}{D^1} \qquad (3.12)$$

The coefficient effect due to the k^{th} explanatory variable is thus

$$D_k = \overset{k}{V}_{\Delta\beta} \times D, \quad \text{i.e., } D = \textstyle\sum_{k=1}^{K} D_k \qquad (3.13)$$

Since the weights considered here are *shares* of respective variables in the *aggregate characteristics* and *aggregate coefficients effect*, for K explanatory variables the following relationships hold:

$$\textstyle\sum_{k=1}^{K} V_{\Delta X}^{k} = 1 \qquad (3.14)$$

$$\textstyle\sum_{k=1}^{K} \overset{k}{V}_{\Delta\beta} = 1 \qquad (3.15)$$

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 above relationships as:

$$H_A - H_B = \left\{\textstyle\sum_{k=1}^{K} V_{\Delta X}^{k}\left[\overline{\Phi(X_A \widehat{\beta_A}^*)} - \overline{\Phi(X_B \widehat{\beta_A}^*)}\right] + \textstyle\sum_{k=1}^{K} V_{\Delta\beta}^{k}\left[\overline{\Phi(X_B \widehat{\beta_A}^*)} - \overline{\Phi(X_B \widehat{\beta_B}^*)}\right]\right\} \qquad (3.16)$$

To test for the significance of the aggregate *characteristics effect* and the aggregate *coefficients effect* the respective variances are derived as follows.

The Variances of Aggregate Characteristics Effect (C) and Aggregate Coefficients Effect (D):

From the estimated variance–covariance structure of the coefficients of model in (3.1), the variances of aggregate *characteristics* and aggregate *coefficients effect*s are obtained using the delta method[7].

From (3.7), the aggregate *characteristics effect*, $C = f(\widehat{\beta_A}^*)$.

Thus, the *asymptotic variance of aggregate characteristics effect*

$$\sigma_C^2 = \left(\frac{\partial C}{\partial \widehat{\beta_A}^*}\right) Asy\ var\ ({\beta_A}^*)\ \left(\frac{\partial C}{\partial \widehat{\beta_A}^*}\right)^T \tag{3.17}$$

Again, from (3.9), the aggregate coefficients effect, D = $f(\widehat{\beta_A}^*, \widehat{\beta_B}^*)$. Thus, the asymptotic variance of aggregate coefficient effect,

$$\sigma_D^2 = \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right) (Asy\ var\ ({\beta_A}^*)) \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial D}{\partial \widehat{\beta_B}^*}\right) (Asy\ var({\beta_B}^*)) \left(\frac{\partial D}{\partial \widehat{\beta_B}^*}\right)^T \tag{3.18}$$

Now, the asymptotic variance of β^* is approximated from the variance-covariance structure of β as obtained from the estimation of (3.1) using the Delta method and σ_C^2 and σ_D^2 are obtained from (3.17) and (3.18) respectively.[8]

The Variances of Specific Characteristic Effect (C_k) and Specific Coefficient Effect (D_k):

From (3.11), the characteristic effect by the k[th] explanatory variable is $C_k = V_{\Delta X}^k \times C$

i.e., $= f\left(\widehat{\beta_A}^*\right)$, using (3.10) and (3.5.1)

The asymptotic variance of C_k is thus,

$$\sigma_{C_K}^2 = \left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right) Asy\ var\ ({\beta_A}^*)\ \left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right)^T. \tag{3.19}$$

Again from (3.13), $D_k = V_{\Delta\beta}^{k} \times D = f\left(\widehat{\beta_A}^*, \widehat{\beta_B}^*\right)$, from (3.12) and (3.5.2)

The asymptotic variance of D_k,

$$\sigma_{D_K}^2 = \left(\frac{\partial D_K}{\partial \widehat{\beta_A}^*}\right) (Asy\ var\ ({\beta_A}^*)) \left(\frac{\partial D_K}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial D_K}{\partial \widehat{\beta_B}^*}\right) (Asy\ var({\beta_B}^*)) \left(\frac{\partial D_K}{\partial \widehat{\beta_B}^*}\right)^T ; \quad (3.20);$$ [9]

[7] See (Bhaumik, Gang, & Yun, 2006), (Yun, 2005)

[8] See Appendix A3.1 for derivation of Asymptotic Variance $({\beta_A}^*)$ and Asymptotic Variance $({\beta_B}^*)$; Appendix A3.2 for the exact form of $\left(\frac{\partial C}{\partial \widehat{\beta_A}^*}\right)$ and Appendix A3.3 for the exact form of $\left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right)$ and $\left(\frac{\partial D}{\partial \widehat{\beta_B}^*}\right)$.

[9] See Appendix A3.4 for the exact form of $\left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right)$ and Appendix A3.5 for the exact form of $\left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right)$ and $\left(\frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right)$.

Testing the Significance of C and D:

For testing the significance of the aggregate *characteristics effect* and the aggregate *coefficients effect* in explaining the difference in poverty estimates between North and South Bengal, the test statistic under the null hypothesis (C= 0, D= 0) are, respectively, given by $t_c = \frac{C}{\sigma_C}$ and $t_D = \frac{D}{\sigma_D}$, which are asymptotically normally distributed.

The significance of the effects at the individual variable level can be tested in a similar fashion as follows: the test statistic under the null hypothesis (C_k= 0, D_k= 0) are, respectively, given by $t_{C_k} = \frac{C_k}{\sigma_{C_k}}$ and $t_{D_k} = \frac{D_k}{\sigma_{D_k}}$, which are asymptotically normally distributed ((Bhaumik, Gang, & Yun, 2006), (Yun, 2005)).

3.3 Data and Results

The analysis in this chapter uses the household level NSS 61st round (2004-05) employment-unemployment data for the rural sector of West Bengal.[10]

The logarithm of ratio of per capita total consumption expenditure[11] to the poverty line, R $\left(= ln\left(\frac{y}{z}\right)\right)$ is the variable under study. Poverty line has been taken to be the official state level poverty line of Rs.382.82 per capita per month for rural West Bengal. The explanatory variables are broadly categorized as:

I. Demographic characteristics of the households.

II. Educational status.

III. Wealth status.

IV. Labour market characteristics.

V. Government aid.

The variables under these broad categories are:

I. Demographic characteristics of the households:

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

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

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

   D_FEMH = 1 if the family is female-headed

[10] The same data set has been used in the subsequent Chapters.

[11] Alternatively, one can take monthly per-equivalent adult consumption expenditure taking into account economies of scale.

= 0 otherwise

II. Educational status of the households

1. The proportion of members having secondary education, (PSECEDU).
2. The proportion of members having tertiary education, (PTERTEDU).
3. The average general educational level, (GENEDU).[12]

III. Wealth status of the households

1. Per Capita Amount of land possessed (measured in Hectares), (PLAND).

IV. Labour market characteristics of the households

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).

V. Government aid

1. Dummy variable, (D_GOVAID),

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

= 0 otherwise.

The variable *R* is regressed on these explanatory variables (equation (3.1) is estimated by maximum likelihood technique[13]), separately for North Bengal and South Bengal. The names and codes of the districts of North Bengal and South Bengal are given in Tables 3.1.1 and Table 3.1.2. The regression estimates are given in Table 3.2. All the coefficients turn out

[12] 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. Since codes/indicators increase with levels of education, these have been taken as proxy for years spent in education.

[13] Alternatively, (3.1) can be estimated using OLS as used by the (World Bank, 2003). The methodology in (Bhaumik, Gang, & Yun, 2006) is followed in this Chapter and (3.1) is estimated by MLE because the asymptotic variance covariance structure of the parameters along with the equation standard error can be obtained simultaneously.

to be positive except for the variable PNSCH[14]. All coefficients except for PSECEDU (for North Bengal) are significant at 5% level.

Taking North Bengal to be Region A and South Bengal to be Region B[15], the incidences of poverty have been estimated for these regions using equations (3.3) and (3.4) and the estimates of the parameters are presented in Table 3.2. The estimates of poverty H_A and H_B are given in Table 3.3. The value for North Bengal is 0.31 and that for South Bengal is 0.17.[16] The difference of the poverty estimates (= 0.14) between Region A and Region B is decomposed into an aggregate characteristics effect, C and an aggregate coefficients effect, D using equation (3.5). Table 3.4 presents the results relating to this decomposition. Both C (= 0.06) and D (= 0.08) turn out to be positive and highly significant.

As observed from the Table, the share[17] of the aggregate characteristics effect (C) in the difference in the incidences of poverty between A and B, i.e., (H_A- H_B), is 41% and the share of the aggregate coefficients effect (D) in (H_A- H_B) is 59%. This means that if the households of North Bengal had the same characteristics as those of South Bengal, given the North Bengal coefficients, the difference in the incidences of poverty, viz., the poverty gap, (H_A- H_B), would have been less by 41%. The poverty gap would have been less by 59% if the coefficients of the variables influencing poverty were same for both the parts, given the South Bengal characteristics.

The Individual Characteristic Effect (C_k's):

Using the relationships in (3.10) and (3.11), C is decomposed into contributions (C_k's) by individual explanatory variables. Coming to the contributions (C_k's) by individual

[14] This variable is expected to have a negative influence on household monthly per-capita expenditure. A possible explanation is that the income earned by joining the labour market is smaller than the gain in income made through increase in efficiency resulting from joining the educational institutions.

[15] The choice is determined by the fact that incidence of poverty is higher in Region A compared to that in Region B and we are interested in decomposing a positive poverty gap. From the decomposition it is clear that reversing the roles of A and B will not produce exactly symmetrically opposite results.

[16] It may be pointed out that these regression based estimates of poverty of North Bengal (0.31) and South Bengal (0.17) are quite close to the conventional direct estimates of 0.33 and 0.15, respectively.
A Chi-square test of independence to determine whether there is a significant relationship between the classification by being poor/non poor and by living in North/South Bengal produces a highly significant $\chi^2_{(1)}$ value of 121.5 indicating that these two classifications are not independent. This provides a further justification for looking at poverty in the two parts of West Bengal separately.
Figure A3.1 in Appendix A3.6 shows the Incidences of poverty across districts of North Bengal and South Bengal.

[17] The share of aggregate characteristics effect in (H_A- H_B) $= \frac{C}{H_A - H_B} \times 100$ and the share of aggregate coefficients effect in (H_A- H_B) $= \frac{D}{H_A - H_B} \times 100$.

explanatory variables in the aggregate characteristics effects C, which has a share of 41% in the poverty gap (H_A-H_B), Educational status has the highest contribution (C_k) with a share $\left(= \frac{C_k}{H_A - H_B} \times 100\right)$ of 27% in the difference in the incidences of poverty. This is followed by Demographic status with a share of 6.4%, Labour market status with a share of 4.5%, Government aid with a share of 2.3% and Wealth status with a share of 0.7%.

A positive value for an individual variable *k*, i.e., $C_k \left(= V_{\Delta X}^{k} \times C\right)$ means that $(\bar{X}_A^k - \bar{X}_B^k)$ and $\widehat{\beta^*}_A^k$ have the same sign.[18] For the k^{th} variable having a positive impact on consumption (and hence $\beta^* = -\frac{\beta}{\sigma}$ is negative), $(\bar{X}_A^k - \bar{X}_B^k)$ is negative. That is, the explanatory variable *k* has a lower average value in North Bengal (Region A) than in South Bengal (Region B). Similarly, for a variable that has a negative impact on consumption, the explanatory variable in Region A would have a higher average value in North Bengal than in South Bengal (for a positive C_k).

A negative value of C_k , on the other hand, implies that $(\bar{X}_A^k - \bar{X}_B^k)$ and $\widehat{\beta^*}_A^k$ have opposite signs. This (using similar arguments) would mean that for a variable that has a positive (negative) impact on consumption, the explanatory variable in Region A would have a higher (lower) average value in North Bengal than in South Bengal.

From Table (3.4) it may be observed that D_FEMH, PNSCH and PEMP have negative values of C_k (hence negative shares in (H_A-H_B)). Given that PNSCH had a negative coefficient and the other two had positive coefficients in the estimation of equation (3.1), North Bengal has higher average values for the variables D_FEMH and PEMP. All the individual shares $\left(= \frac{C_k}{H_A - H_B} \times 100\right)$ as well as the coefficients of the rest of the explanatory variables are positive and hence North Bengal has lower average values for these variables. In general, therefore, the average magnitudes of resources that explain consumption are lower in North Bengal compared to South Bengal.[19] Thus, the *characteristics effect* shows the differential degree of availability of resources given that the degree of utilization of resources is the same in the

[18] $C_k > 0 \Rightarrow V_{\Delta X}^{k} \times C > 0 \Rightarrow \frac{\left\{\left(\bar{X}_A^k - \bar{X}_B^k\right) \beta^{*k}_{A}\right\} \times \phi(\overline{X_A}\widehat{\beta_A}^*)}{C^1} \times C > 0 \Rightarrow \left\{(\bar{X}_A^k - \bar{X}_B^k)\, \beta^{*k}_{A}\right\} \times \phi(\overline{X_A}\widehat{\beta_A}^*) > 0$ [since C and C^1 have the same (positive) signs, as estimated in this Chapter.]

[19] Table 3.5 showing descriptive statistics of the explanatory variables of equation (3.1) corroborates this observation.

two regions. Hence, in a sense the *characteristics effect* can be interpreted as *'Resource effect'*.[20]

The Individual Coefficient Effects (D_k's):

Using the relationships in (3.12) and (3.13), D is decomposed into contributions (D_k's) by individual explanatory variables. In view of the above discussion, the coefficients of the variables in equation (3.1) can be interpreted as the degree of resource utilization in achieving certain level of consumption. Hence, the *coefficients effect* in the Oaxaca decomposition can be interpreted as the '*Efficiency effect'* which gives the differential degree of utilization of resources.

A positive (negative) contribution, D_k by a particular variable, which has a positive coefficient ($\hat{\beta}^k > 0$), would mean that Region A is having a lower coefficient attached to that particular variable compared to Region B.[21] This signifies that Region A is less (more) efficient than Region B with respect to utilization of that particular resource.

Coming to the contributions (D_k's) by individual explanatory variables in the aggregate *coefficients effect* (D), which has a share of 59 % in the poverty gap (H_A-H_B), Demographic status has the highest contribution (D_k) with a share $\left(= \frac{D_k}{H_A - H_B} \times 100\right)$ of 65 % in the difference in the incidences of poverty. This is followed by Educational status (29%), Wealth status (8%), Labour market status (7%) and Government aid (3%). All the individual D_k's turn out to be highly significant.

Except for Educational status, all the variables have negative contributions (D_k) and hence negative shares, meaning that North Bengal is having a higher coefficient associated with the respective variables (for which $\hat{\beta}^k > 0$) compared to South Bengal.[22] Negative shares thus indicate that equalization of the regional coefficients will make North Bengal worse off (Bhaumik,

---

[20] More will be elaborated on this in later chapters.

[21] $D_k > 0 \Rightarrow V^k_{\Delta\beta} \times D > 0 \Rightarrow \frac{\left\{\overline{X}^k_B\left(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B\right)\right\} \times \phi(\overline{X_B}\widehat{\beta_B}^*)}{D^1} \times D > 0$

$\Rightarrow \left\{\overline{X}^k_B\left(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B\right)\right\} \times \phi(\overline{X_B}\widehat{\beta_B}^*) > 0$; [since here D and D^1 have the same (positive) signs (as estimated).]

$\Rightarrow \frac{\hat{\beta}^k_A}{\sigma_A} - \frac{\hat{\beta}^k_B}{\sigma_B} < 0 \quad [\widehat{\beta^*}^k_A = -\frac{\hat{\beta}^k_A}{\sigma_A}; \widehat{\beta^*}^k_B = -\frac{\hat{\beta}^k_B}{\sigma_B}]$

$\Rightarrow \hat{\beta}^k_A - \hat{\beta}^k_B < 0$, as $\sigma_A < \sigma_B$ (the estimated values of σ_A, σ_B are 0.32 and 0.35 , respectively)

[22] This can also be checked from Table 3.2.

Gang, & Yun, 2006) because by increasing $\hat{\beta}^k$ in South Bengal, poverty will decrease in South Bengal and the poverty gap will be widened, because H_A-H_B is positive. For educational status, which has a positive share of coefficient effect, North Bengal is less efficient in utilizing this characteristic. This means that by increasing the coefficient attached to educational status in North Bengal to that in South Bengal, poverty gap can be removed by 29%. The fact that the coefficients effect of the constant term is positive with a share of 112% in the poverty gap (H_A-H_B)), indicates that the average baseline consumption expenditure level is lower in North Bengal.

3.4 Conclusion

The essence of the implementation of the Oaxaca decomposition methodology in the analysis is to determine the effect of the difference in the characteristics of the two regions that cause the regional difference in poverty incidences and to find out the effect of the differential impact of the characteristics over the two regions, so that policy measures can be formulated in terms of enhancement of either the characteristics or the impact of the characteristics over the regions.

The outcome of the above analysis is that specific policy measures can be identified for lowering the poverty gap between the two parts of West Bengal. As the results suggest, there is disparity in the availability of the characteristics (resources) as well as in utilization of resources (efficiency) in the two parts and the latter effect is found to be more prominent in terms of the share in explaining the poverty gap. While the baseline consumption is lower in North Bengal, in terms of both availability of resources and utilization of resources North Bengal lags behind South Bengal. Thus, attention needs to be paid to North Bengal with respect to enhancement of important policy variables like education level, Government aid and employment opportunities. Also, the causes of low resource utilization need to be investigated. The next chapter attempts to address the latter issue.

TABLES

Table 3.1.1 Districts of North Bengal (Region A)

District name	District code no.	NSS Region
Darjiling	1	Himalayan and most parts of Eastern Plains
Jalpaiguri	2	
Kochbihar	3	
Uttar dinajpur	4	
Dakshin dinajpur	5	
Maldah	6	
Murshidabad	7	

Table 3.1.2 Districts of South Bengal (Region B)*

District name	District code no.	NSS Region
Birbhum	8	Western and Central Plains (exception: Nadia: Eastern Plains)
Bardhaman	9	
Nadia	10	
North 24 Paraganas	11	
Hugli	12	
Bankura	13	
Purulia	14	
Medinipur	15	
Howrah	16	
South 24 Paraganas	18	

*As there is no rural sector in Kolkata, it has not been shown in Table 3.12.

Table 3.2 Estimates of the Parameters of Equation (3.1) for North Bengal and South Bengal

$(\boldsymbol{Dependent\ Variable}: \boldsymbol{ln(y/z)})$

Characteristics		Variables under characteristics	North Bengal (Region A)		South Bengal (Region B)	
			Estimate (β_A)	t-value	Estimate (β_B)	t-value
(1)		(2)	(3)	(4)	(5)	(6)
I	Demographic characteristics of the households	1-DEPRAT	0.3206	7.30*	0.1935	8.66*
		D_FEMH	0.0989	3.72*	0.0970	4.59*
II	Educational status of the household	PSECEDU	0.1162	1.49	0.2235	2.22*
		PTERTEDU	0.3946	5.20*	0.3018	7.41*
		GENEDU	0.0467	9.74*	0.0646	12.39*
III	Wealth status	PLAND	0.0005	8.37*	0.0004	10.90*
IV	Labour market characteristics	POWNAC	0.4322	7.28*	0.3625	8.64*
		PNSCH	-0.3425	-4.84*	-0.2628	-6.96*
		PDOMO	0.2402	4.03*	0.2729	5.02*
		PDOM	0.4599	6.38*	0.3285	8.35*
		PEMP	0.1223	2.22*	0.2339	2.88*
V	Government aid	D_GOVAID	0.2394	6.71*	0.2112	9.86*
	Constant		-0.3750	-12.93*	-0.2272	-14.99*

* indicates significance at 5% level.

Table 3.3 Estimates of Poverty in North and South Bengal (from Equations (3.3) and (3.4))

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

* Effective sample size after correction for missing observations.

Table 3.4 Decomposing the Difference of Poverty Incidences: (H_A-H_B) between North Bengal (Region A) and South Bengal (Region B)

(Using Estimates of Table 3.2)

Aggregate effect[***]			Aggregate Characteristics effect (C)			Aggregate Coefficients effect(D)		
			Estimate	Share in (H_A-H_B)		Estimate	Share in (H_A-H_B)	
			0.0570 (0.0030)*	**40.9**		**0.0824** (0.0085)*	**59.1**	
Decomposition of the Aggregate effect			Individual Characteristic effect (C_k)' s			Individual Coefficient effect (D_k)'s		
			Estimates	Shares in (H_A-H_B)**		Estimates	Shares in (H_A-H_B)	
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	Demographic characteristics of the households	1-DEPRAT	0.0106 (0.0015)	7.6	6.4	-0.0898 (0.0118)	-64.4	-65.1
		D_FEMH	-0.0016 (0.0001)	-1.2		-0.0010 (0.0003)	-0.7	
II	Educational status of the household	PSECEDU	0.0015 (0.0004)	1.1	27.1	0.0045 (0.0008)	3.3	29.2
		PTERTEDU	0.0059 (0.0014)	4.2		-0.0071 (0.0028)	-5.1	
		GENEDU	0.0304 (0.0005)	21.8		0.0432 (0.0013)	31.0	
III	Wealth status	PLAND	0.0010 (0.0000002)	0.7	0.7	-0.0107 (0.000004)	-7.7	-7.7
IV	Labour market characteristics	POWNAC	0.0001 (0.00001)	0.1	4.5	-0.0153 (0.0050)	-11.0	-6.6
		PNSCH	-0.0031 (0.0007)	-2.2		0.0063 (0.0023)	4.5	
		PDOMO	0.0045 (0.0008)	3.2		0.0007 (0.0034)	0.5	
		PDOM	0.0049 (0.0011)	3.5		-0.0134 (0.0041)	-9.6	
		PEMP	-0.0002 (0.00003)	-0.1		0.0126 (0.0015)	9.0	
V	Government aid	D_GOVAID	0.0031 (0.0004)	2.3	2.3	-0.0040 (0.0011)	-2.9	-2.9
	Constant		-	-		0.1564 (0.0132)	112.2	

* Figures in parentheses denote the standard errors. Both the aggregate characteristics effect and the aggregate coefficients effect are significant at 5% level. Also the individual characteristic and coefficient effects are highly significant at 5% level.

**share = estimates/ Difference in poverty incidence (Table 3.3)) × 100

*** Aggregate Characteristics effect, $C = \sum_{k=1}^{K} C_k$ and Aggregate Coefficients effect, $D = \sum_{k=1}^{K} D_k$

Table 3.5 Observed Resource Vectors for North Bengal and South Bengal

Characteristics		Variables under characteristics	$\bar{X}^k$	
			Region A (North Bengal)	Region B (South Bengal)
			$\bar{X}_A^k$	$\bar{X}_B^k$
(1)		(2)	(3)	(4)
I	Demographic characteristics of the households	1-DEPRAT	0.6274	0.6685
		D_FEMH	0.1094	0.0886
II	Educational status of the household	PSECEDU	0.0430	0.0594
		PTERTEDU	0.0440	0.0625
		GENEDU	3.4585	4.2696
III	Wealth status	PLAND	82.6713	85.1625
V	Labour market characteristics	POWNAC	0.1555	0.1558
		PNSCH	0.0529	0.0643
		PDOMO	0.1599	0.1832
		PDOM	0.0754	0.0887
		PEMP	0.1608	0.1590
V	Government aid	D_GOVAID	0.0708	0.0872

APPENDICES

Appendix A3.1 Estimation of Asymptotic Variance ($\boldsymbol{\beta_A}^*$) and Asymptotic Variance ($\boldsymbol{\beta_B}^*$)

$\beta^* = -\frac{\beta}{\sigma} = f\ (\beta, \sigma)$; [See equation (3.2)]

Following the Delta method,

Asymptotic Variance (β^*)

$= J_f(\beta)\ [Asy\ var\ (\beta)]\ J_f(\beta)^T$; J denoting the Jacobian matrix.

$= \frac{\partial\left(-\frac{\beta}{\sigma}\right)}{\partial\beta} [Asy\ var\ (\beta)]\ \frac{\partial\left(-\frac{\beta}{\sigma}\right)}{\partial\beta}^T$

$= \left[a_{ij}\ \right]_{K\times K} [Asy\ var\ (\beta)]\ \left[a_{ij}\ \right]^T_{K\times K}$

where $a_{ij} = -\frac{1}{\sigma}$ for $i = j$; (A3.1.1)

= 0, otherwise

Using (A3.1.1), the Asymptotic variance ($\beta_A{}^*$) and Asymptotic variance ($\beta_B{}^*$) can be estimated.

Appendix A3.2 To find $\frac{\partial C}{\partial \widehat{\beta_A}^*}$

From (3.5.1), $C = \left\{ \overline{\Phi(X_A \widehat{\beta_A}^*)} - \overline{\Phi(X_B \widehat{\beta_A}^*)} \right\}$

Writing in terms of individual observations,

$C = \left[\frac{1}{n^A} \sum_{i=1}^{n^A} \Phi\left(X_{i_A} \widehat{\beta_A}^*\right)\right] - \left[\frac{1}{n^B} \sum_{i=1}^{n^B} \Phi(X_{i_B} \widehat{\beta_A}^*)\right]$;

Thus,

$\frac{\partial C}{\partial \widehat{\beta_A}^*} = \vec{\nabla}\, C$, where $\vec{\nabla} C = \left(\frac{\partial C}{\partial \beta^{*1}_A}, \frac{\partial C}{\partial \beta^{*2}_A}, \ldots, \frac{\partial C}{\beta^{*K}_A} \right)$

$= \left[\left(\frac{1}{n^A} \sum_{i=1}^{n^A} X_{i_A}{}^k \phi\left(X_{i_A} \widehat{\beta_A}^*\right) - \frac{1}{n^B} \sum_{i=1}^{n^B} X_{i_B}{}^k \phi\left(X_{i_B} \widehat{\beta_A}^*\right) \right) \right]_{1 \times K}$

$= \left[\left(\overline{X_A{}^k \phi\left(X_A \widehat{\beta_A}^*\right)} - \overline{X_B{}^k \phi\left(X_B \widehat{\beta_A}^*\right)} \right) \right]_{1 \times K}$

Appendix A3.3 To find $\frac{\partial D}{\partial \widehat{\beta_A}^*}$ & $\frac{\partial D}{\partial \widehat{\beta_B}^*}$

From (3.5.2), $D = \left\{\overline{\Phi\,(X_B\widehat{\beta_A}^*)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}$

Writing in terms of individual observations,

$D = \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}\widehat{\beta_A}^*\right)\right] - \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}\widehat{\beta_B}^*\right)\right]$

Thus,

$\frac{\partial D}{\partial \widehat{\beta_A}^*} = \vec{\nabla}\,\mathrm{D}$, where $\vec{\nabla}\mathrm{D} = \left(\frac{\partial D}{\partial \beta^{*1}_A}, \frac{\partial D}{\partial \beta^{*2}_A}, \dots, \frac{\partial D}{\beta^{*K}_A}\right)$

$= \left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B} X_{i_B}{}^k \phi\left(X_{i_B}\widehat{\beta_A}^*\right)\right)\right]_{1\times K}$

$= \left[\left(\overline{X_B{}^k \phi\left(X_B\widehat{\beta_A}^*\right)}\right)\right]_{1\times K}$

Similarly,

$\frac{\partial D}{\partial \widehat{\beta_B}^*} = -\left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B} X_{i_B}{}^k \phi\left(X_{i_B}\widehat{\beta_B}^*\right)\right)\right]_{1\times K}$

$= -\left[\left(\overline{X_B{}^k \phi\left(X_B\widehat{\beta_B}^*\right)}\right)\right]_{1\times K}$

Appendix A3.4 To find $\left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right)$

$$C_k = \left\{\frac{(\bar{X}_A^k - \bar{X}_B^k)\, {\beta^*}_A^k}{(\bar{X}_A - \bar{X}_B)\, {\beta^*}_A}\right\}\left\{\overline{\Phi\,(X_A\widehat{\beta_A}^*)} - \overline{\Phi\,(X_B\widehat{\beta_A}^*)}\right\}$$

$$\left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right) = \left\{\overline{\Phi\,(X_A\widehat{\beta_A}^*)} - \overline{\Phi\,(X_B\widehat{\beta_A}^*)}\right\}\left[\frac{(\bar{X}_A - \bar{X}_B)\, {\beta^*}_A(\bar{X}_A^k - \bar{X}_B^k)\delta_l^k - (\bar{X}_A^k - \bar{X}_B^k)\, {\beta^*}_A^k(\bar{X}_A^l - \bar{X}_B^l)}{\left((\bar{X}_A - \bar{X}_B)\, {\beta^*}_A\right)^2}\right]_{1\times K}$$

$$+\left\{\frac{(\bar{X}_A^k - \bar{X}_B^k)\, {\beta^*}_A^k}{(\bar{X}_A - \bar{X}_B)\, {\beta^*}_A}\right\}\left[\left(\overline{X_A{}^l\phi\,(X_A\widehat{\beta_A}^*)} - \overline{X_B{}^l\phi\,(X_B\widehat{\beta_A}^*)}\right)\right]_{1\times K};$$

[from A3.2,

$$\left[\left(\overline{X_A{}^l\phi\,(X_A\widehat{\beta_A}^*)} - \overline{X_B{}^l\phi\,(X_B\widehat{\beta_A}^*)}\right)\right]_{1\times K} = \frac{\partial C}{\partial \widehat{\beta_A}^*}\ ]$$

$\delta_l^k = 1$ if $l = k$

$= 0$ if $l \neq k$

Appendix A3.5 To find $\left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right)$ & $\left(\frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right)$

$$D_k = \left\{\frac{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)}\right\}\left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}$$

$$\left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right) = \left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}\left[\frac{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\overline{X}_B^k\delta_l^k - \overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\bar{X}_B{}^l}{\left(\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right)^2}\right]_{1\times K}$$

$$+\left\{\frac{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\left(\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right.}\right\}\left[\left(\overline{X_B{}^k\phi\left(X_B\widehat{\beta_A}^*\right)}\right)\right]_{1\times K} \quad ;$$

[from Appendix A3.3, $\frac{\partial\left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}}{\partial \widehat{\beta_A}^*} = \frac{\partial D}{\partial \widehat{\beta_A}^*} = \left[\left(\overline{X_B{}^k\phi\left(X_B\widehat{\beta_A}^*\right)}\right)\right]_{1\times K}$]

δ_l^k = 1 if l=k

= 0 if l≠k

Again,

$$\left(\frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right) = -\left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}\left[\frac{\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\overline{X}_B^k\delta_l^k + \overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\bar{X}_B{}^l}{\left(\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right)^2}\right]_{1\times K}$$

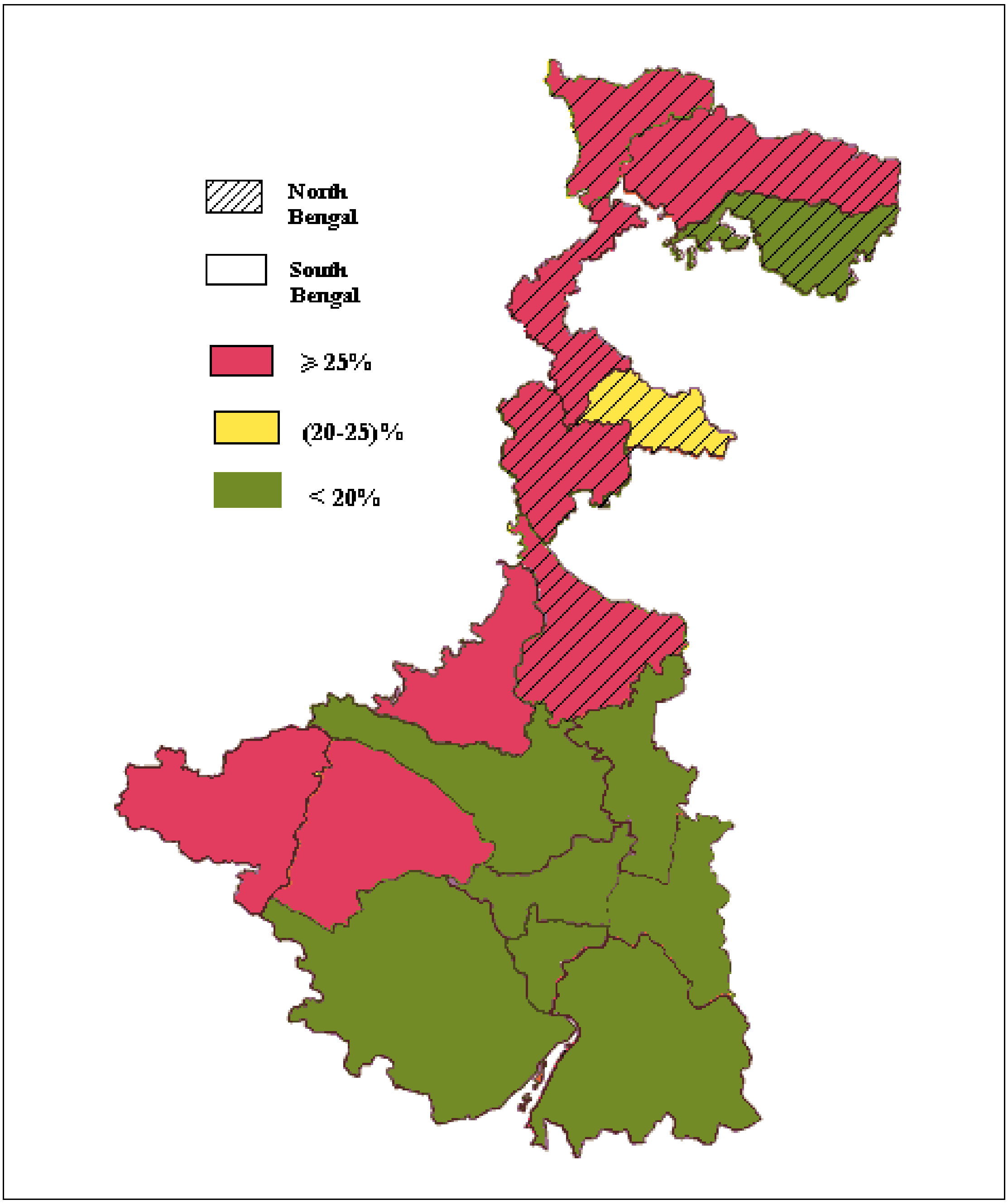
$$-\left\{\frac{\overline{X}_B^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\left(\overline{X}_B\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right.}\right\}\left[\left(\overline{X_B{}^k\phi\left(X_B\widehat{\beta_B}^*\right)}\right)\right]_{1\times K} \quad ;$$

[from Appendix A3.3, $\frac{\partial\left\{\overline{\Phi\left(X_B\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B\widehat{\beta_B}^*\right)}\right\}}{\partial \widehat{\beta_B}^*} = \frac{\partial D}{\partial \widehat{\beta_B}^*} = -\left[\left(\overline{X_B{}^k\phi\left(X_B\widehat{\beta_B}^*\right)}\right)\right]_{1\times K}$]

δ_l^k = 1 if $l = k$

= 0 if $l \neq k$

Appendix A3.6 Figure A3.1 Showing Incidence of Poverty across Districts of North Bengal and South Bengal



CHAPTER 4

ANALYSIS OF POVERTY AND EFFICIENCY: AN EARNINGS FRONTIER APPROACH

4.1 Introduction

Chapter 3 studied the spatial variation in the levels poverty among the geographically segregated units characterized by their intrinsic nature of development status (level of living). The analysis (in Chapter 3) demonstrated that in the decomposition of the poverty gap between North Bengal and South Bengal, both the characteristics effect, that is, the effect due to the differential availability of the characteristics (resource) over the regions and the 'coefficients effect', that is, the effect due to the differential degree of utilization of resources (which has been interpreted as the *efficiency effect*) are significant. Thus both in terms of availability of resources and utilization of resources, there is substantial difference between North Bengal and South Bengal.

In this chapter an attempt has been made to relate the interpretation of *efficiency effect* to the measure of *efficiency* (a measure of '*degree of command over resources'*), used in the context of earnings frontier (defined as the highest potential income associated with a given stock of *human capital* and *endowment*[1]*)*. The relationship between segmentation of West Bengal geographically and that by efficiency scores has been examined.

The concept of *technical efficiency* is an important issue in production function analysis. In the classical production function analysis, each firm, with the objective of maximizing output (subject to the availability of inputs), is expected to operate on its *production frontier*, showing the maximum possible output, given input levels. Empirical studies suggest that given the level of technology, each production unit has a different level of utilization of inputs (see (Tyler, 1979); (Kopp & Smith, 1980); (Fasasi, 2007); (Tong, 2009)). The discrepancy between the potential output (*production frontier*) and the actual output for each firm may thus be attributed to *firm-specific inefficiency*, which can be captured through a random statistical noise taking only positive values (as frontier denotes the maximum possible output, given input levels). *Technical inefficiency* for each firm may

[1] See (Smith, 1759) (Smith, 1776); (Mincer, 1958); (Becker G. S., 1964); (Schultz T. P., 1992) for discussion on human capital theory.

thus be defined in terms of the difference of the actual (estimated) output and its potential (maximum) output. A firm operating below the frontier can increase its output either by increasing input and/or by increasing *technical efficiency,* which, in the current context is the 'output oriented' *technical efficiency.*

Turning to the labour market, the above concept can be borrowed to define the earnings frontier (potential earnings). All individuals are located either on or below this frontier. Individuals who translate their potential earnings into actual earnings enjoy a fully efficient position. In contrast, individuals who earn less than their potential earnings are suffering from some kind of *earnings inefficiency*. A number of studies have estimated the earnings/income frontier using the parametric stochastic frontier approach (SFA) (Jensen, Gartner, & Rassler, 2006)[2]. In this chapter SFA has been used in analysis of poverty, an issue that has not been addressed in any of these studies.[3]

The fact that some individuals are located below the frontier gives rise to a series of questions in the context of poverty: Are the skills, endowments and social opportunities qualitatively different for those who are *efficient* and those who are *inefficient*? Are the *efficient* households clustered in any particular geographical region? Are the *poor* necessarily *inefficient*? Answers to these questions have important policy implications, as it would help assess the performance/applicability of various poverty alleviation programs. The present chapter estimates an earnings frontier using SFA. Splitting the sample into an *efficient* and *inefficient* part based on the estimated frontier and a bench mark efficiency score the status of poverty in the two groups is studied using the Oaxaca decomposition of the poverty gap.[4] The results are compared with those obtained in the previous chapter on the basis of geographical segmentation of West Bengal.

The plan of the chapter is as follows. Section 4.2 discusses the methodology; Section 4.3 describes the data and results and finally section 4.4 concludes. The Appendix at the end of this chapter provides an empirical support for testing a hypothesis.

---

2 See (Aigner, Lovell, & Schmidt, 1977); (Farrell, 1957), (Färe, Grosskopf, & Lovell, 1994) and (Lovell, 1993) for the stochastic frontier approach.

3 (Garfinkel & Haveman, 1977) suggested a measure of economic status called 'earnings capacity' (EC) based on a regression involving socio-economic and demographic determinants and adjustment using returns to assets, an approach different from the stochastic frontier approach. The EC was used to compare and contrast the composition of EC poor with that of the poor based on current income (CY poor). The socio-economic and demographic determinants of EC and CY poverty were compared.

4 The application of Oaxaca decomposition technique has been made in (Bhaumik & Chakrabarty, 2009) to examine the difference in average (log) earnings between two religious groups in Indian and in (Bargain, Bhaumik, Chakrabarty, & Zhao, 2009) to examine the earnings difference between Indian and Chinese wage earners.

4.2 Methodology

The *earnings frontier,* defined as the potential earnings, given the stock of human capital and endowments, is estimated as follows:

$\ln Y_i = \ln \alpha_0 + \sum_{k=1}^{K} \alpha_k \ ln\, x_{ik} - \xi_i,$ (4.1)

where Y_i is the monthly total household consumption expenditure for the i^{th} household, x_{ik} denotes the amount of k^{th} input (human capital/endowment) used by the i^{th} household and ξ_i is an independently and identically distributed *one-sided non-negative error term* with a non-negative mean and constant variance. Therefore, $-\xi_i$ denotes *inefficiency* and the deterministic part denotes the 'frontier'. Parameters of equation (4.1) are estimated using the method of corrected ordinary least square (COLS) technique, described below. Efficiency scores are computed in an analogous manner to estimating technical efficiency from production frontiers.

The Technique of COLS[5]:

Application of Ordinary Least Square (OLS) on (4.1) produce best linear unbiased estimates of slope parameters but *biased* estimate of the constant term, $\ln \alpha_0$, because of the distributional assumption on the stochastic term and $E(\xi_i) \neq 0$. Correction for the bias is made as follows.

Rewrite equation (4.1) as

$\ln Y_i = (\ln \alpha_0 - E(\xi_i)) + \sum_{k=1}^{K} \alpha_k \ ln\, x_{ik} + (E(\xi_i) - \xi_i).$

$= \alpha_0^* + \sum_{k=1}^{K} \alpha_k \ ln\, x_{ik} + e_i$, say. (4.1.1)

Now, noting that $E(e_i) = 0,$ we apply OLS and obtain the largest possible OLS residual, $\widehat{e^*}$, say. Using $\widehat{e^*}$ as the estimate of $E(\xi_i)$, the unbiased (*corrected*) estimate of the intercept parameter is given by,

$\ln \widehat{\alpha_0} = \widehat{\alpha_0}^* + \widehat{e^*},$ (4.1.2)

where $\widehat{e^*} = max\{\widehat{e_i}\}$.

The corrected residuals (COLS residuals) are given by

$\hat{\xi}_i = \widehat{e^*} - \widehat{e_i}$ (4.1.3)

The COLS residuals are non-negative with at least one being zero and can be used to provide technical efficiency score of each firm. The technical efficiency score of the i^{th} firm is derived as:

$TE_i = exp\,(\widehat{-\xi_i})$ (4.1.4)

---

[5] (Greene W. H., 1980) by The Econometric Approach to Efficiency Analysis, *Greene*, (Aghai, Zarafshani, & Behjat, 2008).

Classification of the Households:

Households are categorized as *efficient* or *inefficient* based on a benchmark level, chosen to be the 95th percentile point of the efficiency scores, calculated using (4.1.4). Households with estimated technical efficiency score greater than or equal to the bench mark efficiency level are classified as *efficient* (Group E) and those below the bench mark are classified as *inefficient* (Group I). Actually, there is no a priori rule for fixation of the bench mark level. In fact, the more the difference between mean technical efficiency scores between the groups, the more appropriate will be categorization of the groups. The higher percentile values of the state level efficiency scores are thus preferable, the restriction being that the sample sizes should be adequate for both the groups to run valid regressions. For each group the incidence of poverty is modeled as in the previous chapter.

The Model for Poverty:

The model for estimating the *incidence of poverty* is given by:

$$\left(\frac{y}{z}\right)_i^* = X_i\beta + \varepsilon_i \; ; \forall\, i = 1, 2, \ldots\ldots, n \qquad (4.2)$$

The incidence of poverty for the i^{th} household is given by:

$$p_i = \Phi(X_i\beta^*); \qquad \left[\beta^* = -\frac{\beta}{\sigma}; var(\varepsilon_i) = \sigma^2\right] \qquad (4.3)$$

A regional dummy (= 1 for South Bengal and = 0 for North Bengal) is introduced in the X-vector of Equation (4.2) for each group, i.e., efficient (E) and inefficient (I). The coefficient of this dummy variable would reveal the impact of geographical location of households in determining the probability of being poor.

The FGT0 measures for the efficient (E) and inefficient (I) groups will, respectively, be:

$$H_E = \frac{1}{n^E}\sum_{i=1}^{n^E}\Phi\left(X_{i_E}\widehat{\beta_E}^*\right) \qquad (4.3.1)$$

$$H_I = \frac{1}{n^I}\sum_{i=1}^{n^I}\Phi\left(X_{i_I}\widehat{\beta_I}^*\right) \qquad (4.3.2)$$

$n^E\, and\; n^I$ being the number of households in groups E and I, respectively.

Now the difference of FGT0 poverty measures between the Groups E and I is (following the same procedure as in Chapter 3)

$$\widehat{D}_{EI} = H_E - H_I = C + D, \text{ say}; \qquad (4.4)$$

C being the *characteristics effect* and D being the *coefficients effect*, where C is given by

$$\text{C} = \{\overline{(X_E} - \overline{X_I})\ \widehat{\beta_E}^*\} \times \phi(\overline{X_E}\widehat{\beta_E}^*) + (R_{T_1} + R_{M_1} - R_{M_2}) \tag{4.5}$$

$$= C^1 + C_R \text{ , say} \tag{4.5.1}$$

where $C^1 = \{\overline{(X_E} - \overline{X_I})\ \widehat{\beta_E}^*\} \times \phi(\overline{X_E}\widehat{\beta_E}^*)$ and $C_R = (R_{T_1} + R_{M_1} - R_{M_2})$

$\overline{X_E}$ *and* $\overline{X_I}$ are vectors representing average availability of the explanatory variables (characteristics) in group E and group I, respectively, and R_{T_1} and R_{M_i}'s are the approximation residuals as described in the previous chapter.

For variables with positive estimates of coefficients ($\widehat{\beta_E}$ *and* $\widehat{\beta_I}$), (4.2) can be interpreted as an earnings frontier (analogous to a production frontier) with human capital and endowments as inputs as in (4.1). The associated variables (with positive coefficients and hence with *positive marginal products*) can be considered as inputs (resources) of production. $\overline{X_E}$ *and* $\overline{X_I}$ can thus be considered as the *resource vectors* for Group E and Group I, respectively. Thus, C, a function of the *resource vectors* in two groups, can be considered as *resource effect*.

D is given by:

$$D = \overline{X_I}(\widehat{\beta_E}^* - \widehat{\beta_I}^*) \times \phi(\overline{X_I}\widehat{\beta_I}^*) + (R_{T_2} + R'_{M_1} - R'_{M_2}) \tag{4.6}$$

$$= D^1 + D_R \text{ , say} \tag{4.6.1}$$

where $D^1 = \overline{X_I}(\widehat{\beta_E}^* - \widehat{\beta_I}^*) \times \phi(\overline{X_I}\widehat{\beta_I}^*)$; $D_R = (R_{T_2} + R'_{M_1} - R'_{M_2})$;

$\widehat{\beta_E}^* = -\frac{\beta_E}{\sigma_E}$ and $\widehat{\beta_I}^* = -\frac{\beta_I}{\sigma_I}$ are vectors of transformed regression coefficients, R_{T_2} and $R'_{M_i}{}'s$ are the approximation residuals.

For positive values, the coefficient vector ($\widehat{\beta_E}, \widehat{\beta_I}$) may be interpreted as the *productivity* vector and D may thus be considered as a function of the transformed productivity vectors ($i.e., \widehat{\beta_E}^*$ *and* $\widehat{\beta_I}^*$). As discussed in Chapter 3, the *coefficients effect* in the Oaxaca decomposition can be interpreted as the '*Efficiency effect'*, which gives the differential degree of utilization of resources given the resources. As the magnitudes of the coefficient vectors determine the levels of efficiency, D is expected to be *highly significant* while decomposing the difference in the incidences of poverty between two groups with significant difference in the mean level of *efficiency scores*.[6]

[6] A correspondence (between efficiency and the coefficients effect) can be established using the bootstrap procedure in the following manner. Suppose R samples (with replacement) are drawn from the original data set and let D^r_{pov} be the share of the aggregate coefficients effect in the poverty gap between the Groups E and I for the r^{th} resampled data set. Now as seen from (4.6), D^r_{pov} will be functionally related to $(\widehat{\beta_E}^*, \widehat{\beta_I}^*)$ and thus to

Detailed Decomposition Analysis:

As discussed earlier, the contribution of specific factors can be factored out from the overall contribution of all the variables in making the aggregate characteristics, C and aggregate coefficients effects, D and the poverty gap can thus be written as shares of individual explanatory variables as:

$$H_E\text{-}H_I = \left\{ \sum_{k=1}^{K} V_{\Delta X}^{k} \left[\overline{\Phi(X_E \widehat{\beta_E}^{*})} - \overline{\Phi(X_I \widehat{\beta_E}^{*})} \right] + \sum_{k=1}^{K} V_{\Delta\beta}^{k} \left[\overline{\Phi(X_I \widehat{\beta_E}^{*})} - \overline{\Phi(X_I \widehat{\beta_I}^{*})} \right] \right\}, \tag{4.7}$$

where $V_{\Delta X}^{k}$ and $V_{\Delta\beta}^{k}$ are the shares of the k^{th} explanatory variable in C and D, respectively. The characteristic effect due to the k^{th} explanatory variable is, thus given by:

$$C_k = V_{\Delta X}^{k} \times C \quad ; \quad V_{\Delta X}^{k} = \frac{(\bar{X}_E^k - \bar{X}_I^k)\, {\beta^*}_E^k\, \phi(\overline{X_E \widehat{\beta_E}^{*}})}{(\bar{X}_E - \bar{X}_I)\, {\beta^*}_E\, \phi(\overline{X_E \widehat{\beta_E}^{*}})} \tag{4.8}$$

$$\text{i.e., } C_k = \frac{(\bar{X}_E^k - \bar{X}_I^k)\, {\beta^*}_E^k\, \phi\,(\overline{X_E \widehat{\beta_E}^{*}})}{C^1} C \quad ; \text{[using (4.5.1) and (4.8)].} \tag{4.8.1}$$

Now as long as C and C^1 have the same signs[7], a positive value of C_k implies that $\forall\, \beta_E^k > 0$ (that is, ${\beta^*}_E^k < 0$), $\bar{X}_E^k < \bar{X}_I^k$. i.e., Group I has a higher availability of the k^{th} input, x^k, which has a positive marginal product.

A negative value of C_k implies that Group E has a higher availability of the k^{th} input, x^k, which has a positive marginal product.

The coefficients effect due to the k^{th} explanatory variable is given by:

$$D_k = V_{\Delta\beta}^{k} \times D \quad ; \quad V_{\Delta\beta}^{k} = \frac{\left\{\overline{X}_I^k \left(\widehat{\beta^*}_E^k - \widehat{\beta^*}_I^k\right)\right\} \phi(\overline{X_I \widehat{\beta_I}^{*}})}{\{\overline{X}_I\, (\widehat{\beta^*}_E - \widehat{\beta^*}_I)\}\, \phi(\overline{X_I \widehat{\beta_I}^{*}})} \tag{4.9}$$

$$\text{i.e., } D_k = \frac{\left\{\overline{X}_I^k \left(\widehat{\beta^*}_E^k - \widehat{\beta^*}_I^k\right)\right\} \phi\left(\overline{X_I \widehat{\beta_I}^{*}}\right)}{D^1} D \quad \text{[using (4.6.1) and (4.9)].} \tag{4.9.1}$$

$(\widehat{\beta_E}, \widehat{\beta_I})$. Let δ_{eff}^{r} be the absolute difference of the average efficiency scores for the groups E and I for the r^{th} resampled data set. D_{pov}^{r} should thus be functionally related to δ_{eff}^{r}. Here the correlation coefficient between the values of D_{pov}^{r} and δ_{eff}^{r} is found to be -0.68, indicating a correspondence between *coefficient effect* and *efficiency effect.*

[7] The idea of opposite signs of C and C^1 is not quite appealing intuitively as this means a situation where the effect of the term C_R is more prominent (in absolute terms) compared to C. This is an undesirable situation given the objective of the decomposition analysis and thus not considered. In the present context, however, C and C^1 have the same signs.

As long as D and D^1 have the same signs[8], a positive value of D_k implies that $\hat{\beta}^{*k}_{E} > \hat{\beta}^{*k}_{I}$..
This leads to the following two cases.

Case 1

A positive value of D_k means that $\left(\frac{\widehat{\beta_E}^k}{\widehat{\sigma_E}} - \frac{\widehat{\beta_I}^k}{\widehat{\sigma_I}}\right)\overline{X}_I^k < 0$ [since $\beta^{*k}_{E} = \frac{-\beta_E{}^k}{\sigma_E}$ and $\beta^{*k}_{I} = \frac{-\beta_I{}^k}{\sigma_I}$].

1. If $\widehat{\sigma_E} \leq \widehat{\sigma}_I$, we can definitely say that $\beta_E{}^k - \beta_I{}^k < 0$, i.e., when both $\beta_E{}^k$ *and* $\beta_I{}^k$ are positive, the above result signifies that in terms of utilization of *k*th component of the resource vector (here x^k), Group E is less efficient than Group I.
2. For $\widehat{\sigma_E} > \widehat{\sigma}_I$, $\beta_E{}^k - \beta_I{}^k \gtreqless 0$ and needs to be verified empirically.

Case 2

A negative value of D_k means that $\left(\frac{\widehat{\beta_E}^k}{\widehat{\sigma_E}} - \frac{\widehat{\beta_I}^k}{\widehat{\sigma_I}}\right)\overline{X}_I^k > 0$ [since $\beta^{*k}_{E} = \frac{-\beta_E{}^k}{\sigma_E}$ and $\beta^{*k}_{I} = \frac{-\beta_I{}^k}{\sigma_I}$].

1. If $\widehat{\sigma_E} < \widehat{\sigma}_I$, $\beta_E{}^k - \beta_I{}^k \gtreqless 0$ and needs to be verified empirically.
2. If $\widehat{\sigma_E} \geq \widehat{\sigma}_I$, we can definitely say that $\beta_E{}^k - \beta_I{}^k > 0$, i.e., when both $\beta_E{}^k$ *and* $\beta_I{}^k$ are positive, the above result signifies that in terms of utilization of resource (here x^k), Group E is more efficient than Group I.

Asymptotic Variance of C and D:

Using the Delta method, the asymptotic variance of C ($=\sigma_C^2$) and the asymptotic variance of D ($=\sigma_D^2$) can be obtained as (The derivation is similar to that presented in Chapter 3):

$$\sigma_C^2 = \left(\frac{\partial C}{\partial \widehat{\beta_E}^*}\right) Asymptotic\ Variance\ (\beta_E{}^*) \left(\frac{\partial C}{\partial \widehat{\beta_E}^*}\right)^T \tag{4.10}$$

$$\sigma_D^2 = \left(\frac{\partial D}{\partial \widehat{\beta_E}^*}\right) (\text{Asymptotic Variance } (\beta_E{}^*)) \left(\frac{\partial D}{\partial \widehat{\beta_E}^*}\right)^T + \left(\frac{\partial D}{\partial \widehat{\beta_I}^*}\right) (\text{Asymptotic Variance } (\beta_I{}^*)) \left(\frac{\partial D}{\partial \widehat{\beta_I}^*}\right)^T \tag{4.11}$$

Asymptotic Variance of C_k and D_k:

The asymptotic variance of C_k, $\sigma_{C_K}^2$ and The asymptotic variance of D_k, $\sigma_{D_K}^2$ can be computed as in the previous chapter as follows:

[8] The situation of D and D^1 having the different signs is not discussed due to the reason mentioned in the previous footnote.

$$\sigma_{C_K}^2 = \left(\frac{\partial C_k}{\partial \widehat{\beta_E}^*}\right) Asy\ var\ (\beta_E{}^*)\ \left(\frac{\partial C_k}{\partial \widehat{\beta_E}^*}\right)^T \quad (4.12);$$

$$\sigma_{D_K}^2 = \left(\frac{\partial D_K}{\partial \widehat{\beta_E}^*}\right) (Asy\ var\ (\beta_E{}^*)) \left(\frac{\partial D_K}{\partial \widehat{\beta_E}^*}\right)^T + \left(\frac{\partial D_K}{\partial \widehat{\beta_I}^*}\right) (Asy\ var(\beta_I{}^*)) \left(\frac{\partial D_K}{\partial \widehat{\beta_I}^*}\right)^T \quad (4.13)$$

Testing the Significance of C and D:

As in the previous chapter, the test statistic under the null hypothesis (C=0, D=0) is given by $t_c = \frac{C}{\sigma_C}\ and\ \ t_D = \frac{D}{\sigma_D}$, which are asymptotically normally distributed.

The test statistic under the null hypothesis (C_k=0, D_k=0) are given by $t_{C_k} = \frac{C_k}{\sigma_{C_k}}$ and $t_{D_k} = \frac{D_k}{\sigma_{D_k}}$, which are asymptotically normally distributed.

4.3 Data and Results

The variables for the poverty model are the same as those used in Chapter 3.

The variables used as inputs in the estimation of the earnings frontier are:

A. Endowment

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

(2) Household size, HHSIZE

(3) 1-Dependency, 1-DEPRAT; where

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

B. Human capital

(1) The General Educational level, GENEDU.[9]

The estimates of the parameters of the earnings frontier are reported in Table 4.1. All the parameters have the expected sign and are highly significant.

Table 4.2 presents the location wise distribution of households classified as *efficient* (E) or *inefficient* (I) based on estimates of Table 4.1 and taking the 95th percentile point of the efficiency scores as the bench mark efficiency score (= 0.34) to classify the households. It is observed from Table 4.2 that only 16.7% of the households of rural West Bengal are efficient, given the bench mark efficiency score of 0.34. Among the inefficient group of

[9] 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. Since codes/indicators increase with levels of education, these have been taken as proxy for years spent in education.

83.3%, 70% households belong to South Bengal. A Chi-square test for independence to determine whether there is a significant relationship between the geographical grouping and grouping by efficiency score produces a highly significant $\chi^2_{(1)}$ value of 22.8 indicating that classification by location and that by 'efficiency score' are not independent.[10]

Table 4.3 presents the estimates of the parameters of equation (4.2) for both the e*fficient* and *inefficient* groups. Almost all the parameter estimates are significant at 5% level and have the expected signs. A significant positive estimate of the coefficient of the regional dummy indicates that living in South Bengal has a positive effect on consumption expenditure and hence reduces the probability of being poor. This is in line with the earlier findings that North Bengal is poorer than South Bengal. For most of the other variables the magnitude of the coefficients are higher for Group E compared to Group I.[11] This is expected, given the categorization of the regions in terms of the efficiency scores.

Table 4.4 presents the estimates of poverty (FGT0) for groups E and I. It is observed that poverty (= 0.27) is much higher for Group E as compared to that (= 0.19) for Group I.

Table 4.5 presents the distribution of households classified as *efficient* (E) or *inefficient* (I) and as poor or non poor. A Chi-square test similar to that mentioned above produces a highly significant $\chi^2_{(1)}$ value of 24.9 indicating that classification by 'poverty' and that by 'efficiency' are not independent. Thus, the results combined together point to a link among segmentations of West Bengal (i) geographically (into North and South Bengal), (ii) by incidence of poverty and (iii) by 'efficiency'. Table 4.6 presents the Oaxaca decomposition analysis of the difference in the incidences of poverty between the groups and Table 4.7 presents the observed resource vectors of the two groups.

The summary results from these tables and some other important findings are listed below.

1. Group E, i.e., the group with higher mean efficiency score has a higher incidence of poverty (=0.27) compared to Group I (=0.19)[12], the group with lower mean efficiency score, i.e., as far as the spatial variation in the incidences of poverty are concerned, it varies in a

---

[10] This finding is contingent upon the benchmark value and model specification. It would be an interesting exercise to examine the robustness of this finding.

[11] As mentioned earlier, PNSCH is expected to have a negative influence on household monthly per-capita expenditure. A possible explanation is that the income earned by joining the labour market is smaller than the gain in income made through increased in efficiency resulting from joining the educational institutions.

[12] The conventional non-regression based estimates are almost equal to the regression based estimates of poverty in Regions E and I.

way that poor people are on an average more efficient and enhancement of resources in the efficient group (E) is a possible way to reduce the poverty gap between the groups. This finding is in line with the " small farmers are poor but efficient" theory (Chong, Lizarondo, Cruz, Guerrero, & Smith, 1984), which says that given their knowledge and resources, the traditional farmers are in general good decision makers but scarcity (high price) of capital and non-access to and unavailability of new agricultural technology have hindered their agricultural transformation. This view is strongly supported in (Schultz T. W., 1965). Schultz advocates the concentration on high-payoff new inputs in the form of material and human capital for improvement in the state of the art of production techniques of the farmers. Empirical supports of Schultz's ideas have been found in (Norman, 1977) and (Rask, 1977).[13] Thus, in general, enhancement of resources in Region E could be a possible way to reduce the poverty gap.

2. The share of the characteristics effect in explaining the poverty difference $(\mathrm{H_E} - \mathrm{H_I})$ is 198%. That is,

$$\frac{C}{H_E - H_I} \times 100 = 198$$

$$\Rightarrow C = 1.98\,(H_E - H_I) \approx 2(H_E - H_I)$$

This means that characteristics effect (that is the effect of the differential availability of characteristics (resources)) over the regions is responsible for almost twice of the poverty gap.

Again, as the Characteristics effect, C and the Coefficients effect, D together constitute the poverty gap, $(\mathrm{H_E} - \mathrm{H_I})$, 198% share of C in $(\mathrm{H_E} - \mathrm{H_I})$ has to be compensated by a (-98%) share of D in $(\mathrm{H_E} - \mathrm{H_I})$.

This signifies that differential impact of the characteristics over the regions, i.e., the negative coefficients effect is responsible for almost the entire poverty gap, which is positive. Hence, equalization of the coefficients will make the Region E *worse off* (as explained in Chapter 3, see (Bhaumik, Gang, & Yun, 2006)).

3. The characteristics effect, C is decomposed as contributions ($C_k's$) by specific variables. The share of C_k in $(\mathrm{H_E} - \mathrm{H_I})$ denotes the magnitude and direction of equalization for the k^{th} specific variable. A positive (negative) C_k and hence a positive (negative) share of C_k in $(\mathrm{H_E} - \mathrm{H_I})$ denotes that for a variable with positive coefficient (β), Region E is deficient in the average availability of that particular variable, $\bar{X}^k$ compared to Region I. This can also be

[13] A statistical analysis in support of the 'poor but efficient' hypothesis is given in Appendix A4.1.

verified from Table 4.7 showing the observed resource vectors in Regions E and I. This means that raising the resource level of Group E to the level of Region I will lower the poverty gap (for a positive C_k).

4. The coefficients effect, D is decomposed as contributions (D_k's) by specific variables. The share of D_k in ($H_E - H_I$) denotes the magnitude and direction of equalization for the k^{th} specific variable.

Given that $\widehat{\sigma_E} > \widehat{\sigma_I}$,

A negative D_k and hence a negative share of D_k in $(H_E - H_I)$ denotes that for a variable with positive coefficient (β), Group E is having a higher coefficient attached with the k^{th} explanatory variable x^k. This can also be verified from Table 4.3 showing the coefficient vectors in Groups E and I. The equalization thus, in terms of an increment of the coefficient of Group I to the level of that of Group E will make E worse off (for a negative D_k).

A positive value D_k and hence a positive share of D_k in $(H_E - H_I)$ does not reveal any unambiguous comparison between the regional coefficients. The direction of equalization can, however, be checked from Table 4.3.

5. In light of the arguments above, by enhancing the resource in terms of educational status and wealth status in Group E to that of I, the poverty gap can be reduced.

6. The high share of coefficient effect by the variable (1-DEPRAT) is attributable to the fact that (1-DEPRAT) is a variable in the estimation of the earnings frontier and that Groups E and I have been formed on the basis of efficiency scores. The average efficiency difference between the Groups is expected to be reflected in the Coefficients effect.

7. Highly negative magnitude of the share of the constant term in $(H_E - H_I)$ indicates that there is huge baseline gap in consumption between E and I and baseline consumption is much higher in E compared to I.

4.4 Conclusion

Economic problems in relation to analysis of income can be looked upon in the perspective of potential income analysis. The problem of poverty can be viewed in terms of low income earning potential and can be linked with the notion of technical efficiency. In this chapter we have attempted to capture the relationship between poverty and technical efficiency by estimating an income earnings frontier and by studying the incidences of poverty between two groups differentiated with respect to technical efficiency scores. It turns out that the efficient group (with high mean efficiency score) has a high incidence of poverty compared to the inefficient group (with low mean efficiency score). This is in line, particularly in the context of rural West Bengal, with Schultz' hypothesis that 'poor people are in fact more efficient in a sense they use their resources more efficiently and there is little unutilized resources. That is, the dominant problem faced by the traditional agricultural farmer is the scarcity of resources'.

The Oaxaca decomposition of the poverty gap between the efficient and inefficient groups yields a significant *coefficients effect*, interpreted as *efficiency effect*.

The essence of the findings in this chapter is that poverty (in fact, rural poverty) is mainly a problem originating from the scarcity of resources. So far as the specific resource is concerned, the average household general educational level is the largest contributor in influencing the difference in the incidences of poverty between the groups and enhancement of the education level for Group E would be a possible solution to narrowing the difference of poverty levels between the two groups. It also emerges that geographical location of households does have significant impact in explaining the poverty gap.

Table 4.1 Parameter Estimates of Earnings Frontier Using COLS

Inputs	Estimates	t-values
(1)	(2)	(3)
Land possessed (LAND)	0.0423	13.89*
Household size (HHSIZE)	0.3915	24.09*
1-Dependency Ratio (1-DEPRAT)	0.2116	12.54*
General Education Level (GENEDU)	0.3168	31.31*
Constant	6.2511	242.47*

*Significant at 5% level; R^2= 0.54

Table 4.2 Distribution of Households (Efficient/Inefficient) by Geographical Location

Bench Mark Efficiency = 0.34

	Efficient	Inefficient	Total
(1)	(2)	(3)	(4)
North Bengal	6.35	24.72	31.07 (0.229)*
South Bengal	10.33	58.60	68.93 (0.233)*
Total	16.70	83.30	100

$\chi^2(1)$ = 22.8 > 3.84 (critical value at 5% level)

*Figures in parentheses are the average efficiency scores with respect to the state of West Bengal.

Table 4.3 Factors Influencing Per-Capita Household Consumption: (ML Estimation)

(Dependent Variable: $\left(\frac{y}{z}\right)$)

		Group E		Group I	
Characteristics	Variables	Estimate (β_E)	t value	Estimate (β_I)	t value
(1)	(2)	(3)	(4)	(5)	(6)
Demographic characteristics of the households	1-DEPRAT	0.1026	1.45	0.3302	11.30*
	D_FEMH	0.0802	1.82	0.0580	3.42*
Educational status of the household	PSECEDU	0.2072	0.95	0.0985	2.62*
	PTERTEDU	0.0580	0.31	0.1411	3.57*
	GENEDU	0.1287	11.36*	0.0702	22.18*
Wealth status	PLAND	0.0006	4.10*	0.0004	12.36*
Labour market characteristics	POWNAC	0.4087	4.21*	0.3445	8.78*
	PNSCH	-0.8816	-4.12*	-0.2147	-6.09*
	PDOMO	0.2841	3.09*	0.1815	4.70*
	PDOM	0.5639	4.15*	0.2701	6.41*
	PEMP	0.1466	1.75	0.1484	4.33*
Regional Dummy		0.1871	4.94*	0.1114	11.45*
Government aid	D_GOVAID	0.1699	1.94	0.1745	9.90*
Constant		-0.2177	-3.70*	-0.4399	-23.67*

*Significant at 5% level.

Table 4.4 Poverty Incidence in Group E and Group I

	Sample size	Mean Technical Efficiency	Poverty incidence
(1)	(2)	(3)	(4)
Group E (efficient)	772	0.4403	0.2666
Group I (Inefficient)	4161	0.2205	0.1903
	Difference in poverty incidence: $(H_E\text{-}H_I) = 0.0763$		

Table 4.5 Distribution (Percentage) of Households by Efficiency and Poverty

(Bench Mark Efficiency = 0.34)

	Efficient	Inefficient	Total
(1)	(2)	(3)	(4)
Poor	4.57	16.38	20.95
Non poor	12.13	66.92	79.05
Total	16.70	83.30	100.00

$\chi^2(1) = 24.9 > 3.84$ (critical value at 5% level)

Table 4.6 Decomposing the Poverty Gap between Efficient (Group E) and Inefficient (Group I)

(Using Estimates of Table 4.3)

			Aggregate Characteristics effect (C)			Aggregate Coefficients effect(D)		
Aggregate effect***			Estimate	Share in (H_E-H_I)**		Estimate	Share in (H_E-H_I)	
			0.1515 (0.0108)*	198.5551		-0.0752 (0.0105)*	-98.5551	
Decomposition of the Aggregate effect			Individual Characteristic effect (C_k)' s	Individual Coefficient effect (D_k)'s		Individual Characteristic effect (C_k)' s		
			Shares in (H_E-H_I))**	Estimates		Shares in (H_E-H_I)**	Estimates	
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	Demographic characteristics of the households	1-DEPRAT	-0.0013 (0.0001)	-1.74	-9.16	0.1534 (0.0025)	200.99	201.97
		D_FEMH	-0.0057 (0.0002)	-7.42		0.0008 (0.0001)	0.98	
II	Educational status of the household	PSECEDU	0.0028 (0.0006)	3.68	217.6	-0.0010 (0.0019)	-1.35	-12.06
		PTERTEDU	0.0004 (0.00007)	0.48		0.0055 (0.0004)	7.22	
		GENEDU	0.1629 (0.0019)	213.44		-0.0137 (0.0043)	-17.93	
III	Wealth status	PLAND	0.0046 (0.0000007)	6.07	6.07	0.0061 (0.000005)	7.95	7.95
IV	Labour market characteristics	POWNAC	0.0034 (0.0003)	4.43	-25.37	0.0152 (0.0034)	19.87	51.17
		PNSCH	-0.0185 (0.004)	-24.28		0.0170 (0.0094)	22.32	
		PDOMO	-0.0031 (0.0003)	-4.08		0.0028 (0.0023)	3.68	
		PDOM	0.0055 (0.0008)	7.24		-0.0040 (0.0043)	-5.29	
		PEMP	-0.0066 (0.0006)	-8.68		0.0081 (0.0001)	10.59	
			0.0070 (0.0003)	9.18	9.18	0.0026 (0.0034)	3.37	3.37
V	Government aid	D_GOVAID	0.0002 (0.00002)	0.22	0.22	0.0055 (0.0009)	7.23	7.23
	Constant					-0.2733 (0.0082)	-358.18	-358.18

* Figures in parentheses denote the standard errors. Both the aggregate characteristics effect and the aggregate coefficients effect are significant at 5% level. Also the individual characteristic and coefficient effects are highly significant at 5% level.

**share = estimates/ Difference in poverty incidence (Table 4.4)) × 100

*** Aggregate Characteristics effect, $C = \sum_{k=1}^{K} C_k$ and Aggregate Coefficients effect, $D = \sum_{k=1}^{K} D_k$

Table 4.7 Observed Resource Vectors for Efficient (E) and Inefficient (I) Regions

Characteristics		Variables under characteristics	$\bar{X}^k$	
			Region E	**Region I**
			$\bar{X}_E^k$	$\bar{X}_I^k$
(1)		(2)	(3)	(4)
I	Demographic characteristics of the households	1-DEPRAT	0.6787	0.6516
		D_FEMH	0.2202	0.0719
II	Educational status of the household	PSECEDU	0.0303	0.0588
		PTERTEDU	0.0456	0.0589
		GENEDU	1.7763	4.4347
III	Wealth status	PLAND	69.8634	87.0874
V	Labour market characteristics	POWNAC	0.1410	0.1584
		PNSCH	0.0235	0.0676
		PDOMO	0.1954	0.1724
		PDOM	0.0672	0.0878
		PEMP	0.2396	0.1447
V	Government aid	D_GOVAID	0.6244	0.7030

APPENDICES

Appendix A4.1: Testing the Poor but Efficient Hypothesis

The difference of FGT0 measures between E and I, $\widehat{D}_{EI}$ =H_E-H_I

$$= \left[\frac{1}{n^E}\sum_{i=1}^{n^E} \Phi\left(X_{i_E}\widehat{\beta_E}^*\right)\right] - \left[\frac{1}{n^I}\sum_{i=1}^{n^I} \Phi\ \left(X_{i_I}\widehat{\beta_I}^*\right)\right] = \overline{\Phi\,(X_E\widehat{\beta_E}^*)} - \overline{\Phi(X_I\widehat{\beta_I}^*)}.$$

The test statistic under the null hypothesis ($\widehat{D}_{EI}$ = 0), that there is no difference in poverty estimates between the two groups, is given by $t_{D_{EI}} = \frac{\widehat{D}_{EI}}{\widehat{\sigma}_{D_{EI}}}$, which asymptotically follows a normal distribution. Now, the asymptotic variance $\widehat{\sigma}^2_{D_{EI}}$

$$= \left(\frac{\partial D_{EI}}{\partial \widehat{\beta_E}^*}\right) (\text{Asymptotic Variance } (\beta_E{}^*)) \left(\frac{\partial D_{EI}}{\partial \widehat{\beta_E}^*}\right)^T$$

$$+ \left(\frac{\partial D_{EI}}{\partial \widehat{\beta_I}^*}\right) (\text{Asymptotic Variance } (\beta_I{}^*)) \left(\frac{\partial D_{EI}}{\partial \widehat{\beta_I}^*}\right)^T;$$[14] where

$$\frac{\partial D_{EI}}{\partial \widehat{\beta_E}^*} = \left[\left(\frac{1}{n^E}\sum_{i=1}^{n^E} X_{i_E}{}^k \phi\left(X_{i_E}\widehat{\beta_E}^*\right)\right)\right]_{k=1(1)K} \quad \text{and}$$

$$\frac{\partial D}{\partial \widehat{\beta_I}^*} = -\left[\left(\frac{1}{n^I}\sum_{i=1}^{n^I} X_{i_I}{}^k \phi\left(X_{i_I}\widehat{\beta_I}^*\right)\right)\right]_{k=1(1)K}.$$

Under the assumption of non-normality the statistical precision of the estimate of D_{EI} can also be found by the method of bootstrapping. Let X be the original data set. Observations being drawn with replacement from X, let X^i be the i^{th} resampled data and H_E^i and H_I^i be the estimates of poverty for the efficient and inefficient Groups, respectively, for X^i. For R replications, there will be R realizations of H_E^i and H_I^i and correspondingly R realizations of D_{EI}^i, where $D_{EI}^i = H_E^i - H_I^i$.

The estimate of the standard error of $\widehat{D}_{EI}$ can be derived as,

$$\widehat{\sigma}_{D_{EI}} = \sqrt{\frac{\sum_{i=1}^{R}\left(D_{EI}^i - \overline{D}_{EI}\right)^2}{R-1}}\ , \text{ where } \overline{D}_{EI} = \frac{\sum_{i=1}^{R}(D_{EI}^i)}{R}$$

---

[14] $\beta^* = -\frac{\beta}{\sigma} = f(\beta,\sigma)$. Thus, asymptotic Variance (β^*) can be obtained in a similar manner as described in Appendix A3.1

Data and Results:

Households have been categorized as *efficient* or *inefficient* depending on a number of benchmark levels, which have been set based on percentile values of efficiency scores[15]. Standard error of $\widehat{D}_{EI}$ has been calculated in each case using both the Delta and bootstrap methods. Poverty gap turns out to be positive and significant at 5% level, validating the poor but efficient hypothesis. The table below shows the result of the empirical exercise.

Table A4.1: Difference in Poverty Incidences at Different Bench Mark Values

Bench Mark Efficiency Score	Difference in poverty incidences: D_{EI}	$t_{D_{EI}}$ (under normality)	$t_{D_{EI}}$ (bootstrapped)
(1)	(2)	(3)	(4)
Mean efficiency score = 0.2316	0.1307	15.31	12.91
50th Percentile score = 0.2133	0.1378	18.95	14.15
75th Percentile score = 0.2498	0.1187	11.79	9.89
90th Percentile score = 0.3000	0.0871	6.10	7.02
95th Percentile score = 0.3436	0.0868	5.71	5.90

[15] Since there is no fixed bench mark value for efficiency scores, several percentile values have been tried subject to the fact that adequate sample size is found for the two groups.

CHAPTER 5

DECOMPOSING DIFFERENCE IN POVERTY INCIDENCES: A SPATIAL REFORMULATION

5.1 Introduction

The estimation of the incidences of poverty studied in the previous chapters is based on the assumption that the observations of the dependent variable under study, i.e., the incidences of poverty, are independent. That is, the probability of a person being poor is independent of the probability of his neighbours being poor. This, however, may not be a realistic assumption because often variables ordered in geo-space exhibit a correlation pattern, known as *spatial autocorrelation*.[1] The genesis of this correlation pattern may be linked intuitively with a fundamental notion in Geography, which says that nearby entities often share more similarities than entities which are far apart, an idea which is known as *Tobler's First Law of Geography*. The essence of the above law is that "everything is related to everything else, but nearby things are more related than distant things" (Tobler, 1970).

The term *spatial autocorrelation* (Anselin L. , 1988) can be formally expressed by the moment condition as: $cov[y_i, y_j] = E[y_i, y_j] - E[y_i]E[y_j] \neq 0$, for $i \neq j$, where i, j refer to individual observations (locations) and y_i is the value of the random variable of interest at location i. 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.[2] Spatial autocorrelation can take both positive and negative values. In the positive case, the value of a variable at a given location tends to be similar to the values of that variable in nearby locations. This means that if the value of some variable is low in a given location, the presence of spatial autocorrelation indicates that nearby values are also low. On the other hand, negative spatial autocorrelation is characterized by dissimilar values in nearby

---

[1] It is observed from the Map in Chapter 3 that there is a pattern in the incidences of poverty across districts of North Bengal and South Bengal.

[2] Spatial autocorrelation may be compared apparently with the familiar temporal dependence in time series analysis modelled either by a lagged dependent variable or via the error term. However, the dependence in time series is uni dimensional in the sense that past influences the present, while the dependence is conceived to be multidimensional in the cross-sectional spatial case as neighbours influence the behavior of their neighbours and vice versa.

locations. Incorporating this feature in the analysis of poverty is important because the self-reinforcing effect of development that 'development often appears to induce more development' (Burger, Van Der Berg, Van Der Walt, & Yu, 2004) can take a serious dimension resulting in widening gaps between regions, creating *pockets of poverty* that are so extreme that migration may be the ultimate solution of escaping poverty. Certainly, anti-poverty policies should be designed so as to meet the unique features of each region's low-income population and should vary across types and regions (Cushing & Zheng, 1999). It is thus important to examine the significance of spatial effect in the context of analysis of poverty in rural West Bengal.

The existing literature introduces the *spatial effect* in the incidences of poverty, which is the dependent variable in a regression on a set of poverty correlates (*spatial regression*) ((Petrucci, Salvati, & Seghieri, 2004), (Voss P. R., Long, Hammer, & Friedman, 2006)). In the spatial regression model s*patial dependence* can be incorporated in *two* distinct ways: one is by introducing a regressor in the form of spatially lagged dependent variable and the other is by introducing a spatial structure in the error term. The former is referred to as *spatial lag* model and is appropriate when the focus of interest is the assessment of the existence and strength of spatial interaction.

The *spatial lag* model approach starts from theory and posits a structure for spatial dependence a priori. 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, which is of the form:

$y = \rho Wy + X\beta + \varepsilon$.

This is similar to the standard regression model with an additional term being constructed from a pre-defined spatial weight matrix, W, applied to the observed variable y together with a spatial autoregression parameter, ρ, which typically has to be estimated from the data. Here y is the $(n \times 1)$ vector representing the spatially dependent variable. Wy is thus the *spatial lag* term and $W(n \times n)$ 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 weights are either $w_{ij}=1$ (binary weights), or a function of something else, such as $w_{ij}=1/d_{ij}$ $(i \neq j)$, being the distance between observation i and observation j. The diagonal elements of a spatial weight matrix are conventionally set to zero and typically standardized such that the elements of a row sum to one ((Anselin & Bera, 1998) ; (Gibson & Olivia, 2007)). Hence, the spatial weight matrix allows all of the interactions between observation i and each of its neighbours to be parameterized in the form of a weighted average. Specifically, for the random variable of interest y, each element of the spatially lagged variable Wy equals $\sum_j w_{ij} y_j$ which is a weighted average of the *y* values in the neighbourhood of point *i*. The null hypothesis of no autocorrelation corresponds to: $H_0 : \rho = 0$. Ignoring this form of spatial autocorrelation will lead to biased OLS estimates and this is similar to omitting a significant explanatory variable in the regression model.

In contrast to the first approach, the second approach is exploratory in nature in the sense that rather than starting from theory, it attempts to infer an appropriate form for the dependence from cross correlation statistics. Assuming a spatial process for the error terms, either of an autoregressive or moving average form, the *spatial error* model can be formally expressed as: $Y = X\beta + \varepsilon$; $\varepsilon = \lambda W\varepsilon + \xi$,where $W\varepsilon$ 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 $\sigma^2 I$. A spatial moving average process in the error terms takes the form $\varepsilon = \lambda W\xi + \xi$, where now the spatial lag pertains to the errors ξ and not to the original ε. In both the cases the null hypothesis takes the form: $H_0 : \lambda = 0$. The consequences of ignoring spatial error dependence are not quite as severe as those of ignoring spatial lag dependence (Patton & Mcerlean, 2005). The main problem is that the OLS estimates become inefficient, but they are still unbiased (Anselin L. , 1999). In the presence of spatial error dependence, standard error estimates will be biased downward, producing Type I errors (Anselin 1988). The loss of information implicit in this spatial error dependence is accounted for in estimation in order to produce unbiased standard error estimates via maximum likelihood estimation (Darmofal, 2006).

Two other sources of spatial autocorrelation are the *spatial heterogeneity in parameters* (Anselin L. , 1992) and *functional form heterogeneity* (Darmofal, 2006). While the former takes into account the particular features of each location (spatial unit) by explicitly considering varying parameters, random coefficients or various forms of structural change in place of ordinary regression analysis; the latter deals with the issue of spatial heterogeneity in functional form, where different functional forms are valid in different spatially indexed subsets of the data.[3]

In the present chapter *spatial effect* is introduced in the study of poverty incidences using the spatial lag model by introducing spatial dependence in *household level monthly per-*

[3] A third form of spatial heterogeneity is spatial heterogeneity in error variance, leading to spatial heteroskedasticity (Anselin & Griffith, 1988).

capita expenditure values (to be precise, in the logarithm of the ratio of monthly per-capita expenditure values and the poverty line). This is a deviation from the existing literature, which introduces spatial effect in the *incidence of poverty* directly. Spatial dependence has been captured through the spatial correlation (among households) based on geographic and economic distances. From the probability of incidence of poverty estimated at the household level the regional incidences of poverty are obtained.[4]

Next, the disparity in the estimates of poverty between North Bengal and South Bengal is reexamined using a spatial lag model and the difference between the poverty estimates is decomposed into a *characteristics effect* (resource constraint) and a *coefficients effect* (efficiency constraint) using the Oaxaca decomposition method (Oaxaca, 1973) as in Chapter 3. The impact of introduction of spatial effect on the interpretation of the traditional Oaxaca decomposition results is studied.[5]

The plan of the chapter is as follows: Section 5.2 discusses the model; Section 5.3 describes the data and results and finally the Section 5.4 concludes. Appendices A5.1-A5.6 show the detailed derivations of results.

5.2 The Model

It is assumed that consumption of the i^{th} household is influenced by the consumptions of *similar* households in its neighbourhood[6]. This leads to a spatial lagged dependent model of the form

$$\left(\frac{y}{z}\right)_i^* = \rho \sum_{j=1}^{n} w_{ij} \left(\frac{y}{z}\right)_j^* + X_i\beta + \varepsilon_i \,;\, \forall\, i = 1, 2, \ldots\ldots, n\,. \quad (5.1)$$

The subscript *i* denotes 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 (w_{ij}) is the weight defined for the i^{th} household with respect to j^{th} household. The vector of explanatory variables X_i contains an element '1', so that equation (5.1) has a constant term that captures the unobservable neighbourhood effect (in respect of i^{th} household). Thus, while the spatial effect represents spill-over effects of local interactions of households, the

---

4 Application of spatial model in analyzing household demand may be found in (Case, 1991). (Ayadi & Amara, 2008), in their analysis of poverty in Tunisia using spatial techniques find that models with spatially correlated and unobserved spatial heterogeneity are preferred to the traditional non spatial regression model, and give a better approximation of the Tunisian poverty map.

5 This Chapter is based on a paper that has been revised incorporating the comments by two anonymous referees of a Journal. Comments from the referees are highly appreciated.

6 The definition of neighbourhood in the present context is given in section 5.3 of this chapter.

unobservable effect represents a common environment of the neighbourhood, which has a global impact on the neighbourhood. Assuming that the relationship holds for each household, we get from (5.1)

$$\left(\frac{Y}{Z}\right)^{*} = \rho W \left(\frac{Y}{Z}\right)^{*} + X\beta + \varepsilon. \qquad (5.2)$$

Here $\left(\frac{Y}{Z}\right)^{*}$ is the $(n \times 1)$ vector of spatially dependent variable, $W(n \times n)$ the spatial weight matrix with $W\boldsymbol{\iota} = \boldsymbol{\iota}$, where $\boldsymbol{\iota}$ is the $n \times 1$ unit vector. $W\left(\frac{Y}{Z}\right)^{*}$ is the *spatial lag* term. It is assumed that

$\varepsilon \sim N(0, \sigma^2 I_n)$.

From (5.2) we get

$$\left(\frac{Y}{Z}\right)^{*} = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon. \qquad (5.3)$$

Here $(I - \rho W)^{-1}$ is the Leontief inverse, acting like a spatial multiplier in the sense that it links the spatially dependent variable to explanatory variables and errors at all locations.[7] As pointed out by (Anselin & Bera, 1998) the matrix inverse is a full matrix producing an infinite series and thus,

$$\left(\frac{Y}{Z}\right)^{*} = (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots)\, X\beta + (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots)\, \varepsilon. \qquad (5.4)$$

Since the inverse is expanded into an infinite series including both the explanatory variables and the error terms at all locations, the spatial lag term in (5.2), should be treated as an endogenous variable and proper estimation method must account for this endogeneity (Anselin L. , 1999). Since OLS will be biased and inconsistent for this type of model due to the simultaneity bias, either a maximum likelihood (ML) or an instrumental variables estimator is needed for the estimation of a spatial lag model.

Following (Anselin L. , 1992), the maximum likelihood estimation of the spatial lag model is used. It 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. The likelihood function is of the form:

[7] (Anselin L. , 2003) refers to (5.2) and (5.3) as a model with "spatial externalities in both modeled and unmodeled effects" (p. 161) because its reduced form applies a spatial multiplier to both the independent variable and the errors. He also points out that it is constrained by postulating a single multiplier matrix for both (Small & Steimetz, 2009).Given $W\boldsymbol{\iota} = \boldsymbol{\iota}$, $(I - \rho W)^{-1}\boldsymbol{\iota} = (1 - \rho)^{-1}\boldsymbol{\iota}$ (see (Kim, Phipps, & Anselin, 2003)), i.e., the sum of each row of the inverse of the matrix $(I - \rho W)$ sum to $\frac{1}{1-\rho}$, which is the spatial multiplier.

$L=\sum_i \ln(1-\rho\omega_i)-\frac{n}{2}ln(2\pi)-\frac{n}{2}ln(\sigma^2)-\left(\left(\frac{Y}{Z}\right)^*-\rho W\left(\frac{Y}{Z}\right)^*-X\beta\right)^t\left(\left(\frac{Y}{Z}\right)^*-\rho W\left(\frac{Y}{Z}\right)^*-X\beta\right)/2\sigma^2$

; where ω_i is the i^{th} eigen value of the weight matrix W. 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. The concentrated likelihood function takes on the form:

$L_C=-{}^n/_2\, ln[(e_0-\rho e_L)^t\,(e_0-\rho e_L)/n]+\sum_i \ln(1-\rho\omega_i)$,

where e_0 and e_L are the residuals in OLS regression of $\left(\frac{Y}{Z}\right)^*$on X and $W\left(\frac{Y}{Z}\right)^*$, respectively.[8] 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 $(\frac{1}{\omega_{min}},\frac{1}{\omega_{max}})$ is implemented, where ω_{min} $and\ \omega_{max}$ are, respectively, the smallest and the largest eigen values of the weight matrix.[9] 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.

Now, let us rewrite equation (5.3) as follows:

$$\left(\frac{Y}{Z}\right)^*=MX\beta+M\,\varepsilon\ =X^*\beta+\varepsilon^*, \text{ say} \qquad (5.5)$$

where M=$(I-\rho W)^{-1}$ is the Leontief inverse ; $X^*=MX,\ \varepsilon^*=M\,\varepsilon$.

Let p_i be the probability of the i^{th} household being poor.

Then, p_i= prob $\left(\left(\frac{y}{z}\right)_i^*<0\right)$ = prob $(X_i^*\beta+\varepsilon_i^*<0)$ [from (5.5)]

$$= \text{prob } (\varepsilon_i^*<-X_i^*\beta) = \text{prob}\left(\frac{\varepsilon_i^*-E(\varepsilon_i^*)}{\sqrt{var(\varepsilon_i^*)}}<\frac{-X_i^*\beta-E(\varepsilon_i^*)}{\sqrt{var(\varepsilon_i^*)}}\right) = \Phi\left(\frac{-X_i^*\beta-E(\varepsilon_i^*)}{\sqrt{var(\varepsilon_i^*)}}\right)$$

$$=\Phi\left(\frac{-X_i^*\beta}{\sigma_i}\right) \quad ; \quad [E(\varepsilon_i^*)=0\,;\ var(\varepsilon_i^*)\ =\ \sigma_i^2, \text{say}]$$

$$=\Phi\,(X_i^{**}\beta^*); \quad \left[\ \beta^*=-\beta, X_i^{**}=\frac{X_i^*}{\sigma_i}\right] \qquad (5.6)$$

The FGT0 measure for region A will be,

$$\mathrm{H_A}=\frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(X_{i_A}^{**}\widehat{\beta_A}^*\right) \quad [\text{using } (5.6)], \qquad (5.7)$$

---

[8] For further details see ((Ord, 1975), (Anselin L. , 1980), (Anselin L. , 1988), (Anselin & Hudak, 1992)).

[9] See (Anselin & Hudak, 1992) for technical details on the implementation of this bisection search.

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

FGT0 for region B will be,

$$H_B = \frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}^{**}\widehat{\beta_B}^*\right) \quad (5.8)$$

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

Now, the difference of FGT0′s between the regions A and B is given by:

$$\delta_{pov} = H_A - H_B$$

$$= \left[\frac{1}{n^A}\sum_{i=1}^{n^A}\Phi\left(X_{i_A}^{**}\widehat{\beta_A}^*\right)\right] - \left[\frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}^{**}\widehat{\beta_B}^*\right)\right]; \quad \text{[using (5.7) and (5.8)]}$$

$$= \overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)} \;\; \text{; (the over bar denotes sample average)}$$

$$= \left\{\overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)}\right\} + \left\{\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right\} \quad (5.9)$$

= C + D, say

Where C = $\overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)}$ and D = $\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}$.

$\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} = \frac{1}{n^B}\sum_{i=1}^{n^B}\Phi\left(X_{i_B}^{**}\widehat{\beta_A}^*\right) = H_B^C$ is the counter factual poverty in Region B with Region A's coefficients. Note that $\widehat{\beta^*} = -\beta$ (unlike $\beta^* = -\frac{\beta}{\sigma}$ in Chapter 3) and $X_i^{**} = \frac{X_i^*}{\sigma_i}$ $= \frac{(MX)_i}{\sigma_i}$ is the heteroscedasticity and spatial weight adjusted resource vector for the i^{th} household.[10]

Now, following (Yun, 2004) and as has been done in Chapter 3, C can be expressed in an alternative form by evaluating the function $\overline{\Phi(.)}$ s at the mean values and by evaluating the difference of the function $\Phi(\cdot)$ ′s using the first order Taylor expansion as follows:

$$C = \overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} \quad (5.10)$$

$= \left[\Phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right) + R_{M_1}\right] - \left[\Phi\left(\overline{X_B^{**}}\widehat{\beta_A}^*\right) + R_{M_2}\right]$; R_{M_i}'s being approximation residuals.

$= \left\{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right) + \left(R_{T_1} + R_{M_1} - R_{M_2}\right)$; [$R_{T_i}$ is the approximation residual resulting from linearization of the difference of the function $\Phi(\cdot)$ ′s around $\overline{X_A^{**}}\widehat{\beta_A}^*$ by using the first order Taylor expansion ,ϕ being the first derivative of the function Φ].[11]

$= C^1 + C^R$, say;

where $C^1 = \left\{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$ and $C^R = \left(R_{T_1} + R_{M_1} - R_{M_2}\right)$ (5.11)

---

[10] $(MX)_i$ denotes the the i^{th} row of the matrix $(MX) = X^*$

[11] $R_{M_1} = \overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \Phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$; $R_{M_2} = \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \Phi\left(\overline{X_B^{**}}\widehat{\beta_A}^*\right)$

$R_{T_1} = \left\{\Phi(\overline{X_A^{**}}\widehat{\beta_A}^*) - \Phi(\overline{X_B^{**}}\widehat{\beta_A}^*)\right\} - \left\{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$

Now substituting $X^{**}=\Sigma^{-1}MX$ from (5.5) and (5.6) with $\Sigma=\text{diag}(\sigma_1,\sigma_2,\ldots\ldots,\sigma_n)$, where $\sigma_i = var(\varepsilon_i^*)$, C^1 can be written as

$$C^1 = \left[\left\{\overline{\Sigma_A^{-1}(I+\rho_A W_A+\rho_A^2 W_A^2+\cdots)X_A} - \overline{\Sigma_B^{-1}(I+\rho_B W_B+\rho_B^2 W_B^2+\cdots)X_B}\right\}\widehat{\beta_A}^*\right] \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$$

$$; [\because X^* = (I-\rho W)^{-1}X]$$

$$= \left[\left(\overline{\Sigma_A^{-1}X_A} - \overline{\Sigma_B^{-1}X_B}\right)\widehat{\beta_A}^* + \left(\rho_A\overline{\Sigma_A^{-1}W_A X_A} - \rho_B\overline{\Sigma_B^{-1}W_B X_B}\right)\widehat{\beta_A}^* + \left(\rho_A^2\overline{\Sigma_A^{-1}W_A^2 X_A} - \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}\right)\widehat{\beta_A}^*\right] \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$$

[Higher ordered terms involving ρW are ignored in empirical exercises (Lesage & Charles, 2008)] (5.12)

$$= \left\{\left(\overline{\Sigma_A^{-1}X_A} - \overline{\Sigma_B^{-1}X_B}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right) + \left\{\left(\rho_A\overline{\Sigma_A^{-1}W_A X_A} - \rho_B\overline{\Sigma_B^{-1}W_B X_B}\right)\widehat{\beta_A}^* + \left(\rho_A^2\overline{\Sigma_A^{-1}W_A^2 X_A} - \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$$

(5.13)

$= C_{con}^1 + C_{spat}^1$, say. (5.14)

$C_{con}^1 = \left\{\left(\overline{\Sigma_A^{-1}X_A} - \overline{\Sigma_B^{-1}X_B}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$ may be interpreted as the conventional characteristics effect showing the effect of regional characteristics in influencing the disparity in regional poverty incidences δ_{pov}.[12]

$C_{spat}^1 = \left\{\left(\rho_A\overline{\Sigma_A^{-1}W_A X_A} - \rho_B\overline{\Sigma_B^{-1}W_B X_B}\right)\widehat{\beta_A}^* + \left(\rho_A^2\overline{\Sigma_A^{-1}W_A^2 X_A} - \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}\right)\widehat{\beta_A}^*\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^*\right)$ may be interpreted as the spatial characteristics effect. It signifies the impact of the differences in the values of average neighbouring characteristics (weighted by the magnitude of spatial dependence) on the disparity in poverty incidences between two regions.

Similarly, approximating the function $\overline{\Phi(.)}$ s at the mean values and evaluating the difference of the function $\Phi(\cdot)$'s by using the first order Taylor expansion around $\overline{X_B^{**}}\widehat{\beta_B}^*$, D can be written as follows:

$D = \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}$ (5.15)

$= \left\{\Phi\left(\overline{X_B^{**}}\widehat{\beta_A}^*\right) + R_{M_1}'\right\} - \left\{\Phi\left(\overline{X_B^{**}}\widehat{\beta_B}^*\right) + R_{M_2}'\right\}$;

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

---

[12] Strictly speaking, this is not the conventional characteristics effect as it involves terms containing Σ_A^{-1} and Σ_B^{-1}, where $\sum$'s are the *spatial* MLE estimates.

$= \overline{X_B^{**}}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\phi(\overline{X_B^{**}}\widehat{\beta_B}^*) + (R_{T_1}{}' + R_{M_1}{}' - R_{M_2}{}')$ (5.16)

[R_{T_i} is the approximation residual resulting from linearization of the difference of the function $\Phi(\cdot)$'s around $\overline{X_B^{**}}\widehat{\beta_B}^*$ by using the first order Taylor expansion]

$= D^1 + D^R$, say;

where $D^1 = \overline{X_B^{**}}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$ and $D^R = (R_{T_1}{}' + R_{M_1}{}' - R_{M_2}{}')$

Again writing X^{**}=Σ^{-1}MX, D^1 can be written as:

$D^1 = \overline{\Sigma_B^{-1}(I + \rho_B W_B + \rho_B^2 W_B^2 + \cdots)X_B}\ (\widehat{\beta_A}^* - \widehat{\beta_B}^*))\phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$

$= \left[\overline{\Sigma_B^{-1}X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) + \rho_B\overline{\Sigma_B^{-1}W_B X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) + \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\right]\phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$;

[Higher ordered terms involving ρW are ignored in empirical exercises (Lesage & Charles, 2008)] (5.17)

$= \left\{\overline{\Sigma_B^{-1}X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\right\} \times \phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*) + \left\{\rho_B\overline{\Sigma_B^{-1}W_B X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) + \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\right\}\phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$

(5.18)

$=D_{con}^1 + D_{spat}^1$ (5.19)

$D_{con}^1 = \left\{\overline{\Sigma_B^{-1}X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\right\} \times \phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$ is the conventional coefficients effect,[13] decomposing the poverty gap, δ_{pov} in terms of the differential effect of the regional characteristics over the regions.

$D_{spat}^1 = \left\{\rho_B\overline{\Sigma_B^{-1}W_B X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*) + \rho_B^2\overline{\Sigma_B^{-1}W_B^2 X_B}(\widehat{\beta_A}^* - \widehat{\beta_B}^*)\right\}\phi\,(\overline{X_B^{**}}\widehat{\beta_B}^*)$

may be interpreted as the spatial coefficients effects, showing the differential impact of the spatially weighted regional characteristics over the regions.

The poverty gap, δ_{pov} may thus be written using the above relationships (5.9), (5.14) and (5.19) as:

$\delta_{pov} = (C_{con}^1 + D_{con}^1) + (C_{spat}^1 + D_{spat}^1) + (C^R + D^R)$

The first term in bracket constitutes the conventional characteristics and coefficients effects. The second term in bracket constitutes the spatial effect. In case there is no spatial effect, i.e., $\widehat{\rho_A} = 0, \widehat{\rho_B} = 0$, we get the usual non-spatial decomposition analysis.

Detailed Decomposition Analysis of C and D:

[13] As mentioned earlier, in a strict sense it is not the conventional coefficients effect as it involves the term $X_B^1 = \Sigma_B X_B$.

C contains the effects of all the explanatory variables and is the *aggregate characteristics effect*. The contribution of specific factors can be factored out from the overall contribution of all the variables constituting the characteristics effect.

Following (Yun, 2004), the weight (the share of the particular variable in the aggregate characteristics effect) of the k^{th} *explanatory variable* as derived from (5.13) is:

$$V_{\Delta X}^{k} = \frac{\left\{\left(\overline{(\Sigma_A^{-1}X_A)}^{k} - \overline{(\Sigma_B^{-1}X_B)}^{k}\right)\widehat{\beta^*}_A^{k}\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^{*}\right) + \left\{\left(\rho_A\overline{(\Sigma_A^{-1}W_AX_A)}^{k} - \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^{k}\right)\widehat{\beta^*}_A^{k} + \left(\rho_A^2\overline{(\Sigma_A^{-1}W_A^2X_A)}^{k} - \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^{k}\right)\widehat{\beta^*}_A^{k}\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^{*}\right)}{\left\{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right)\widehat{\beta_A}^{*}\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^{*}\right)} \tag{5.20}$$

$$= \frac{\left(\overline{(\Sigma_A^{-1}X_A)^{k}} - \overline{(\Sigma_B^{-1}X_B)^{k}}\right)\widehat{\beta^*}_A^{k} + \left\{\rho_A\overline{\left(\Sigma_A^{-1}W_AX_A\right)}^{k} - \rho_B\overline{\left(\Sigma_B^{-1}W_BX_B\right)}^{k}\right\}\widehat{\beta^*}_A^{k} + \left(\rho_A^2\overline{\left(\Sigma_A^{-1}W_A^2X_A\right)}^{k} - \rho_B^2\overline{\left(\Sigma_B^{-1}W_B^2X_B\right)}^{k}\right)\widehat{\beta^*}_A^{k}}{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right)\widehat{\beta_A}^{*}}. \tag{5.21}$$

Here, $\overline{(.)^k}$ denotes the mean of the k^{th} column of the matrix (.). The characteristic effect due to the k^{th} explanatory variable is, thus given by:

$$C_k = V_{\Delta X}^{k} \times C$$

i.e., $C_k =$

$$\frac{\left\{\left(\overline{(\Sigma_A^{-1}X_A)}^{k} - \overline{(\Sigma_B^{-1}X_B)}^{k}\right)\widehat{\beta^*}_A^{k}\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^{*}\right) + \left\{\left(\rho_A\overline{(\Sigma_A^{-1}W_AX_A)}^{k} - \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^{k}\right)\widehat{\beta^*}_A^{k} + \left(\rho_A^2\overline{(\Sigma_A^{-1}W_A^2X_A)}^{k} - \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^{k}\right)\widehat{\beta^*}_A^{k}\right\} \times \phi\left(\overline{X_A^{**}}\widehat{\beta_A}^{*}\right)}{C^1} C \tag{5.22}$$

Now, as long as C and C^1 have the same signs (*which is the case in the present context*)[14], a positive value of C_k implies that $\forall\ \beta_A^k > 0$ (that is, $\beta^{*k}_A < 0$),

$$\left(\overline{(\Sigma_A^{-1}X_A)^k} - \overline{(\Sigma_B^{-1}X_B)^k}\right)\widehat{\beta}_A^k + \left(\rho_A\overline{(\Sigma_A^{-1}W_AX_A)^k} - \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k}\right)\widehat{\beta}_A^k +$$

$$\left(\rho_A^2\overline{\left(\Sigma_A^{-1}W_A^2X_A\right)}^k - \rho_B^2\overline{\left(\Sigma_B^{-1}W_B^2X_B\right)}^k\right)\hat{\beta}_A^k < 0$$

$$\Rightarrow \overline{(\Sigma_A^{-1}X_A)^k} + \rho_A\overline{(\Sigma_A^{-1}W_AX_A)^k} + \rho_A^2\overline{(\Sigma_A^{-1}W_A^2X_A)^k} < \overline{(\Sigma_B^{-1}X_B)^k} + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k} + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}. \tag{5.23}$$

This signifies that the average spatially adjusted resource level for the k^{th} explanatory variable is scarce in Region A compared to that in Region B. The average availability of the k^{th} explanatory variable in terms of own household characteristic and weighted neighbouring household characteristic is less in Region A compared to Region B. On each side the first part denotes the average availability of own household k^{th} explanatory variable (resource if β_A^k>0,

[14] The idea of opposite signs of C and C^1 is not quite appealing intuitively as explained in the previous chapter.

β_B^k>0) in the region. The other two terms constitute the spatially weighted neighbouring household characteristics for the k^{th} explanatory variable for the region.

A negative value of C_k implies (the other conditions remaining the same) that the average spatially adjusted resource level for the k^{th} explanatory variable is more in Region A compared to that in Region B.

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 (5.18) is:

$$\overset{k}{V}_{\Delta\beta} = \frac{\left[\overline{(\Sigma_B^{-1}X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right] \times \Phi\left(\overline{X_B^{**}}\widehat{\beta_B}^*\right)}{\overline{X_B^{**}}\left(\widehat{\beta_A}^* - \widehat{\beta_B}^*\right)\varphi\left(\overline{X_B^{**}}\widehat{\beta_B}^*\right)} \quad (5.24)$$

$$= \frac{\overline{(\Sigma_B^{-1}X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\overline{X_B^{**}}\left(\widehat{\beta_A}^* - \widehat{\beta_B}^*\right)} \quad (5.25)$$

The coefficient effect due to the k^{th} explanatory variable is, thus given by:

$D_k = \overset{k}{V}_{\Delta\beta} \times D,$ i.e.,

$$D_k = \frac{\left[\overline{(\Sigma_B^{-1}X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)\right] \times \Phi\left(\overline{X_B^{**}}\widehat{\beta_B}^*\right)}{D^1} \mathrm{D}. \quad (5.26)$$

With D and D^1 having the same signs, a positive (negative) value of D_k implies that $\forall\, \beta_A^k, \beta_B^k > 0$ (that is, $\beta^{*k}_A, \beta^{*k}_B < 0$),

$$\overline{(\Sigma_B^{-1}X_B)^k}\left(\hat{\beta}_A^k - \hat{\beta}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k}\left(\hat{\beta}_A^k - \hat{\beta}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}\left(\hat{\beta}_A^k - \hat{\beta}_B^k\right) < (>)0$$

$$\Rightarrow \left(\overline{(\Sigma_B^{-1}X_B)^k} + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k} + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}\right)\left(\hat{\beta}_A^k - \hat{\beta}_B^k\right) < (>)0 \quad (5.27)$$

$\Rightarrow \left(\hat{\beta}_A^k - \hat{\beta}_B^k\right) < (>)0$; [since $\overline{(\Sigma_B^{-1}X_B)^k} + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k} + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k} > 0$]

i.e., as far as the utilization of resource is concerned, Region A is less (more) efficient in utilizing the k^{th} explanatory variable compared to Region B.

<u>The Variance-covariance Structure of ($\boldsymbol{\rho}, \boldsymbol{\beta}^*$):</u>

From the maximum likelihood estimation of the model as given in (5.2), 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.[15]

Now, Asymptotic Variance (ρ, β, σ^2) is given by[16]

$$\begin{bmatrix} tr[W_A]^2 + tr\left[{W_A}^T W_A\right] + \frac{[W_A X\beta]^T[W_A X\beta]}{\sigma^2} & \frac{(X^T W_A X\beta)^T}{\sigma^2} & \frac{tr[W_A]}{\sigma^2} \\ \frac{X^T W_A X\beta}{\sigma^2} & \frac{X^T X}{\sigma^2} & 0 \\ \frac{tr[W_A]}{\sigma^2} & 0 & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \tag{5.28}$$

$$= \begin{bmatrix} A_{1\times1} & B_{1\times K} & C_{1\times1} \\ D_{K\times1} & E_{K\times K} & F_{K\times1} \\ G_{1\times1} & H_{1\times K} & I_{1\times1} \end{bmatrix}, \text{(say)} \tag{5.29}$$

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

Now $\beta^* = -\beta$; hence,

Asymptotic Variance (ρ, β^*)

$$= \begin{bmatrix} \frac{\delta}{\delta\rho}(\rho \quad -\beta) \\ \frac{\delta}{\delta\beta}(\rho \quad -\beta) \end{bmatrix} \text{Asymptotic Variance } (\rho, \beta) \begin{bmatrix} \frac{\delta}{\delta\rho}(\rho \quad -\beta) \\ \frac{\delta}{\delta\beta}(\rho \quad -\beta) \end{bmatrix}^T \tag{5.30}$$

= Asymptotic Variance (ρ, β)

From the estimated variance–covariance structure of the coefficients of spatial regression model in (5.2), the variances of C and D are found using the delta method and the asymptotic variance of C_k, $\sigma^2_{C_K}$ and that of D_k, $\sigma^2_{D_K}$ are computed as in Chapter 3.[17]

Testing the Significance of C and D:

The test statistic under the null hypothesis (C=0, D=0) are respectively given by $t_C = \frac{C}{\sigma_C}$ and $t_D = \frac{D}{\sigma_D}$, which are asymptotically normally distributed. The hypothesis testing can be implemented in a similar fashion at the individual variable level.

The test statistic under the null hypothesis (C_k=0, D_k=0) are, respectively, given by $t_{C_k} = \frac{C_k}{\sigma_{C_k}}$ *and* $t_{D_k} = \frac{D_k}{\sigma_{D_k}}$, which are asymptotically normally distributed.

---

[15] See (Powell L. A., 2007); (Seber, 1982).

[16] See (Anselin & Bera, 1998).

[17] The exact forms of σ^2_C, σ^2_D, $\sigma^2_{C_k}$ and $\sigma^2_{D_k}$ are given in Appendix A5.1 – A5.4.

5.3 Data and Results

The explanatory variables and the dependent variable are the same as in Chapter 3. North Bengal is taken to be Region A and South Bengal is taken to be Region B. The distinctive feature of this 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 (diagonal elements of the weight matrix are set to zero).

For incorporating the household specific spatial effects based on adjacency criterion[18] 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[19], spatial weight being equal to zero, if otherwise.
2. For households living in different districts, spatial weight is equal to unity if both of 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 categorization of the districts in terms number of adjacent districts has been shown in Table 5.1 and Table 5.2, separately for North Bengal and South Bengal. The map of West Bengal is given in the Appendix A2.9 (Figure A2.1).

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). Using two measures of spatial autocorrelation: Moran's I (Moran, 1948) and Geary's C (Geary, 1954)[20],

---

[18]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 & Gleditsch, 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.

[19] The sampled households of rural West Bengal are obtained using a two-stage stratified sampling design. The first stage strata are the districts and the three second stage strata are as follows:

SSS 1: relatively affluent households

SSS 2: Households not belonging to SSS1and having principal earning from non- agricultural activity

SSS 3: other households

[20] A short description about the measures is given in Appendix A5.5.

presence of strong spatial autocorrelation in the data is suggested. The whole analysis as implemented in Chapter 3 to find the difference in the incidences of poverty between North Bengal and South Bengal is repeated in the framework of spatial regression using a spatial autoregressive lag model. The important findings from the analysis are:

1. The incidence of poverty is slightly higher (compared to the non-spatial analysis in Chapter 3) in the spatial analysis, particularly for South Bengal. There is marked difference (0.13) in the incidences of poverty between North Bengal and South Bengal, the difference being more or less the same (0.14) as in Chapter 3.
2. As found from Tables 5.3 and 5.4, within each region the two measures of spatial autocorrelation ($\mathrm{I} > E(\mathrm{I}), C < 1$) suggest that significant spatial effect is present in the incidence of poverty.
3. Tables 5.7 and 5.8 show that the lag coefficient (ρ) is significant 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. The share of aggregate characteristics effect is much higher (124%) compared to that of the aggregate coefficients effect (-24%) (see Table 5.9).

It is to be noted that the aggregate characteristics effect and the aggregate coefficients effect had a more or less balanced share in the case of non-spatial analysis as performed in Chapter 3. The much higher share of the aggregate characteristics effect signify that poverty disparity between North Bengal and South Bengal is indeed a result of a resource disparity than of the differential capacity of utilization of the resources. From Table 5.10, showing the (spatially weighted) average resource vectors in North and South Bengal, it may be observed that the severity of the deficiency is captured more closely using a spatial analysis.[21]

The summary of the findings of the spatial analysis is that (Table 5.11) North Bengal has a resource and efficiency deficiency, specifically in terms of the *educational status.* By enhancing the educational status and the associated coefficient vector in North Bengal poverty gap can be eliminated respectively, by 107% and 93%. This is a finding in line with Chapter 3, but the spatial approach provides a closer picture of the actual situation in terms of severity of the deficiency.[22] The aggregate coefficients effect turns out to be negative and

---

[21] A negative share of aggregate coefficients effect indicates that North Bengal will be worse off if coefficients are equalized over the regions, as explained in Chapter 3.

[22] A comparison of Table 5.10 with Table 3.5 (Chapter 3) shows the increased severity of the deficiency in the average resource level of North Bengal (due the incorporation of the spatial effect).

non-significant (at 5% level). This is in contrast to our earlier findings in Chapter 3, which showed a positive and significant coefficients effect. Thus, the spatial analysis, while enhancing the characteristics effect, reduces the coefficients effects to a negligible one, reinforcing the importance of the characteristics effect in explaining the poverty gap and indicating scarcity of resource vector in North Bengal.

5.4 Conclusion

The distinctive feature of the present chapter is the generalisation of the Oaxaca decomposition technique in the poverty incidence analysis using a spatial framework. The traditional decomposition analysis can be regarded as a special case of the present analysis.

It is noted from the present analysis that there is marked difference in the shares and magnitudes of characteristics effect and coefficients effect from those obtained in the non-spatial analysis in Chapter 3. In the non-spatial decomposition, the aggregate characteristics and aggregate coefficients effect had a more or less balanced share signifying the need for enhancement of both efficiency and resource levels in North Bengal. The analysis of this chapter reveals that resource scarcity in North Bengal is the dominant factor influencing the difference in incidences of poverty between the two parts of West Bengal. A limitation of this spatial model, however, is that the manner in which spatial effect has been modeled here, may give rise to "reflection problem" ((Manski, 1993), (Pinkse & $Slade$, 2009)). It may, therefore, be possible that the insignificance of the coefficients effect in the spatial model is due to spatial correlation of the coefficients.

TABLES

Table 5.1 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 5.2 Districts of South Bengal (Region B)*

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

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

Table 5.3 Tests of Spatial Autocorrelation for North Bengal (Region A)

(Variable: $\boldsymbol{ln}\left(\frac{y}{z}\right)$)

Variable $\boldsymbol{ln}\left(\frac{y}{z}\right)$	I	$E(I)$	$sd(I)$	z	p-value (one tailed test)
(1)	(2)	(3)	(4)	(5)	(6)
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 5.4 Tests of Spatial Autocorrelation for South Bengal (Region B)

(Variable: $\boldsymbol{ln}\left(\frac{y}{z}\right)$)

Variable $\boldsymbol{ln}\left(\frac{y}{z}\right)$	I	$E(I)$	$sd(I)$	z	p-value (one tailed test)
(1)	(2)	(3)	(4)	(5)	(6)
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

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

Table 5.5 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*	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	PLAND	0.0004	7.0275*	0.0003	7.5681*
Labour market characteristics	POWNAC	0.4575	7.9319*	0.4004	8.2792*
	PNSCH**	-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*

* indicates significance at 5% level.

** As explained earlier in Chapter 3 the influence of this variable is expected to be negative.

Table 5.6 Poverty Incidences in North and South Bengal (Spatial Framework)

	Sample size	Spatial rho (ρ)	Spatial sigma (σ)	Poverty incidence
(1)	(2)	(3)	(4)	(5)
North Bengal	1526	0.4215	0.3100	$H_A = 0.3127$
South Bengal	3407	0.5945	0.3400	$H_B = 0.1849$
	Difference in poverty incidence: H_A-H_B = 0.1278			

Table 5.7 Statistical Test for Spatial Rho (ρ_A): North Bengal

Coefficient (ρ)	Standard error	z	P>|z|	95% Confidence Interval
(1)	(2)	(3)	(4)	(5)
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 5.8 Statistical Test for Spatial Rho (ρ_B): South Bengal

Coefficient (ρ)	Standard error	z	P>|z|	95% Confidence Interval
(1)	(2)	(3)	(4)	(5)
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 5.9 Decomposing the Difference of Poverty Incidences: (HA-HB) between North Bengal (Region A) and South Bengal (Region B)

(Using Estimates of Table 8)

			Aggregate Characteristics effect (C)			Aggregate Coefficients effect(D)		
Aggregate effect***			Estimate	Share in (H_A-H_B)**		Estimate	Share in (H_A-H_B)	
			0.1583 (0.0128)*	123.96		-0.0306 (0.0375)	-23.96	
Decomposition of the Aggregate effect			Individual Characteristic effect (C_k)' s			Individual Coefficient effect (D_k)'s		
			Estimates	Shares in (H_A-H_B)**		Estimates	Shares in (H_A-H_B)	
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	Demographic characteristics of the households	1-DEPRAT	0.0939 (0.0026)	73.49	73.88	-0.0902 (0.0063)	-70.64	-71.8
		D_FEMH	0.0005 (0.000008)	0.39		-0.0015 (0.0002)	-1.16	
II	Educational status of the household	PSECEDU	0.0035 (0.0001)	2.78	107.36	0.0117 (0.0036)	9.18	93.09
		PTERTEDU	0.0172 (0.0008)	13.47		-0.0089 (0.0015)	-6.94	
		GENEDU	0.1164 (0.0003)	91.11		0.1161 (0.0007)	90.85	
III	Wealth status	PLAND	0.0140 (0.0000005)	10.95	10.95	-0.0093 (0.000001)	-7.27	-7.27
IV	Labour market characteristics	POWNAC	0.0290 (0.001)	22.69	60.36	-0.0131 (0.0031)	-10.27	0.62
		PNSCH	-0.0158 (0.0007)	-12.38		0.0076 (0.0016)	5.94	
		PDOMO	0.0304 (0.0011)	23.83		-0.0048 (0.0026)	-3.76	
		PDOM	0.0218 (0.0009)	17.03		-0.0152 (0.0023)	-11.87	
		PEMP	0.0117 (0.0004)	9.19		0.0263 (0.0014)	20.58	
V	Government aid	D_GOVAID	0.0098 (0.0002)	7.69	7.69	-0.0067 (0.0005)	-5.27	-5.27
	Constant		-0.1741 (0.0032)	-136.28	-136.28	-0.0426 (0.0031	-33.31	-33.31

* Figures in parentheses denote the standard errors. The aggregate characteristics effect is significant at 5% level. But the aggregate coefficients effect is not significant at 5% level. All the individual characteristic and coefficient effects are highly significant at 5% level.

**share = estimates/ Difference in poverty incidence (Table 5.6)) × 100

*** Aggregate Characteristics effect, $C = \sum_{k=1}^{K} C_k$ and Aggregate Coefficients effect, $D = \sum_{k=1}^{K} D_k$

Table 5.10 Observed Resource Vectors for North Bengal and South Bengal*

Characteristics		Variables under characteristics	$\bar{X}^k$	
			Region A (North Bengal) $\bar{X}_A^k$	Region B (South Bengal) $\bar{X}_B^k$
(1)		(2)	(3)	(4)
I	Demographic characteristics of the households	1-DEPRAT	3.4942	4.7868
		D_FEMH	0.6102	0.6292
II	Educational status of the household	PSECEDU	0.2393	0.4199
		PTERTEDU	0.2442	0.4422
		GENEDU	19.2466	30.5555
III	Wealth status	PLAND	459.9241	593.4503
V	Labour market characteristics	POWNAC	0.8662	1.1169
		PNSCH	0.2943	0.4720
		PDOMO	0.8919	1.3036
		PDOM	0.4191	0.6310
		PEMP	0.8958	1.1403
V	Government aid	D_GOVAID	0.3931	0.6004

* Resource vectors considered here is the spatially weighted resource vector

Table 5.11 A Comparative Summarization of the Spatial and Non-spatial Analysis

		Ordinary Regression Analysis	Spatial Regression Analysis
(1)	(2)	(3)	(4)
Incidence of poverty	North Bengal	0.3108	0.3127
	South Bengal	0.1714	0.1849
Difference of poverty incidences			
Oaxaca Decomposition of the difference of poverty incidences	Share* of Aggregate Characteristics effect	40.9	123.96
	Share of Aggregate Coefficients effect	59.1	-23.96
Individual Characteristic effect (C_k)' s		**Ordinary Regression Analysis**	**Spatial Regression Analysis**
Demographic characteristics of the households		6.4	73.88
Educational status of the household		27.1	107.36
Wealth status		0.7	10.95
Labour market characteristics		4.5	60.36
Government aid		2.3	7.69
Constant		-	-136.28
Individual Coefficient effect (D_k)'s			
Demographic characteristics of the households		-65.1	-71.8
Educational status of the household		29.2	93.09
Wealth status		-7.7	-7.27
Labour market characteristics		-6.6	0.62
Government aid		-2.9	-5.27
Constant		112.2	-33.31

* Share is given in percentages

APPENDICES

Appendix A5.1 To find $\frac{\partial C}{\partial \widehat{\rho_A}}$

C= $f\left(\widehat{\rho_A},\ \widehat{\beta_A}^*, \widehat{\rho_B}\right)$; [from (5.12)]

Hence, the asymptotic variance of C,

$$\sigma_C^2 = \left(\frac{\partial C}{\partial \widehat{\rho_A}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^*}\right) (\text{Asymptotic Variance } (\rho_A, \beta_A{}^*)) \left(\frac{\partial C}{\partial \widehat{\rho_A}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial C}{\partial \widehat{\rho_B}}\right) (\text{Asymptotic Variance } (\rho_B)) \left(\frac{\partial C}{\partial \widehat{\rho_B}}\right)^T$$

$$\Rightarrow \sigma_C^2 = \left(\frac{\partial C}{\partial \widehat{\rho_A}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^*}\right) \begin{bmatrix} A_{1\times1} & B_{1\times K} \\ D_{K\times1} & E_{K\times K} \end{bmatrix}^A \left(\frac{\partial C}{\partial \widehat{\rho_A}} \quad \frac{\partial C}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial C}{\partial \widehat{\rho_B}}\right) (\text{Asymptotic Variance } (\rho_B)) \left(\frac{\partial C}{\partial \widehat{\rho_B}}\right)^T$$

C= $\overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)}$; [from (5.10)]

$$\Rightarrow C = \overline{\Phi\left\{\left(\Sigma_A^{-1}(I + \rho_A W_A + \rho_A^2 W_A^2 + \cdots)X_A\right)\widehat{\beta_A}^*\right\}} - \overline{\Phi\left\{\left(\Sigma_B^{-1}(I + \rho_B W_B + \rho_B^2 W_B^2 + \cdots)X_B\right)\widehat{\beta_A}^*\right\}}$$

$$= \overline{\Phi\left(\Sigma_A^{-1}X_A\widehat{\beta_A}^* + \rho_A\Sigma_A^{-1}W_AX_A\widehat{\beta_A}^* + \rho_A^2\Sigma_A^{-1}W_A^2X_A\widehat{\beta_A}^*\right)} -$$

$$\overline{\Phi\left(\Sigma_B^{-1}X_B\widehat{\beta_A}^* + \rho_B\Sigma_B^{-1}W_BX_B\widehat{\beta_A}^* + \rho_B^2\Sigma_B^{-1}W_B^2X_B\widehat{\beta_A}^*\right)}$$

$$\frac{\partial C}{\partial \widehat{\rho_A}} = \overline{\left(\Sigma_A^{-1}W_AX_A\widehat{\beta_A}^* + 2\rho_A\Sigma_A^{-1}W_A^2X_A\widehat{\beta_A}^*\right)\phi\left(\Sigma_A^{-1}X_A\widehat{\beta_A}^* + \rho_A\Sigma_A^{-1}W_AX_A\widehat{\beta_A}^* + \rho_A^2\Sigma_A^{-1}W_A^2X_A\widehat{\beta_A}^*\right)}$$

(Here over bar means sample average).

Appendix A5.2 To find $\frac{\partial C}{\partial \widehat{\beta_A}^*}$

C=$\overline{\Phi\left(X_A^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)}$; [from (5.9.1)]

$\frac{\partial C}{\partial \widehat{\beta_A}^*} = \vec{\nabla}\,C$, where $\vec{\nabla}=\left(\frac{\partial C}{\partial \beta^{*1}_A}, \frac{\partial C}{\partial \beta^{*2}_A}, \dots, \frac{\partial C}{\partial \beta^{*K}_A}\right)$

$=\left[\left(\frac{1}{n^A}\sum_{i=1}^{n^A} X^{**}{}_{i_A}{}^{k}\phi\left(X^{**}{}_{i_A}\widehat{\beta_A}^*\right) - \frac{1}{n^B}\sum_{i=1}^{n^B} X^{**}{}_{i_B}{}^{k}\phi\left(X^{**}{}_{i_B}\widehat{\beta_A}^*\right)\right)\right]_{1\times K}$; (as in Appendix A3.2)

$=\left[\left(\overline{X^{**}{}_A{}^{k}\phi\left(X^{**}{}_A\widehat{\beta_A}^*\right)} - \overline{X^{**}{}_B{}^{k}\phi\left(X^{**}{}_B\widehat{\beta_A}^*\right)}\right)\right]_{1\times K}$;

$X^{**}{}_A{}^{k}$ and $X^{**}{}_B{}^{k}$ are the k[th] columns of the matrices $X^{**}{}_A$ and $X^{**}{}_B$ respectively.

Appendix A5.3 To find $\left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right)$ & $\left(\frac{\partial D}{\partial \widehat{\beta_B}^*}\right)$

Using (5.17), D = $f\left(\widehat{\beta_A}^*; \widehat{\rho_B}, \widehat{\beta_B}^*\right)$

Thus, the asymptotic variance of coefficients effect

$$\sigma_D^2 = \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right) \text{Asymptotic Variance } ({\beta_A}^*) \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right)^T$$

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

$$= \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right) [E_{K\times K}]^A \left(\frac{\partial D}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial D}{\partial \widehat{\rho_B}} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right) \begin{bmatrix} A_{1\times 1} & B_{1\times K} \\ D_{K\times 1} & E_{K\times K} \end{bmatrix}^B \left(\frac{\partial D}{\partial \widehat{\rho_B}} \quad \frac{\partial D}{\partial \widehat{\beta_B}^*}\right)^T;$$

[using (5.29)]

Now,

$$\text{D} = \overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}$$

Writing in terms of individual observations,

$$D = \left[\frac{1}{n^B}\sum_{i=1}^{n^B} \Phi\left({X^{**}}_{i_B}\widehat{\beta_A}^*\right)\right] - \left[\frac{1}{n^B}\sum_{i=1}^{n^B} \Phi\left({X^{**}}_{i_B}\widehat{\beta_B}^*\right)\right];$$

Thus,

$$\frac{\partial D}{\partial \widehat{\beta_A}^*} = \vec{\nabla}\,\text{D}, \text{ where } \vec{\nabla}\text{D} = \left(\frac{\partial D}{\partial {\beta^*}_A^1}, \frac{\partial D}{\partial {\beta^*}_A^2}, \dots, \frac{\partial D}{{\beta^*}_A^K}\right)$$

$$\left(\frac{\partial \boldsymbol{D}}{\partial \widehat{\boldsymbol{\beta}_A}^*}\right) = \left[\left(\frac{1}{n^B}\sum_{i=1}^{n^B} {{X^{**}}_{i_B}}^k \phi\left({X^{**}}_{i_B}\widehat{\beta_A}^*\right)\right)\right]_{1\times K}$$

$$= \left[\left(\overline{{X_B^{**}}^k \phi\left(X_B^{**}\widehat{\beta_A}^*\right)}\right)\right]_{1\times K}; \; {X_B^{**}}^k \text{ is the k}^{\text{th}} \text{ column of the matrix } X_B^{**}.$$

Similarly,

$$\left(\frac{\partial \boldsymbol{D}}{\partial \widehat{\boldsymbol{\beta}_B}^*}\right) = -\left[\left(\overline{{X_B^{**}}^k \phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right)\right]_{1\times K}$$

Appendix A5.4 To find $\frac{\partial C_k}{\partial \widehat{\beta_A}^*}$

$$\sigma^2_{C_k} = \left(\frac{\partial C_k}{\partial \widehat{\rho_A}} \quad \frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right) (\text{Asymptotic Variance } \rho_A, \beta_A{}^*)) \left(\frac{\partial C_k}{\partial \widehat{\rho_A}} \quad \frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial C_k}{\partial \widehat{\rho_B}}\right) (\text{Asymptotic Variance } (\rho_B)) \left(\frac{\partial C_k}{\partial \widehat{\rho_B}}\right)^T$$

$$\Rightarrow \ \sigma^2_{C_k} = \left(\frac{\partial C_k}{\partial \widehat{\rho_A}} \quad \frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right) \begin{bmatrix} A_{1\times 1} & B_{1\times K} \\ D_{K\times 1} & E_{K\times K} \end{bmatrix}^A \left(\frac{\partial C_k}{\partial \widehat{\rho_A}} \quad \frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right)^T + \left(\frac{\partial C_k}{\partial \widehat{\rho_B}}\right) (\text{Asymptotic Variance } (\rho_B)) \left(\frac{\partial C_k}{\partial \widehat{\rho_B}}\right)^T$$

Now,

$$C_k = \left[\frac{\left\{\left(\overline{(\Sigma_A^{-1} X_A)}^k - \overline{(\Sigma_B^{-1} X_B)}^k\right) \widehat{\beta^*}_A^k\right\} + \left\{\left(\rho_A \overline{(\Sigma_A^{-1} W_A X_A)}^k - \rho_B \overline{(\Sigma_B^{-1} W_B X_B)}^k\right) \widehat{\beta^*}_A^k + \left(\rho_A^2 \overline{(\Sigma_A^{-1} W_A^2 X_A)}^k - \rho_B^2 \overline{(\Sigma_B^{-1} W_B^2 X_B)}^k\right) \widehat{\beta^*}_A^k\right\}}{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right) \widehat{\beta_A}^*}\right] \times$$

$$\left\{\overline{\Phi\left(X_A^{**} \widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**} \widehat{\beta_A}^*\right)}\right\}$$

Hence,

$$\left(\frac{\partial C_k}{\partial \widehat{\beta_A}^*}\right) = \left\{\overline{\Phi\left(X_A^{**} \widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**} \widehat{\beta_A}^*\right)}\right\} \left[\left(\frac{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right) \widehat{\beta_A}^* Z_l - Z\left(\overline{X_A^{**}}^l - \overline{X_B^{**}}^l\right)}{\left(\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right) \widehat{\beta_A}^*\right)^2}\right)\right]_{1\times K}$$

$$+ \left\{\frac{\left(\overline{(\Sigma_A^{-1} X_A)^k} - \overline{(\Sigma_B^{-1} X_B)^k}\right) \widehat{\beta^*}_A^k + \left(\rho_A \overline{(\Sigma_A^{-1} W_A X_A)}^k - \rho_B \overline{(\Sigma_B^{-1} W_B X_B)}^k\right) \widehat{\beta^*}_A^k + \left(\rho_A^2 \overline{(\Sigma_A^{-1} W_A^2 X_A)}^k - \rho_B^2 \overline{(\Sigma_B^{-1} W_B^2 X_B)}^k\right) \widehat{\beta^*}_A^k}{\left(\overline{X_A^{**}} - \overline{X_B^{**}}\right) \widehat{\beta_A}^*}\right\} \frac{\partial C}{\partial \widehat{\beta_A}^*} \quad \text{; where}$$

$$Z_l = \left[\left(\overline{(\Sigma_A^{-1} X_A)^k} - \overline{(\Sigma_B^{-1} X_B)^k}\right) + \left(\rho_A \overline{(\Sigma_A^{-1} W_A X_A)}^k - \rho_B \overline{(\Sigma_B^{-1} W_B X_B)}^k\right) + \left(\rho_A^2 \overline{(\Sigma_A^{-1} W_A^2 X_A)}^k - \rho_B^2 \overline{(\Sigma_B^{-1} W_B^2 X_B)}^k\right)\right] \delta_l^k \ ;$$

$\delta_l^k = 1$ if l=k
$= 0$, if $l \neq$k

Appendix A5.5: To find $\left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right)$ & $\left(\frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right)$

$$\sigma^2_{D_k} = \left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right) \text{Asymptotic Variance } ({\beta_A}^*) \left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right)^T$$

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

$$\Rightarrow \sigma^2_{D_k} = \left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right) [E_{K\times K}]^A \left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right)^T$$

$$+ \left(\frac{\partial D_k}{\partial \widehat{\rho_B}} \quad \frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right) \begin{bmatrix} A_{1\times 1} & B_{1\times K} \\ D_{K\times 1} & E_{K\times K} \end{bmatrix}^B \left(\frac{\partial D_k}{\partial \widehat{\rho_B}} \quad \frac{\partial D_k}{\partial \widehat{\beta_B}^*}\right)^T$$

Now,

$$D_k = \left\{\frac{\overline{(\Sigma_B^{-1}X_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B)}{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)(\widehat{\beta^*}_A - \widehat{\beta^*}_B)}\right\}\left\{\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right\}.$$

Hence,

$$\left(\frac{\partial D_k}{\partial \widehat{\beta_A}^*}\right) = \left\{\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right\} \times$$

$$\left[\frac{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)(\widehat{\beta^*}_A - \widehat{\beta^*}_B)U_l - U\left(\left(\overline{\Sigma_B^{-1}X_B}\right)^l + \rho_B\left(\overline{\Sigma_B^{-1}W_BX_B}\right)^l + \rho_B^2\left(\overline{\Sigma_B^{-1}W_B^2X_B}\right)^l\right)}{\left(\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)(\widehat{\beta^*}_A - \widehat{\beta^*}_B)\right)^2}\right]_{1\times K}$$

$$+\left\{\left\{\frac{\overline{(\Sigma_B^{-1}X_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k(\widehat{\beta^*}^k_A - \widehat{\beta^*}^k_B)}{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)(\widehat{\beta^*}_A - \widehat{\beta^*}_B)}\right\}\right\}\frac{\partial D}{\partial \widehat{\beta_A}^*};$$

$$U_l = \left(\overline{(\Sigma_B^{-1}X_B)^k} + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k} + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}\right)\delta_l^k$$

$\delta_l^k = 1$ if $l = \text{k}$

$= 0$, if $l \neq \text{k}$

Again,

$$D_k = \left\{\frac{\overline{(\Sigma_B^{-1}X_B)^k}\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k}\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)}\right\}\times$$

$$\left\{\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right\}$$

$$\Rightarrow\left(\frac{\partial D_k}{\partial\widehat{\beta_B}^*}\right)$$

$$=\left[\frac{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\left(\frac{\partial U}{\partial\widehat{\beta_B}^*}\right) + U\left(\left(\overline{\Sigma_B^{-1}X_B}\right)^l + \rho_B\left(\overline{\Sigma_B^{-1}W_BX_B}\right)^l + \rho_B^2\left(\overline{\Sigma_B^{-1}W_B^2X_B}\right)^l\right)}{\left(\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)\right)^2}\right]\times$$

$$\left\{\overline{\Phi\left(X_B^{**}\widehat{\beta_A}^*\right)} - \overline{\Phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right\} + \left\{\left\{\frac{\overline{(\Sigma_B^{-1}X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right) + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)}^k\left(\widehat{\beta^*}_A^k - \widehat{\beta^*}_B^k\right)}{\left(\overline{\Sigma_B^{-1}X_B} + \rho_B\overline{\Sigma_B^{-1}W_BX_B} + \rho_B^2\overline{\Sigma_B^{-1}W_B^2X_B}\right)\left(\widehat{\beta^*}_A - \widehat{\beta^*}_B\right)}\right\}\right\}\frac{\partial D}{\partial\widehat{\beta_B}^*};$$

$$\left(\frac{\partial U}{\partial\widehat{\beta_B}^*}\right) = -\left(\overline{(\Sigma_B^{-1}X_B)^k} + \rho_B\overline{(\Sigma_B^{-1}W_BX_B)^k} + \rho_B^2\overline{(\Sigma_B^{-1}W_B^2X_B)^k}\right)\delta_l^k;$$

$$\left(\frac{\partial D}{\partial\widehat{\beta_B}^*}\right) = -\left[\left(\overline{X_B^{**k}\phi\left(X_B^{**}\widehat{\beta_B}^*\right)}\right)\right]_{1\times K},$$ [from A5.3] ;

$\delta_l^k = 1$ if $l = \mathrm{k}$

$= 0$, if $l \neq \mathrm{k}$

Spatial autocorrelation can be defined as the phenomenon that occurs when the spatial distribution of the variable of interest exhibits a systematic pattern (Cliff & Ord, 1981).
Two measures of spatial autocorrelation have been analyzed in this chapter, viz., Moran's I (Moran, 1948), Geary's C (Geary, 1954).

Moran's I:

Moran's I is defined as
$I = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}Z_iZ_j}{S_0 m_2}$,
N being the number of spatial units indexed by i and j. w_{ij} denotes elements of spatial weight matrix W corresponding to the location pair (i,j). $Z_i = Y_i - \bar{Y}$; Y_i denotes the value taken on the variable Y at location i; $\bar{Y}$ denotes the mean of the variable Y. $S_0 = \sum_i \sum_j w_{ij}$; $m_2 = \sum_i \frac{Z_i^2}{N}$
Under the null hypothesis of no global spatial autocorrelation, the expected value of I is given by $E(I) = -\frac{1}{N-1}$.

If I is larger than its expected value, then the overall distribution of variable Y can be seen as characterized by positive spatial autocorrelation, meaning that the value taken on by Y at each location i tends to be similar to the values taken on by Y at spatially contiguous locations. On the other hand, if I is smaller than its expected value, then the overall distribution of variable can be seen as characterized by negative spatial autocorrelation, meaning that the value taken on by Y at each location i tends to be different from the values taken on by Y at spatially contiguous locations. Inference is based on z-values, computed by subtracting $E(I)$ from I and dividing the result by the standard deviation of I.

$Z_I = \frac{I - E(I)}{sd(I)}$;

Now,

$VAR(I) = \frac{NS_4 - S_3S_5}{(N-1)(N-2)(N-3)(\sum_i \sum_j w_{ij})^2}$, where

$S_4 = (N^2 - 3N + 3)S_1 - NS_2 + 3\left(\sum_i \sum_j w_{ij}\right)^2$;

$S_3 = \frac{N^{-1}\sum_i (x_i - \bar{x})^4}{(N^{-1}\sum_i (x_i - \bar{x})^2)^2}$;

$S_5 = S_1 - 2NS_1$;

$S_1 = \frac{1}{2}\sum_i \sum_j (w_{ij} + w_{ji})^2$;

$S_2 = \sum_i \left(\sum_j w_{ij} + \sum_i w_{ji}\right)^2$

<u>Geary's C:</u>

Geary's C is computed by,

$C = (N-1)\ \frac{\sum_i \sum_j w_{ij}(z_i - z_j)}{2NS_0 m_2}$

Under the null hypothesis of no global spatial autocorrelation, the expected value of C equals 1. If C is larger than 1, then the overall distribution of variable Y can be seen as characterized by negative spatial autocorrelation; on the other hand, if C is smaller than 1, then the overall distribution of variable Y can be seen as characterized by positive spatial autocorrelation.

As in the case of Moran's I, inference is based on z-values, computed by subtracting 1 from c and dividing the result by the standard deviation of C.

For details and formulas, see (Sokal, Oden, & Thomson, 1998), which implies that Z_C is asymptotically distributed as a standard normal variate (Anselin L. , 1992).

CHAPTER 6

CONCLUDING REMARKS

The thesis has addressed the problem of estimation of poverty at the district level in the Indian context using NSS data, with particular reference to the rural sector of the state of West Bengal. It has attempted to analyse the problem of poverty from the perspective of policy formulation using methods proposed herein and also using existing econometric methods. The analysis here is spatial in nature, as it involves cross sectional comparison of districts (without any time dimension).

Chapter 1 has dealt with inadequacy of the data at the district level and proposes a method for tackling the problem. The performances of the proposed and conventional poverty estimates have been compared. It is observed that, in general, in majority of the cases the Relative Standard Error (RSE), computed as the ratio of the standard error and the point estimate, is less for the proposed method than the corresponding RSE of the conventional method. This is more clearly observed for cases with higher discrepancy between the two estimates, which occur when the sample sizes are small. This indicates that for such districts the proposed method yields better estimates.

Next, the issue of *spatial aspect* of poverty with respect to price has been addressed in Chapter 2. One source of spatial variation in poverty estimates is the spatial difference in prices. To examine the extent to which the spatial difference in prices affect the poverty estimates, Chapter 2 has proposed a method of estimating spatial price indices, using which district level price indices (with state as base) and the corresponding district level poverty lines are obtained. The method does not require item-specific price or unit-value data and hence overcomes the problem of data inadequacy in the context of prices. Estimates of district level poverty based on *district level poverty lines* and those using the *conventional state level poverty line* are compared. It is noted that there is ample variation in values of the estimated indices across the districts as well as across definitions of 'region'. It is also observed that taking 'West Bengal' as numeraire, the Northern districts have lower price levels and the Southern districts have higher price levels. It is interesting to note that except for few districts this segregation largely coincides with the traditional (geographical) division of North and South Bengal with respect to River Hooghly.

An alternative source of spatial variation in estimates of poverty is the geographically segregated units characterized by their intrinsic nature of development status (level of living). Given the fact that there is considerable difference in the levels of economic well being in two parts of Bengal mentioned above, viz., North and South Bengal, Chapter 3 has identified the sources and characteristics affecting the differential levels of economic well being (poverty) in the two parts. The results of the Oaxaca decomposition analysis suggest that resource deficiency is an important factor that makes North Bengal poorer compared to South Bengal. This suggests policy prescription in terms of enhancement of resource level in North Bengal, specifically in terms of educational attainment.

Chapter 4 has introduced the earnings frontier approach in explaining monthly consumption expenditure (a proxy for income) in terms of *human capital* and *endowments* of a household and has examined an alternative source of spatial variation in the incidences of poverty on the basis of efficiency based segregation. Splitting the sample into an *efficient* and *inefficient* part based on the estimated frontier, the status of poverty in the two groups has been studied using the Oaxaca decomposition of the poverty gap. An important finding in this chapter is that the group with higher incidence of poverty has a higher level of efficiency. This corroborates the poor but efficient doctrine postulated in (Chong, Lizarondo, Cruz, Guerrero, & Smith, 1984) and also supported in (Schultz T. W., 1965). The essence of the findings in this chapter is that rural poverty is mainly a problem originating from the scarcity of resources. So far as the specific resource is concerned, the average household general educational level is the largest contributor in influencing the difference in the incidences of poverty between the groups and enhancement of the education level for the efficient group will be a possible solution to narrowing the difference of poverty levels between the two groups. It also emerges that geographical location of households does have significant impact in explaining the poverty gap.

Chapter 5 has looked into the problem of poverty from the perspective of spatial regression analysis and in a sense it is spatial reformulation of Chapter 3 introducing the spatial autoregressive dependence in the monthly consumption expenditure values within North and South Bengal. A comparison of the results with those of Chapter 3 shows that there is marked difference in the shares and magnitudes of characteristics effect and coefficients effect from those obtained in the non-spatial analysis in Chapter 3. The summary of the findings of the spatial analysis is that North Bengal has a resource and efficiency deficiency, specifically in terms of the *educational status.* This is a finding in line with

Chapter 3, but the spatial approach provides a closer picture of the actual situation in terms of severity of the deficiency. To be specific, the spatial analysis, while enhancing the characteristics effect, reduces the coefficients effects to a negligible one, reinforcing the importance of the characteristics effect in explaining the poverty gap and indicating scarcity of resource vector in North Bengal.

APPENDIX

<u>*Sampling Design of the NSS Employment – Unemployment Schedule:*</u>

The thesis is based on unit-level employment–unemployment data collected by the Indian National Sample Survey Organization in their 55th and 61st round of survey operations, relating to the periods 1999-2000 and 2004-2005, respectively. A short description of the structure of NSS data (rural) used in this study is provided in this appendix. It may be mentioned here that while Chapter 1 is based on the 55th and 61st round data, the rest of the chapters are based on the 61st round of data only.

<u>*Sampling design of NSS 55th Round (Rural):*</u>

A two stage stratified sampling design is adopted in the 55th round for selection of the sample first stage units (FSU's), which are villages for the rural areas. Two special strata are formed in the *first stage* at the state level, viz.,

Stratum 1: all FSU's with population between 1 and 100,

Stratum 2: FSU's with population more than 15,000.

These strata are formed if at least 50 FSU's are there in the respective frames. Otherwise they are merged with the general strata. While forming the general strata (consisting of FSU's other than those covered under strata 1 and 2) efforts have been made to treat each district as a separate stratum. If limitation of sample size does not allow forming so many strata, smaller districts within a particular NSS region are merged to form a stratum. The round, covering the period July 1999 – June 2000, is made up of 4 sub-rounds, mainly representing four seasons. Sample size for the whole round for each State is allocated equally among the 4 sub-rounds. For each sub-round, sample FSU's from each stratum are selected in the form of 2 independent sub-samples by following circular systematic sampling with (a) probability proportional to population for all strata other than stratum 1, and (b) equal probability for stratum 1.

Households, which are the ultimate stage units (USU's), are selected (10-12 in number) from the FSU's using a *second stage stratification* (SSS) of 'Affluent households' and 'Rest of the households'. Information on all members of a household are collected through interviewing a representative from each household in the sample.

Sampling design of NSS 61st Round (Rural):

The NSS 61st round survey is characterized by a multi-stage design with the sample first stage units (FSU) being the villages for the rural areas and the ultimate stage units (USU) being the households, as in case of 55th round data. The period covered for this survey is July 2004 – June 2005. Each district is formed of two basic strata, rural stratum and urban stratum. Each stratum is sub-divided into a number of sub-strata: With ‘r’ being the sample size allocated for a rural stratum, the number of sub-strata formed will be ‘r/2’. The villages within a district as per frame are first arranged in ascending order of population. Then sub-strata 1 to ‘r/2’ are demarcated in such a way that each sub-stratum comprises a group of villages of the arranged frame and has more or less equal population. Two FSUs are selected from each sub-stratum of a district of rural sector with Probability Proportional to Size with Replacement (PPSWR), size being the population as per Population Census 2001. Households are again classified into three second stage strata (SSS) as follows.

SSS 1: Relatively affluent households

SSS2 2: Households not belonging to SSS1 having principal earning from non-agricultural activity

SSS 3: Other households

From each SSS the sample households are selected by SRSWOR and information on all members of a household are collected through interviewing a representative from each household in the sample.

REFERENCES

1. Aghai, H., Zarafshani, K., & Behjat, A. M. (2008). The Effect of Crop Insurance on Technical Efficiency of Wheat Farmers in Kermanshah Province: A Corrected Ordinary Least Square Approach. *Journal of Applied Sciences , 8* (5), 891-894.
2. Aigner, D. J., Lovell, C. A., & Schmidt, P. (1977). Formulation And Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics , 6*, 21-37.
3. Albacea, Z. V. (2009). Small Area Estimation of Poverty Statistics. *PIDS Discussion Paper* (16).
4. Almas, I. (2008). International Income Inequality: Measuring PPP Bias By Estimating Engel Curves For Food. *30th General Conference of The International Association for Research in Income and Wealth, August 24-30.* Portoroz, Slovenia.
5. Anselin, A. (2000). Spatial Econometrics. In B. Baltagi (Ed.), *A Companion to Theoretical Econometrics* (pp. 310-330). Oxford: Basil Blackwell.
6. Anselin, L. (1980). Estimation Methods For Spatial Autoregressive Structures. *Regional Science Dissertation And Monograph Series.* Ithaca, NY.
7. Anselin, L. (1992). *Space Stat Tutorial: A Workbook For Using Space Stat in The Analysis of Spatial Data.* Urbana: University of Illinois.
8. Anselin, L. (1999). *Spatial Econometrics.* Dallas: University of Texas.
9. Anselin, L. (2000). Spatial Econometrics. In B. Baltagi (Ed.). Oxford: Basil Blackwell.
10. Anselin, L. (1988). *Spatial Econometrics: Methods And Models.* Dordrecht: Kluwer Academic Publishers.
11. Anselin, L. (2003). Spatial Externalities, Spatial Multipliers and Spatial Econometrics. *International Regional Science Review , 26* (2), 152-166.
12. Anselin, L., & Bera, A. K. (1998). Spatial Dependence in Linear Regression Models With An Introduction to Spatial Econometrics. In A. Ullah, & D. E. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237-289). New York: Marcel Dekker.
13. Anselin, L., & Griffith, D. (1988). Do Spatial Effects Really Matter in Regression Analysis? *Papers of the Regional Science Association , 65*, 11-34.
14. Anselin, L., & Hudak, S. (1992). Spatial Econometrics in Practice: A Review of Software Options. *Regional Science and Urban Economics , 22*, 509-536.
15. Aten, B., & Menezes, T. (2002). Poverty Price Levels: An Application to Brazilian Metropolitan Areas". *World Bank ICP Conference, March 11-15* . Washington, D C.

16. Ayadi, M., & Amara, M. (2008, (23-25) November). Spatial Patterns and Geographic Determinants of Welfare and Poverty in Tunisia. *Equity and Economic Development, ERF 15th Annual Conference* .

17. Baker, J. L., & Grosh, M. E. (1994). Measuring the Effects of Geographic Targeting on Poverty Reduction. *World Bank Living Standards Measurement Study Working Paper , 99*.

18. Balk, B. M. (1996). A Comparison of Ten Methods For Multilateral International Price and Volume Comparison. *Journal of Official Statistics , 12*, 199-222.

19. Ballini, F., Betti, G., Carrette, S., & Neri, L. (2009). Poverty and Inequality Mapping in The Commonwealth of Dominica. *Estudios Economicos , 2*, 123-162.

20. Baltagi, B. H., & Li, D. (2001). LM Tests for Functional Form and Spatial Error Correlation. *International Regional Science Review , 24*, 194-225.

21. Baltagi, B. H., & Li, D. (2005). Testing for Linear and Log-Linear Models Against Box-Cox Alternatives with Spatial Lag Dependence. In J. P. Lesage, & R. K. Pace, *Advances in Econometrics, Volume 18:Spatial and Spatiotemporal Econometrics.* Oxford: Elsivier.

22. Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel Curves and Consumer Demand. *The Review of Economics and Statistics , 79* (4), 527-539.

23. Bargain, O., Bhaumik, S. K., Chakrabarty, M., & Zhao, Z. (2009). Earnings Differences Between Chinese And Indian Wage Earners, 1987-2004. *Review of Income and Wealth , 55* (s1), 562-587.

24. Barma, S. (2007). *Socio-Political Movements in North Bengal (A Sub-Himalayan Tract).* Global Vision Publishing House.

25. Barrett, G. F., & Brzozowski, M. (2010). Using Engel Curves to Estimate The Bias in The Australian CPI. *Economic Record , 86* (272), 1-14.

26. Battese, G. E., Harter, R. M., & Fuller, W. A. (1988). An Error-Component Model For Prediction of County Crop Areas Using Survey and Satellite Data. *Journal of The American Statistical Association , 83*, 28-36.

27. BBS. (2004). *Local Estimation of Poverty and Malnutrition in Bangladesh.* United Nations World Food Programme, Bangladesh Bureau of Statistics, Dhaka.

28. Beatty, T., & Larsen, E. R. (2005). Using Engel Curves to Estimate Bias in The Canadian CPI As A Cost of Living Index. *Canadian Journal of Economics , 38* (2), 482-499.

29. Beck, N., & Gleditsch, K. S. (2006). Space Is More Than Geography: Using Spatial Econometrics in The Study of Political Economy. *50*, 27-44.

30. Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis With Special Reference to Education.* National Bureau of Economic Research, New York.

31. Becker, G. S. (1964). *Human Capital: A Theoritical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

32. Betti, G., & Ballini, F. (2008). Variance Estimates of Poverty and Inequality Measures in Albania. *Eastern European Economics , 46* (6), 87-101.

33. Betti, G., & Neri, L. (2010). *Further Updating Poverty Mapping in Albania.* World Bank.

34. Bhattacharya, N., Chatterjee, G. S., & Pal, P. (1988). Variations in Level of Living Across Regions and Social Groups in India, 1963-64 and 1973-74. In T. N. Srinivasan, & P. K. Bardhan (Eds.), *Rural Poverty in South Asia.* Oxford University Press.

35. Bhattacharyya, S. S., Joshi, P. D., & Roychowdhury, A. B. (1980). Regional Price Indices Based on NSS 25th Round Consumer Expenditure Data. *Sarvekshana , 3* (4), 107-121.

36. Bhaumik, S. K., & Chakrabarty, M. (2009). Is Education the Panacea for Economic Deprivation of Muslims?: Evidence from Wage Earners in India, 1987-2005. *Journal of Asian Economics , 20* (2), 137-149.

37. 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).

38. 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.

39. Bishop, J. A., Chow, K. V., & Zheng, B. (1995). Statistical Inference and Decomposable Poverty Measures. *Bulletin of Economic Research , 47*, 329-340.

40. Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources , 8* (4), 436-455.

41. Bokosi, F. K. (2007). Household Poverty Dynamics in Malawi: A Bivariate Probit Analysis". *Journal of Applied Science , 7* (2), 258-262.

42. Burger, R., Van Der Berg, S., Van Der Walt, S., & Yu, D. (2004). Geography As Destiny: Considering The Spatial Dimensions of Poverty and Deprivation in South

Africa. *Cornell University Conference on African Development and Poverty Reduction, 13 to 15 October* .

43. Cameron, A. C., & Trivedi, P. K. (2005). *Micoeconometrics: Methods and Applications.* New York: Cambridge University Press.

44. Case, A. C. (1991). Spatial Patterns in Household Demand. *Econometrica , 59* (4), 953-965.

45. Chambers, R. L., & Tzavidis, N. (2006). M-Quantile Models For Small Area Estimation. *Biometrika , 93* (2), 255-268.

46. Chambers, R., Tzavidis, N., Salvati, N., & Pratesi, M. (2008). M-Quantile Small Area Estimation, Poverty Mapping and Geographically Weighted Regression. *Presentation at "Creating Socio-Economic Data For Small Areas: Methods and Outcomes", September 19.* Canberra.

47. Chandrasiri, G. W., & Samarakoon, L. (2008). Spatial Patterns and Geographic Determinants of Poverty in Sri Lanka: Linking Poverty Mapping With Geoinformatics. *Asian Conference on Remote Sensing.* Colombo.

48. Chattopadhyay, S. (2010). District level Poverty Estimation : A Spatial Approach. *Economics Bulletin , 30* (4), 2962-2977.

49. Chong, K.-C., Lizarondo, M. S., Cruz, Z. S., Guerrero, C. V., & Smith, I. R. (1984). Milkfish Production Dualism in the Philippines:A Multidisciplinary Perspective On Continuous Low Yields and Constraints to Aquaculture Development. *ICLARM Technical Reports 15* .

50. Cliff, A. D., & Ord, J. K. (1981). *Spatial Processes: Models and Applications.* London: Pion.

51. Cochran, W. G. (1953). *Sampling Techniques.* Canada: John Wiley & Sons, Inc.

52. Coondoo, D., & Saha, S. (1990). Between-State Differentials in Rural Consumer Prices in India: An Analysis of Intertemporal Variations. *Sankhya, Series B , 52* (3), 347-360.

53. Coondoo, D., Majumder, A., & Chattopadhyay, S. (DOI: 10.1111/j.1475-4991.2010.00382.x, Published Online : March 25, 2010). Estimating Spatial Consumer Price Indices Through Engel Curve Analysis. *Review of Income and Wealth* .

54. Coondoo, D., Majumder, A., & Ray, R. (2004). A Method of Calculating Regional Consumer Price Differentials With Illustrative Evidence From India. *Review of Income and Wealth , 50* (1), 51-68.

55. Costa, A., Satorra, A., & Ventura, E. (2009). On The Performance of Small-Area Estimators: Fixed Vs. Random Area Parameters. *SORT , 33* (1), 85-104.

56. Costa, D. L. (2001). Estimating Real Income in The United States From 1888 to 1994: Correcting CPI Bias Using Engel Curves. *Journal of Political Economy , 109* (6), 1288-1310.

57. Costa, L., Satorra, A., & Ventura, E. (2003). An Empirical Evaluation of Small Area Estimators. *SORT , 27* (1), 113-136.

58. Coudouel, A., Hentschel, J. S., & Wodon, Q. T. (2002). Using Linear Regressions For Analyzing The Determinants of Poverty. In J. Klugman (Ed.), *A Sourcebook For Poverty Reduction Strategies (2-Volume set)* (Vol. 2, p. Annex A.8). World Bank.

59. Cushing, B., & Zheng, B. (1999). Locational Difference in the Underlying Characteristics of Poverty. *Paper presented at the Southern Regional Science Association Conference .*

60. Dabalen, A., & Ferrè, C. (2008). *Updating Poverty Maps: A Case Study of Albania.* Report to the World Bank.

61. Darmofal, D. (2006). Spatial Econometrics and Political Science. *Presented at the Annual Meeting of the Southern Political Association, Atlanta, GA.*

62. Davidson, R., & Flachaire, E. (2007). Asymptotic and Bootstrap Inference For Inequality and Poverty Measures. *Journal of Econometrics , 141* (1), 141-166.

63. Davis, B. (2003). *Choosing A Method of Poverty Mapping.* Food and Agriculture Organization of The United Nations.

64. Deaton, A. S. (2003). Prices and Poverty in India, 1987-2000. *Economic and Political Weekly , 37* (4), 362-368.

65. Deaton, A. S. (1997). *The Analysis of Household Surveys – A Microeconometric Approach to Development Policy.* Baltimore and London: Johns Hopkins University Press.

66. Deaton, A. S., & Muellbauer, J. (1980b). An Almost Ideal Demand System. *American Economic Review , 70*, 312-326.

67. Deaton, A. S., & Muellbauer, J. (1980a). *Economics and Consumer Behaviour.* Cambridge: Cambridge University Press.

68. Deaton, A. S., & Tarozzi, A. (2005). Prices and Poverty in India"; Chapter 16 of 'Data and Dogma: The Great Indian Poverty Debate'. In A. S. Deaton, & V. Kozel (Eds.), *Data and Dogma: The Great Indian Poverty Debate.* New Delhi, India: Macmillan.

69. *Delta Method*. (n.d.). Retrieved from Wikipedia.

70. Demombynes, G., Elbers, C., Lanjouw, J. O., Lanjouw, P., Mistiaen, J., & OZler, B. (2002). Producing an Improved Geographic Profile of Poverty: Methodology and Evidence From Three Developing Countries. *WIDER Discussion Paper* (39).

71. Diewert, W. E. (1999). Axiomatic and Economic Approaches to International Comparisons. In A. Heston, & R. E. Lipsey (Eds.), *International and Interarea Comparisons of Income, Output and Prices* (pp. 13-87). Chicago: University of Chicago Press.

72. Durbin, J. (1959). A Note on The Application of Quenouille's Method of Bias Reduction to The Estimation of Ratios. *Biometrika , 46*, 477-480.

73. Efron, B., & Stein, C. (1981). The Jackknife Estimate of Variance. *Annals of Statistics , 9*, 586-596.

74. Efron, B., & Tibshirani, R. (1986). Bootstrap Methods For Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science , 1*, 54-77.

75. Efron, B., & Tibshirani, R. J. (1993). *An Introduction to The Bootstrap.* Chapman & Hall.

76. Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro-Level Estimation of Poverty and Inequality. *Econometrica , 71* (1), 355-364.

77. 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.

78. Färe, R., Grosskopf, S., & Lovell, C. A. (1994). *Production Frontiers.* Cambridge: Cambridge University Press.

79. Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of Royal Statistical Society , 120* (3), 253-281.

80. Fasasi, A. R. (2007). Technical Efficiency in Food Crop Production in Oyo State, Nigeria. *J. Hum. Ecol. , 22* (3), 245-249.

81. Fay, R. E., & Herriot, R. A. (1979). Estimates of Income For Small Places: An Application of James-Stein Procedures to Census Data. *Journal of The American Statistical Association , 74* (366), 269-277.

82. Foster, J., Greer, J., & Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica , 52* (3), 761-766.

83. Fry, V., & Pashardes, P. (1989). Constructing The True Cost of Living Index From The Engel Curves of The PIGLOG Model. *Journal of Applied Econometrics , 4*, 41-56.

84. 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.

85. Ganguly, R. (2005). Poverty, Malgovernance and Ethnopolitical Mobilization: Gorkha Nationalism and The Gorkhaland Agitation in India. *Nationalism and Ethnic Politics , 11* (4), 467-502.

86. Garfinkel, I., & Haveman, R. (1977). Earnings Capacity, Economic Status, and Poverty. *Journal of Human Resources , 12* (1), 49-70.

87. Geary, R. C. (1958). A Note on Comparison of Exchange Rate and Purchasing Power Parities between Countries. *Journal of The Royal Statistical Society , 121 (Part 1)*, 97-99.

88. Geary, R. C. (1954). The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician , 5*, 115-145.

89. Geda, A., Jong, N., Mwabu, G., & Kimenyi, M. (2001). Determinants of Poverty in Kenya: Household-Level Analysis. *KIPPRA Discussion Paper , 9* . Nairobi: Kenya Institute For Public Policy Research and Analysis.

90. Ghosh, M., & Rao, J. N. (1994). Small Area Estimation: An Appraisal. *Statistical Science , 9* (1), 55-93.

91. Gibson, J., & Olivia, S. (2007). Spatial Autocorrelation and Non-Farm Rural Enterprises in Indonesia. *51st Conference, Australian Agricultural and Resource Economics Society, February 13-16* .

92. Gonzalez, M. E., & Hoza, C. (1978). Small-Area Estimation With Application to Unemployment and Housing Estimates. *Journal of The American Statistical Association , 73* (361), 7-15.

93. Greene, W. H. (2003). *Econometric Analysis* (5th ed.). Delhi, India: Pearson Education (Singapore) Pte. Ltd., Indian Branch.

94. Greene, W. H. (1980). Maximum Likelihood Estimation of Econometric Frontier Functions. *J Econ. , 13*, 27-56.

95. Greene, W. (1993). The Econometric Approach to Efficiency Measurement. In H. Fried, C. A. Lovell, & S. Schmidt (Eds.), *The Measurement of Productive Efficiency: Theory and Applications.* New York: Oxford University Press.

96. Hamilton, B. (2001). Using Engel's Law to Estimate CPI Bias. *American Economic Review , 91* (3), 619-630.

97. Heinrich, G. (1988). Changing Times, Testing Times: A Bootstrap Analysis of Poverty and Inequality Using The PACO Database. *CERT Discussion Paper* (9802).

98. Hentschel, J., Lanjouw, J. O., Lanjouw, P., & Poggi, J. (2000). Combining Census and Survey Data to Trace The Spatial Dimensions of Poverty: A Case Study of Ecuador. *World Bank Economic Review , 14* (1), 147-165.

99. Hill, R. J. (1997). A Taxonomy of Multilateral Methods For Making International Comparisons of Prices and Quantities. *Review of Income and Wealth , 43* (1), 49-69.

100. Holt, D., Smith, T. M., & Tomberlin, T. J. (1979). Small-Area Estimation With Application to Unemployment and Housing Estimates. *Journal of The American Statistical Association , 74* (366), 405-410.

101. Howes, S., & Lanjouw, J. O. (1998). Does Sample Design Matter For Poverty Rate Comparisons? *Review of Income and Wealth , 44* (1), 99-109.

102. *http://Ec.Europa.Eu/Employment_Social/Social_Inclusion/Docs/Regionalindicators_En.Pdf.* (n.d.).

103. Jensen, U., Gartner, H., & Rassler, S. (2006). Measuring Overeducation with Earnings Frontiers and Multiply Imputed Censored Income Data. *IAB Discussion Paper , 11*.

104. Jiang, J., Lahiri, P., Wan, S., & Wu, C. (2001). *Jackknifing in the Fay-Herriot Model with an Example.* Technical Report, University of Nebraska-Lincoln, Mathematics and Statistics, Lincoln.

105. Johnston, J., & Dinardo, J. (1997). *Econometric Methods.* The Mcgraw-Hill Companies, Inc.

106. Kelejian, H. H., & Prucha, I. R. (1998). A Generalized Spatial Two-Stage Least Squares Procedure For Estimating A Spatial Autoregressive Model With Autoregressive Disturbances. *Journal of Real Estate Finance and Economics , 17*, 99-121.

107. Khamis, S. (1972). Properties and Conditions for The Existence of A New Type of Index Numbers. *Sankhya, Series B , 32 (Parts 1 & 2)*, 81-98.

108. Kim, C. W., Phipps, T. T., & Anselin, L. (2003). Measuring the Benefits Air Quality Improvement: A Spatial Hedonic Approach. *Journal of Enviornmental Economics and Management , 45* (1), 24-39.

109. Kokoski, M. F., Moulton, B. R., & Zeischang, K. D. (1999). Interarea Price Comparisons For Heterogeneous Goods and Several Levels of Commodity

Aggregation. In R. E. Lipsey, & A. Heston (Eds.), *International and Interarea Comparisons of Prices, Income and Output* (pp. 327-364). National Bureau of Economic Research, Chicago University Press.

110. Kopp, R. J., & Smith, V. K. (1980). Frontier Production Function Estimates For Steam Electric Generation: A Comparative Analysis. *Southern Economic Journal , 46* (4), 163-173.

111. Kravis, I. B., Heston, A., & Summers, R. (1978). *International Comparison of Real Product and Purchasing Power.* Baltimore: John Hopkins University Press.

112. Lancaster, J., & Ray, R. (1998). Comparison of Alternative Models of Household Equivalence Scales: The Australian Evidence on Unit Record Data. *Economic Record , 74* (1-14).

113. Larse, M. D. (2003). Estimation of Small-Area Proportions Using Covariates and Survey Data. *Journal of Statistical Planning and Inference , 112* (1-2), 89-98.

114. Larsen, E. R. (2007). Does The CPI Mirror Costs-of-Living? Engel's Law Suggests Not in Norway. *Scandinavian Journal of Economics , 109* (1), 177-195.

115. Lemieux, T. (2002). Decomposing Changes in Wage Distributions: A Unified Approach. *Canadian Journal of Economics , 35* (4), 646-688.

116. Lesage, J. P., & Charles, J. S. (2008). Using Home Buyers' Revealed Preferences to Define The Urban–Rural Fringe. *J Geograph Syst , 10*, 1-21.

117. Levy, P. S., & Lameshow, S. (1991). *Sample of Populations: Methods and Applications.* Canada: John Wiley & Sons, Inc.

118. Lin, E. S. (2007). On the Standard Errors of Oaxaca-type Decompositions for Inter-industry Gender Wage Differentials. *Economics Bulletin , 10* (6), 1-11.

119. Logan, T. D. (2008). Are Engel Curve Estimates of CPI Bias Biased? *NBER Working Paper Series , 13870.*

120. Lovell, C. A. (1993). Production Frontier and Productive Efficiency. In H. O. Fried, C. A. Lovell, & S. S. Schmidt (Eds.). Oxford: Oxford University Press.

121. Manski, C. Economic Analysis of Social Interactions. *The Journal of Economic Perspectives , 14* (3), 115-136.

122. Manski, C. (1993). Identification of Endogenous Social Effects: the Reflection Problem. *The Review of Economic Studies , 60* (3), 531-542.

123. McEwin, M., & Elazar, D. (2006). *Statistics Division: UNESCAP.* Retrieved from UNESCAP Web site: http://www.unescap.org/stat/apex

124. Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy , 66* (4), 281-302.

125. Mincer, J. (1958). Investment in Human Capital and the Personal Income Distribution. *Journal of Political Economy , 66* (4), 281-302.

126. Mishra, A., & Ray, R. (Forthcoming in 2010). Prices, Inequality and Poverty: Methodology and Indian Evidence. *Review of Income and Wealth* .

127. Molina, I., & Rao, J. N. (2010). Small Area Estimation of Poverty Indicators. *(forthcoming in) Canadian Journal of Statistics* .

128. Moran, P. (1948). The Interpretation of Statistical Maps. *Journal of The Royal Statistical Society-Series B , 10*, 243-251.

129. Murthy, M. N. (1967). *Sampling Theory and Methods.* Calcutta: Statistical Publishing Society.

130. Neary, J. P. (2004). Rationalizing The Penn World Table: True Multilateral Indices For International Comparisons of Real Income. *American Economic Review , 94* (5), 1411-1428.

131. Neri, L., Ballini, F., & Betti, G. (2005). Poverty and Inequality Mapping in Transition Countries. *Statistics in Transition , 7* (1), 135-157.

132. Nielsen, H. S. (1998). Discrimination and Detailed Decomposition in A Logit Model. *Economics Letters , 61*, 115-120.

133. Norman, D. W. (1977). Economic Rationality of Traditional Hausa Dryland Farmers in the North of Nigeria. In R. D. Stevens, *Tradition and Dynamics in Small-farm Agriculture:Economic Studies in Asia , Africa and Latin America* (pp. 63-91). Ames: Iowa State University Press.

134. Oaxaca, R. (1973). *International Economic Review , 14*, 693-709.

135. Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review , 14*, 693-709.

136. Oaxaca, R., & Ransom, M. R. (1999). Identification in Detailed Wage Decomposition. *Review of Economics and Statistics , 81* (1), 154-157.

137. Oehlert, G. W. (1992). A Note on The Delta Method. *The American Statistician , 46* (1), 27-29.

138. Opsomer, J. D., Claeskens, G., Ranalli, M. G., Kauermann, G., & Breidt, F. J. (2008). Nonparametric Small Area Estimation Using Penalized Spline Regression. *Journal of The Royal Statistical Society-Series B , 70* (1), 265-286.

139. Ord, K. (1975). Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association , 70*, 120-126.

140. Osberg, L., & Xu, K. (2000). International Comparison of Poverty Intensity: Index Decomposition and Bootstrap Inference. *Journal of Human Resources , 35*, 51-81.

141. Papalia, R. B. (2006). Estimating Real Income at Regional Level Using A CPI Bias Correction. *Proceedings of Q2006, European Conference on Quality in Survey Statistics.*

142. Patton, M., & Mcerlean, S. (2005). Spatial Effects Within The Agricultural Land Market in Northen Ireland. *Journal of Agricultural Economics , 54* (1), 35-54.

143. Petrucci, A., Salvati, N., & Seghieri, C. (2004). Autologistic Regression Model for Poverty Mapping and Analysis. *Metodoloˇski zvezki , 1* (1), 225-234.

144. Pinkse, J., & Slade, M. E. (2009). The Future of Spatial Econometrics. *Prepared for the 5oth anniversary symposium of the Journal of Regional Science held at the New York Federal Reserve.*

145. Pisati, M. (2001). Tools For Spatial Data Analysis. *Stata Technical Bulletin Reprints , 10*, 277-281.

146. Powell, L. A. (2007). Approximating Variance of Demographic Parameters Using The Delta Method: A Reference For Avian Biologists. *The Condor , 109* (4), 949-954.

147. Powell, L. A. (2007). Approximating Variance of Demographic Parameters Using The Delta Method: A Reference For Avian Biologists. *The Condor , 109*, 949-954.

148. Pratesi, M., & Salvati, N. (2008). Small Area Estimation: The EBLUP Estimator Based on Spatially Correlated Random Area Effects. *Statistical Methods and Applications , 17* (1), 113-141.

149. Pratesi, M., Ranalli, M. G., & Salvati, N. (2008). Semiparametric M-quantile Regression for Estimating the Proportion of Acidic Lakes in 8-digit HUCs of the Northeastern US. *Environmetrics , 19*, 687-701.

150. Purcell, N. J., & Kish, L. (1979). Estimation for Small Domains. *Biometrics , 35*, 365-385.

151. Quintano, C., Castellano, R., & Punzo, G. (2007). Estimating Poverty in The Italian Provinces Using Small Area Estimation Models. *Metodološki Zvezki , 4* (1), 37-70.

152. Rao, D. S. (1997). Aggregation Methods For International Comparison of Purchasing Power Parities and Real Income: Analytical Issues and Some Recent Developments. *Proceedings of The International Statistical Institute, 51st Session*, (pp. 197-2000).

153. Rao, D. S. (2001). Weighted EKS and Generalised CPD Methods For Aggregation at Basic Heading Level and Above Basic Heading Level. *Joint World Bank-OECD Seminar on Purchasing Power Parities – Recent Advances in Methods and Applications* .

154. Rao, J. N. (2003). *Small Area Estimation.* London: Wiley.

155. Rao, J. N. (2003). Some New Developments in Small Area Estimation. *JIRSS , 2* (2), 145-169.

156. Rask, N. (1977). Factors limiting change on traditional small farmers in South Brazil. In R. D. Stevens, *Tradition and Dynamics in Small-farm Agriculture:Economic Studies in Asia , Africa and Latin America* (pp. 92-114). Ames: Iowa State university Press.

157. Rongve, I. (1995). Statistical Inference For Poverty Indices With Fixed Poverty Lines. *Economics Discussion Paper, University of Regina, Canada* (54).

158. Sastry, N. S. (2003). District Level Poverty Estimates: Feasibility of Using NSS Household Consumer Expenditure Survey Data. *Economic and Political Weekly , 38* (4), 409-412.

159. Schultz, T. P. (1992). The Role of Education and Human Capital in Economic Development: An Empirical Assessment. *paper presented at the conference on Economic Growth in the World Economy, Institut fur Weltwirtschaft, Kiel* .

160. Schultz, T. W. (1965). *Transforming Traditional Agriculture.* New Haven: Yale University Press.

161. Seber, G. A. *The Estimation of Animal Abundance and Related Parameters.* (2nd, Ed.) London: Griffin C and Co.

162. Seber, G. A. (1982). *The Estimation of Animal Abundance and Related Parameters.* London, New York: Chapman, Macmillan.

163. Sen, A. (1976). Real National Income. *The Review of Economic Studies , 43* (1), 19-39.

164. Sinha, S. K., & Rao, J. N. (2008). Robust Small Area Estimation Under Unit Level Models. *Section on Survey Research Methods - JSM* .

165. Small, K. A., & Steimetz, S. S. (2009). Spatial Hedonics and The Willingness to Pay For Residential Amenities.

166. Smith, A. (1759). The Theory of Moral Sentiments.

167. Smith, A. (1776). *The Wealth of Nations.* London: Penguin.

168. Sokal, R. R., Oden, N. L., & Thomson, B. A. (1998). Local spatial autocorrelation in a biological model. *Geographical Analysis , 30*, 331-354.

169. Som, R. K. (1965). Use of Interpenetrating Samples in Demographic Studies. *Sankhya, Series B , 27* (3/4), 329-342.

170. Srivastava, A. K. (2007). Small Area Estimation – A Perspective and Some Applications. *Journal of The Indian Society of Agricultural Statistics , 61* (3), 295-309.

171. Summers, R. (1973). International Price Comparisons Based Upon Incomplete Data. *Review of Income and Wealth , 19* (1), 1-16.

172. Tarozzi, A., & Deaton, A. (2009). Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas. *Review of Economics and Statistics , 91* (4), 773-792.

173. Tobler, W. (1970). A Computer Movie Simulating Urban Growth in The Detroit Region. *Economic Geography , 46* (2), 234-240.

174. Tong, Y. K. (2009). A Study of E-Recruitment Technology Adoption in Malaysia. *Industrial Management & Data Systems , 109* (2), 281-300.

175. Tyler, W. G. (1979). Technical Efficiency in Production in A Developing Country: An Empirical Examination of The Brazilian Plastics and Steel Industries. *Oxford Economic Papers , 31* (3), 477-495.

176. 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.

177. Voss, P. R., Long, D. D., Hammer, R. B., & Friedman, S. (2006). County child poverty rates in the US: a spatial regression approach. *25* (4).

178. White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and A Direct Test For Heteroscedasticity. *Econometrica , 48*, 817-838.

179. Won, K. C., Phipps, T. T., & Anselin, L. (2003). Measuring The Benefits of Air Quality Improvement: A Spatial Hedonic Approach. *Journal of Environmental Economics and Management , 45*, 24-39.

180. World Bank. (2003). *Poverty Reduction Strategy Sourcebook.*

181. Xu, J., & Long, J. S. (2005). Confidence Intervals for Predicted Outcomes in Regression Models for Categorical Outcomes. *The Stata Journal , 5* (4), 537-559.

182. Yun, M. S. (2004). Decomposing Differences in The First Moment. *Economic Letters , 82* (2), 275-280.

183. Yun, M. S. (2005). Hypothesis Tests When Decomposing Differences in The First Moment. *Journal of Economic and Social Measurement , 30*, 295-304.