Economics Bulletin

Volume 31, Issue 3

Earnings efficiency and poverty dominance analysis: a spatial approach

Somnath Chattopadhyay
Indian Statistical Institute

Abstract

The paper estimates an earnings frontier by the method of Corrected Ordinary Least Squares (COLS) and categorizes households as efficient or inefficient based on some benchmark efficiency score and the estimated frontier. The spatial distribution of the poor and non poor households is then explored by construction of a poverty segregation curve across efficiency zones. Robust poverty comparisons across the efficient and inefficient groups reveal that poverty is in fact higher for the efficient group compared to the inefficient one. The paper thus indirectly supports the "poor but efficient hypothesis".

The paper being an extension work of a chapter of my PhD thesis, naturally there are coincidences in part for which I once again express my heart-felt gratitude to Prof Amita Majumder of Indian Statistical Institute, Kolkata for her overall guidance. I am also grateful to an anonymous referee for the valuable comments. The usual disclaimer applies.
Citation: Somnath Chattopadhyay, (2011) "Earnings efficiency and poverty dominance analysis: a spatial approach", *Economics Bulletin*, Vol. 31 no.3 pp. 2298-2318.
Submitted: May 01 2011. **Published:** August 13, 2011.

1. Introduction

The concept of *technical efficiency* is an issue of fundamental importance in production function analysis. In the classical production function analysis (Afrait, 1972; Truett & Roberts, 1973) each firm, with the objective of maximizing output (subject to the availability of inputs), operates on its *production frontier*, showing the maximum possible output, given input levels. Empirical studies (Tyler, 1979; Kopp & Smith, 1980; Fasasi, 2007; Tong, 2009) however suggest that given the level of technology, production units differ in terms of utilization of inputs. The difference between the potential output (*production frontier*) and the actual output for each firm may thus be ascribed 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 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.* The concept of technical efficiency considered in this paper is 'output oriented' *technical efficiency,* which shows the ability of the firm to maximize output from a given amount of inputs[1].

The above concept can be used in the context of the labour market to define an *earnings frontier* (potential earnings). The concept of *potential earnings* is based on human capital theory, which is built upon the neoclassical theories of investment and production (Smith, 1759; Mincer, 1958; Becker G. S., 1964; Schultz T. P., 1992). Investment in human capital in the form of schooling, work experience etc., translates into earned income. The *earnings frontier* (potential earnings) describes the highest potential income associated with a given stock of human capital, endowment and social opportunities. All individuals are located either on or below this frontier. Individuals translating their potential earnings into actual earnings enjoy a fully efficient position. On the other hand individuals earning less than their potentials suffer from some kind of *earnings inefficiency*. A closely related concept in connection with this issue of potential earnings is the problem of *poverty*, i.e., the state of a household's income being less than some subsistence level (poverty line). Some poor households having income less than the poverty line might have an earnings frontier that is above the poverty line. But for some poor households even the 'frontiers' may be below the poverty line. That is, even if they are fully efficient, given their stock of human capital and endowment, their potential incomes are below the poverty line. Thus these households cannot possibly be pulled out of poverty unless they are given some *extraneous assistance* (say in terms of 'social opportunities'). The present study attempts to look into the status of poverty and *earnings efficiency* for the rural households of West Bengal, an eastern State of India.

Estimating an earnings frontier using monthly consumption expenditure as a proxy for income, the paper classifies the households in terms of their efficiency scores measuring the difference of their potential earnings from the actual earnings. Earnings frontier is estimated in the literature (Landeau, 2003; Jensen, Gartner, & Rassler, 2006) using a parametric *stochastic*

[1] There is another concept of technical efficiency, i.e., the input-oriented technical efficiency signifying the firm's ability to minimize the inputs required to produce a given amount of output (Kumbhakar & Lovell, 2003; Fried, Lovell, & Schimdt).

frontier approach (SFA).[2] An alternative simpler methodology of "*Corrected Ordinary Least Square*" is adopted in this paper instead of SFA. Arranging the households in terms of their efficiency scores, the variation in the incidences of poverty is observed over the efficiency classes. The analysis is *spatial* in the sense that it considers only the *cross-sectional* variation among the observational units. To make the approach robust to the choice of poverty line, a *stochastic dominance*[3] technique to the comparison of poverty across efficiency zones is adopted. Stochastic dominance in relation to analysis of poverty examines an *unambiguous comparison* between two income/expenditure distributions over a range of poverty lines (Madden, 2000; Deaton, 1997). The spatial pattern in the incidences of poverty is again explored by the construction of a *poverty segregation curve*[4] (Dhongde, 2011) across the efficiency zones.

The distinctive feature of the present paper is the combination of the household level efficiency scores with *stochastic dominance* and construction of a *poverty segregation curve* to look into the regional heterogeneity in the distribution of the poor. An important finding which comes out is that the efficient group has almost unambiguously more incidence of poverty than the inefficient group. This implies that poverty; particularly rural poverty in India can be reduced by an enhancement of only the resource level (physical endowment and human capital). This finding indirectly supports the "poor but efficient" hypothesis popularized by (Schultz T. W., 1965), which says that poor people are in fact more efficient because inflicted with sufferings they just do the best they can under the difficult circumstances and their fields are as productive as they can be; they just cannot be very productive (Tax, 1953; Duflo, 2003).

The plan of the paper is as follows: Section 2 describes the methodology; Section 3 describes data and results and finally Section 4 concludes.

2. Methodology

The *earnings frontier,* defined as the potential earnings, given the stock of human capital and endowments, has been estimated in this paper as follows:

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

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 $(\ln \alpha_0 + \sum_{k=1}^{K} \alpha_k \ ln\, x_{ik})$ denotes the 'frontier'. Parameters of equation (1) have been estimated in this paper using the method of corrected ordinary least square (COLS) technique (Richmond, 1974; Greene, 1980).[5] Efficiency scores are computed in an analogous manner to estimating technical efficiency from the production frontiers.

[2] See (Aigner, Lovell, & Schmidt, 1977; Farrell, 1957; Färe, Grosskopf, & Lovell, 1994; Lovell, 1993) for discussion about SFA.

[3] Stochastic dominance is a form of stochastic ordering quantifying the concept of one random variable being bigger than another (Bawa, 1975; Hader & Russel, 1969).

[4] Segregation, a concept measuring the unevenness of distributions for groups categorized in some respects, was first used by (Duncan & Duncan, 1955). The recent applications of segregation curve have been made in (Hutchens, 2004; Dygalo, 2007).

[5] Application of COLS technique can be found in (Aghai, Zarafshani, & Behjat, 2008), (Ramos & Silber, 2005). For further information on COLS and other possible estimation methods, see Appendix A3 in (Deutsch, Ramos, & Silber, 2003). The advantage of the COLS method is that it is easy to implement and generates an unbiased production (earnings) frontier that lies above the data, though the simplicity comes at the cost that structure of the "best practice" technology is the same as the structure of "central tendency" production technology (Kumbhakar &

2.1. The Technique of COLS:

Application of Ordinary Least Square (OLS) on (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 $E(\xi_i) \neq 0$. Following (Greene, 2008) correction for the bias is made as follows:
Rewrite equation (1) as

$$\ln Y_i = (\ln \alpha_0 - E(\xi_i)) + \sum_{k=1}^{K} \alpha_k \; ln\, x_{ik} + (E(\xi_i) - \xi_i).$$
$$= \alpha_0^* + \sum_{k=1}^{K} \alpha_k \; ln\, x_{ik} + e_i, \text{ say.} \qquad (1.1)$$

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,

$$\ln \widehat{\alpha_0} = \widehat{\alpha_0}^* + \widehat{e^*}, \text{ where } \widehat{e^*} = \max\{\widehat{e_i}\} \qquad (1.2)$$

The corrected residuals (COLS residuals) are given by

$$\hat{\xi}_i = \widehat{e^*} - \widehat{e_i} \qquad (1.3)$$

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:

$$TE_i = exp\,(\widehat{-\xi_i}) \qquad (1.4)$$

2.2. Classification of the Households:

Households have been categorized as *efficient* or *inefficient* based on some benchmark level of efficiency scores. Households with an 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 not any a priori rule for fixation of the bench mark level in the literature. As the objective of this paper is to compare the status of poverty between the efficient (E) and inefficient (I) groups, the more the difference between mean technical efficiency scores between the groups, the more appropriate should be categorization of the groups. Higher percentile values of the state level efficiency scores will thus be preferable as the bench mark scores. But with fixation of higher percentile values as bench marks, the number of observations for group E gets decreased. Thus to satisfy these two criteria that appropriate characterization of the groups is possible and at the same time enough sample size is found for getting reliable poverty estimates, I have used the 95th percentile value of the state level efficiency scores (calculated using (1.4)) as the benchmark.

2.3. Measures of Poverty:

The head count index is the simplest measure of poverty measuring the proportion of poor persons (households) in the population. This index is criticized on ground that it fails to take into account the intensity of poverty among the poor. For a population (or sample) of N persons (households), the headcount index is: $P_0 = \frac{1}{N}\sum_{i=1}^{N} I(y_i \leq z)$; y_i denoting the income of the i^{th} person (household) and z being the official poverty line. $I(.)$ is an indicator function that takes on a value of 1 if the bracketed expression is true, and 0 otherwise.

The *poverty gap index* measures the extent to which individual person's (household's) income fall below the poverty line as a proportion of the poverty line. The sum of these

Lovell, 2003). Given the objective of the present study and the fact that only cross-sectional household level data are available, it is expected that results will be robust to alternative model specifications and estimation techniques.

shortfalls, i.e., *poverty gaps* gives the minimum cost of eliminating of poverty, if transfers are perfectly targeted. This measure does not reflect the changes in inequality among the poor. The poverty gap index is given as: $P_1 = \frac{1}{N}\sum_{i=1}^{N}\frac{G_i}{z}$; where $G_i = (z - y_i)\, I(y_i \le z)$ is the income shortfall of the i^{th} person (household).

To construct a measure of poverty that takes into account *inequality* among the poor, squared poverty gap (*poverty severity*) index is used. This is simply a weighted sum of poverty gaps (as a proportion of the poverty line), where the weights are the proportionate poverty gaps themselves. This implicitly gives more weight to persons (households) that are far below the poverty line. The measure lacking any intuitive appeal cannot be interpreted very easily and seldom used in the literature. Formally, squared poverty gap index, $P_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{G_i}{z}\right)^2$

All the above three poverty measures fall under a more general family of poverty measures, viz. Foster, Greer, Thorbecke (FGT) measure of poverty (Foster, Greer, & Thorbecke, 1984)[6]. In its continuous form, the measure is given by: $F_\alpha = \int_0^z \left(\frac{z-y}{z}\right)^\alpha dy$

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. The measure in its discrete form is written as: $F_\alpha = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z-y_i}{z}\right)^\alpha I(y_i \le z)$

Comparisons are made in the levels of living among households in terms of these poverty measures following a stochastic dominance approach as described below:

2.4. Poverty dominance analysis

Poverty dominance analysis is a tool that primarily deals with the sensitivity of ranking of income distributions in terms of poverty with respect to the specification of a poverty line and the same with respect to the specifications of a poverty measure (Quirk & Saposnik, 1962; Hader & Russel, 1969; Rothschild & Stiglitz, 1970).

For making an unambiguous comparison of the intensity of poverty between two groups, it is needed that status of poverty, measured in some respect is always lower/higher in one group than in the other for a relevant range of poverty lines. Dominance ordering with respect to the head count ratio is termed as First Order Dominance (FOD), which can be established by constructing a *Poverty Incidence curve* (PI). A PI is traced by plotting the head count ratio on the vertical axis and the poverty line on the horizontal axis allowing the latter to vary from zero to the maximum consumption. Each point on the curve thus giving the proportion of the population consuming less than the amount given on the horizontal axis, the poverty incidence curve is simply the *cumulative distribution function* (CDF) of consumption (income). Now if the PI of any distribution (A) is nowhere above the PI of another distribution (B), the former distribution has unambiguously lower incidence of poverty as measured by P_0 compared to the latter for the relevant range of poverty lines and A is said to FOD over B. In the context of Welfare Economics, let F_A and F_B denote the CDF s corresponding to two welfare indicators. Assuming support in the non-negative real numbers, let $D_A^1(x) = F_A(x) = \int_0^x dF_A(y)$. Then FOD implies $D_A^1(x) \le D_B^1(x)$ for all $x \in R_+$.

When the *poverty incidence* curves of two income distributions cross, the unambiguous ordering in terms of head count ratio is not possible. Dominance relationship can here be

[6] See (Coudouel, Hentschel, & Wodon, 2002), (Ravallion, 1992) for a discussion about the poverty measures.

established constructing a new curve from the poverty incidence curve, viz. the *Poverty Deficit* curve (PD). A PD is obtained by plotting the areas up to each point on the PI against the poverty line allowing the latter to vary from zero to the maximum consumption. Thus if $D^2(x)$ be the area under F up to x, i.e., $D^2(x) = \int_0^x D^1(y)\, dy$; the Distribution A is said to dominate over Distribution B if $D_A^2(x) \leq D_B^2(x) \forall$ x. This implies that PD of A is beneath that of B for any poverty line. This is called *Second Order Dominance* (SOD).

If the *poverty deficit* curves intersect, dominance relationship should be established by defining another curve, viz. the *Poverty Severity* curve, which calculates the area under the *poverty deficit* curve at each point. The *Third Order Dominance* (TOD) requires that the *poverty severity* curve is everywhere lower in one of the two distributions being compared. If necessary, one can go on to test higher order dominance, though the interpretation of the (increasingly) restricted class of measures becomes less clear. Now the existence of the higher order of dominance implies; $D_A^s(x) \leq D_B^s(x)$, where $D_A^s(x) = \int_0^x D_A^{s-1}(y)\, dy$; for any integer, s≥ 3. Following (Davidson & Duclos, 2000) $D^s(x)$ can alternatively be expressed as:

$$D^s(x) = \frac{1}{N(s-1)!}\int_0^x (x-y)^{s-1}\, dF(y)$$

Now, for a random sample of N independent observations on the welfare variable, y a natural estimator of $D^s(x)$ is: $\widehat{D}^s(x) = \frac{1}{N(s-1)!}\int_0^x (x-y)^{s-1}\, d\widehat{F}(y) = \frac{1}{N(s-1)!}\sum_{i=1}^{n}(x-y_i)^{s-1}\, I(y_i \leq x)$; $\widehat{F}$ being the empirical distribution function of the sample and $\mathrm{I}(.)$ being an indicator function, which is equal to one when the argument in it is true and equal to zero, when otherwise.

3. Data and results

As an illustrative application, the procedure presented above has been applied to Indian National Sample Survey Organization's (NSSO) 61st round of employment-unemployment data conducted during the years 2004-05. The analysis of poverty and efficiency has been done for the rural sector of West Bengal, an eastern state of India. Monthly consumption expenditure has been taken as a proxy for household income (earning).

According to the "canonical model", earnings are determined by human capital, which consists of capacities to contribute to production, generically called skills (Bowles, Gintis, & Osborne, 2001). Investment in human capital in the form of schooling, work experience etc., translates into earned income. The household general education level has been taken as a proxy for human capital. 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.

The other component determining the earnings is the "endowment" in the form of land possessed (Land) and labour. The household size (HHSIZE) has been taken to be proxy for labour. The variables have been made unidirectional in the sense that they have been transformed (wherever necessary) so that these may be interpreted as 'inputs' having positive effects on income.

Another variable treated as input in the present analysis is (1-Dependency Ratio), where dependency ratio measures proportion of dependents, i.e., the proportion of children (aged less than 15 years) and old persons (aged above 65 years) in the household. The more the dependency ratio, the less will be the earnings. Dependency Ratio having a negative impact on earnings, (1-Dependency Ratio) will have a positive impact and thus can be treated as an input (endowment) in the earnings frontier (Chattopadhyay, 2012).

The estimates of the parameters of the earnings frontier are reported in Table 1. All the parameters have the positive signs and are statistically significant at 5 % level of significance.

The objective of the study is to observe the status of poverty and efficiency for the rural households. The efficiency scores are accordingly computed (using the steps as given Section 2.1) and the households are classified on the basis of the estimated efficiency scores into different efficiency zones. The level of poverty is observed for each efficiency zone. Taking poverty line to be the official state level poverty line of Rs.382.82 per capita per month for rural West Bengal, it is noted that the higher the level of efficiency the higher the intensity of poverty (see Table 2). The spatial variation in the status of poverty is explored in view of the distributional heterogeneity among the poor and non poor in the efficiency framework. The efficiency zones are first arranged in terms of increasing order of poverty. The no. of poor and non poor are observed for each efficiency zone. If x_{tk} denotes the number of type "t" household in the efficiency class "k", the proportional share of any specific type of household is given by ${d_k}^t = \frac{x_{tk}}{\sum_{k=1}^{K} x_{tk}}$. Suppose t=1 denotes the state of being poor and t=2 denotes the state of being non poor. The cumulative share of the poor households up to any efficiency class, say, r will thus be: $M_r = \sum_{k=1}^{r} {d_k}^1$ and the corresponding share for the non poor households will be: $N_r = \sum_{k=1}^{r} {d_k}^2$. Now joining the points (N_r, M_r) ∀ r=1, 2..., K will give us a curve that resembles a segregation curve. In a sense it indicates how heterogeneously poor and non poor are distributed across the classes and by construction it lies between zero and one. If there is not any heterogeneity at all, i.e., ${d_k}^1 = {d_k}^2\ \forall\ k = 1(1)K$; segregation curve will be a straight line with slope being equal to one. With no unevenness in the distribution, this in fact is a situation of complete integration of the two types of households, which essentially means that there is no variation in the incidences of poverty across the zones. In contrast, if the variation in the incidences of poverty is such that some regions have hundred per cent poor population and other regions have zero percent poverty, the curve will be L-shaped lying along the axes. This is the case of complete segregation of poor and non-poor. The present study finds an upward sloping segregation curve that lies in between these two extremes (Figure 1). This indicates that the poor households are not homogenously distributed. The amount of heterogeneity quantified in terms of Duncan's dissimilarity index (Duncan & Duncan, 1955) turns out to be of magnitude 0.25. [7]

The essence of the above findings is that the poor are not homogenously distributed across efficiency classes and the intensity of poverty is more in higher efficiency zone. Fixing a bench mark efficiency score to classify households as efficient or inefficient (see Table 3), all the three measures of poverty, viz. head count ratio (FGT0), income gap ratio (FGT1), squared income gap ratio (FGT2) give the status of poverty to be more in the efficient group (E) than in the inefficient group (I)[8]. This implies that as far as the *spatial variation* in the incidences of poverty is concerned; it varies in a way that poor people are on an average more efficient. This finding is in line with the *" small farmers are poor but efficient"* theory (Chong, Lizarondo,

---

[7] The Duncan's dissimilarity index: $D = 0.5 \times \sum_{i=1}^{K} |d_k^1 - d_k^2|$, gives the maximum vertical distance between the segregation curve and the diagonal line.

[8] The specifications of the groups have been clarified in Section 2.2.

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) , (Rask, 1977). The enhancement of the resource level is thus the only way that poor people can be made better off.

The finding as in above is however contingent upon some fixed *official state level poverty line*. The result can be made robust by implementing a *poverty dominance* analysis as outlined in the previous section. Let $\widehat{D}^s{}_E(x)$ be the estimate of $D^s(x)$ for the efficient region, E and $\widehat{D}^s{}_I(x)$ be the corresponding estimate for the inefficient region, I. Now for an FGT(α) type poverty measure, the statistical precision of $\widehat{D}^s(x)$ can be obtained following (Davidson & Duclos, 1997) as: Variance$\left(\widehat{D}^s(x)\right) = \left(\frac{x}{\alpha!}\right)^2 \left[\frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{x-y_i}{x}\right)^{\alpha}\right]^2 - (P_\alpha)^2\right]$; where $P_\alpha = \text{FGT}(\alpha)$. The statistical precision of the difference of the estimate, $\widehat{D}^s(x)$ for Regions E and I is thus given by: Variance$\left(\widehat{D}_E{}^s(x) - \widehat{D}_I{}^s(x)\right) = \text{Variance}\left(\widehat{D}_E{}^s(x)\right) + \text{Variance}\left(\widehat{D}_I{}^s(x)\right)$

Now for each poverty line, *x* up to some arbitrarily defined highest poverty line, simple "t" statistics are computed for testing the null hypothesis, $H_0\colon \widehat{D}_E{}^s(x) - \widehat{D}_I{}^s(x) = 0$.

Taking the state-level mean efficiency score as the bench mark value, it is found that $H_0\colon \widehat{D}_E{}^s(x) - \widehat{D}_I{}^s(x) = 0$ is rejected and the signs are the same for all *t* values for *s*=1. This indicates I *first order dominates over* E. In other words, the incidence of poverty is more in Region, E than in Region, I. for all poverty lines up to Rs 525.875[9] and for all poverty measures. Table 4 shows that the difference $\left(\widehat{D}_E{}^1(x) - \widehat{D}_I{}^1(x)\right)$ is statistically significant and negative over the range of poverty lines from Rs 227 to Rs 526 per capita per month. Setting the bench mark to *75th percentile* of the state level efficiency score, it is observed that dominance is achieved at order 2 for region I over region E (Tables 5 and 6). In other words, though the poverty incidence curve of I is not everywhere lying below that of E, the poverty deficit curve of Region I is lying below that of Region E for the relevant range of poverty lines.

The implication of the above findings is that rural poor are closer to their frontier and their frontier is beneath the poverty line so that enhancement of the efficiency (earnings potential) will not improve their status. The result can be strengthened by applying a "treatment effect" model testing the effect of efficiency on the status of poverty. The *treatment effect* (Caliendo & Kopeinig, 2008; Becker & Ichino, 2002) signifies the effect of getting any treatment (T) on some outcome variable (Y). Suppose Y_{0i} and Y_{1i} are the values of Y under the two values of T for the i^{th} individual with T=0 signifying no treatment and T=1 signifying the treatment. The effect of the treatment can be measured by the *average treatment effect on the treated* (ATT), which is defined as: ATT=$E(Y_{1i} - Y_{0i}\ /W_i = 1)$, where $W_i = 1$ denotes the receipt of treatment by the i^{th} individual. The first term in bracket denotes the average outcome in the population of trainees (T=1), a potentially observable quantity. The second term in bracket

[9] This is the 50th percentile score of the monthly per capita expenditure values. In fact strictly speaking this is the restricted form of stochastic dominance, dominance relationship holding for a certain reasonable range of poverty lines.

signifies (counterfactual) average earnings of trainees had they not been trained (T=0). The effect of treatment can be established by noting whether ATT is statistically significant or not.

In the context of the present analysis, the basic idea is to assume that being efficient is like receiving a treatment so that we may estimate an ATT on the probability of being in poverty. In this way, we want to compare the probability of being in poverty for efficient households (T=1) with that of the inefficient households (T=0). Now, the effect of the treatment (efficiency) can be found out by eliminating the effects of other factors influencing the outcome (incidence of poverty). This can be done by comparing (matching) the outcomes of the treated and non-treated households, which are as similar as possible. The matching mechanism is based on scores ("propensity score") assigned to each subject (household), which are generated from their pre-treatment characteristics (Esquivel & Alejandra, 2007).

The outcome variable is the probability(p_i) of being in poverty for each household. $p_i{}'$s have been obtained by using the parameter estimates (γ) of a regression of the logarithm of ratio of income (y) to poverty line (z) on a set of socio-economic variables (V) as: $p_i = \Phi(V_i\gamma^*)$; Φ being the CDF of standard normal distribution; $\gamma^* = -\frac{\gamma}{\sigma}$, assuming $var(\varepsilon_i)= \sigma^2$, γ being the regression parameter vector and ε_i being the stochastic regression error term.[10] Households above some bench mark level of efficiency score are considered as the treatment group (T=1) and these are matched with control group households (T=0) that have similar characteristics but lower level of efficiency than the bench mark. The weighted average of the difference in $p_i's$ between the matched households gives the net effect of treatment. It is observed from Table 7 that there is not any significant difference in the status of poverty between the groups. i.e., ATT is statistically insignificant at 5% level of significance. [11]

4. Conclusion

The paper has explored the status of poverty and efficiency for the rural households of West Bengal, an eastern state of India. The spatial pattern in the distribution of the poor in the efficiency framework has been depicted by construction of a poverty segregation curve. The curve shows that rural poor are not homogenously distributed across different efficiency classes. Robust poverty comparisons reveal that incidence of poverty is more in the efficient region than in the inefficient region. The finding is pertinent particularly in the context of developing countries like India. It means that rural poor households are already efficient. The dearth of resource level (endowment) is thus the dominant factor causing the poverty in the rural sector. The average treatment effect of the level of efficiency on the probability of being in poverty has been found to be statistically insignificant. The paper thus indirectly supports the much debated "poor but efficient" hypothesis.

It would be an interesting exercise to reformulate the present analysis in a fuzzy logic framework (Zadeh, 1965). Instead of using a single bench mark, two critical values of bench marks can be used for categorization of the households. Each household will have a membership function in the domain of efficiency and poverty (Abdullah, 2010; Cerioli & Zani, 1990).

[10] This is the (World Bank) methodology. See (Coudouel, Hentschel, & Wodon, Using Linear Regressions For Analyzing The Determinants of Poverty, 2002) for a discussion about the World Bank method *using linear regression in analyzing the determinants of poverty* and (Chattopadhyay, 2012) for an application of the methodology in the context of rural West Bengal.

[11] The detailed derivations and tables are not given in this paper. These will be supplied to interested readers on request.

Studying the spatial variation in household richness (Brzezinski, 2010; Peichl, 2010; Medeiros, 2006) indices alongside variation in their earnings potential is another area that can be explored.

References

1. Abdullah, L. (2010) "Poverty Lines Based on Fuzzy Sets Theory and its Application to Malayasian Data". *Social Indicators Research* .
2. Afrait, S. N. (1972) "Efficiency Estimation of Production Function". *International Economic Review* **13 (3)**.
3. 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.
4. Aigner, D. J., Lovell, C. A., & Schmidt, P. (1977) "Formulation And Estimation of Stochastic Frontier Production Function Models". *Journal of Econometrics* **6,** 21-37.
5. 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.
6. Bawa, V. S. (1975) "Optimal Rules for Ordering Uncertain Prospects". *Journal of Financial Economics* **2,** 95-121.
7. Becker, G. S. (1964) "*Human Capital: A Theoretical and Empirical Analysis With Special Reference to Education".* National Bureau of Economic Research, New York.
8. Becker, S. O., & Ichino, A. (2002) "Estimation of Average Treatment Effects Based on Propensity Scores". *The Stata Journal* **2 (4)**, 358-377.
9. Bellù, L. G., & Liberati, P. (2005) "Poverty Analysis". EASYPol .
10. 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.
11. Bowles, S., Gintis, H., & Osborne, M. (2001) "The Determinants of Earnings: A Behavioural Approach". *Journal of Economic Literature* **39 (4)**, 1137-1176.
12. Brzezinski, M. (2010) "Income Affluence in Poland". *Social Indicators Research* **99 (2)**, 285-299.
13. Caliendo, M., & Kopeinig, S. (2008) "Some Practical Guidance for The Implementation Of Propensity Score Matching". *Journal of Economic Surveys* **22 (1)**, 31-72.

14. Cerioli, A., & Zani, S. (1990) "A Fuzzy Approach to the Measurement of Poverty". In C. Dagum, & M. Zenga, *Income and Wealth Distribution, Inequality and Poverty* (pp. 272-284). Berlin: Springer Verlag,.

15. Chattopadhyay, S. (2012) Inter-regional Poverty Comparisons in India: Case of West Bengal. *forthcoming in Journal of Quantitative Economics* .

16. 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 .

17. Coudouel, A., Hentschel, J. S., & Wodon, Q. T. (2002) "Poverty Measurement and Analysis".

18. 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.

19. Daneshvary, N., Herzog, H. W., Hofler, R. A., & Schlottmann, A. N. (1992) "Job Search and Immigrant Assimilation: An Earnings Frontier Approach". *Review Of Economics and Statistics* **74 (3)**.

20. Davidson, R., & Duclos, J. (1997) "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers". *Econometrica* **65**, 1453-1465.

21. Davidson, R., & Duclos, J. Y. (2000) "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality". *Econometrica* **68**, 1435-1464.

22. Deaton, A. S. (1997) *"The Analysis of Household Surveys – A Microeconometric Approach to Development Policy".* Baltimore and London: Johns Hopkins University Press.

23. Deutsch, J., Ramos, X., & Silber, J. (2003) "Poverty and Inequality of Standard of Living and Quality of Life in Great Britain". In J. Sirgy, D. Rahtz, & A. C. Samli, *Advances in Quality-of-Life Theory and Research,* (pp. 99-128). Dordrecht: Kluwer Academic Publishers.

24. Dhongde, S. (2011) "A Poverty Segregation Curve". *Spring Seminar Series, School of Economics* .

25. Duflo, E. (2003) "Poor but Rational". Department of Economics, MIT.

26. Duncan, O. D., & Duncan, B. (1955) "A Methodological Analysis of Segregation Indices". *American Sociological Review* **20 (2)**, 210-217.
27. Dygalo, N. (2007) "On Unemployment Duration and Narrowing Job Opportunities at Older Ages". *Economics Bulletin* **10 (21)**, 1-6.
28. Efron, B., & Tibshirani, R. (1986) "Bootstrap Methods For Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy". *Statistical Science* **1**, 54-77.
29. Esquivel, G., & Alejandra, H.-P. (2007) "Remittances and Poverty in Mexico: A Propensity Score Matching Approach". *Integration and Trade Journal* .
30. Färe, R., Grosskopf, S., & Lovell, C. A. (1994) "*Production Frontiers".* Cambridge: Cambridge University Press.
31. Farrell, M. J. (1957) "The Measurement of Productive Efficiency". *Journal of Royal Statistical Society* **120 (3)**, 253-281.
32. Fasasi, A. R. (2007) "Technical Efficiency in Food Crop Production in Oyo State, Nigeria". *J. Hum. Ecol.* **22 (3)**, 245-249.
33. Foster, J., Greer, J., & Thorbecke, E. (1984) "A Class of Decomposable Poverty Measures". *Econometrica* **52 (3)**, 761-766.
34. Fried, H. O., Lovell, C. A., & Schimdt, S. S. (2008) "*Efficiency and Productivity".* In Fried, H.O., Lovell, C.A.K .& Schmidt, S.S. (Eds.), *The Measurement of Productive Efficiency and Productivity Growth* (pp.3-91). Oxford University press.
35. Gonzalez, C. A. (2011) "Technical Efficiency of Organic Fertilizer in Small Farms of Nicaragua:1998-2005". *African Journal of Business Management* **5 (3)**, 967-973.
36. Greene, W. H. (1980) "Maximum Likelihood Estimation of Econometric Frontier Functions". *J Econ.* **13**, 27-56.
37. Greene, W. H. (2008) "The Econometric Approach to Efficiency Analysis". In H. Fried, C. Lovell, & S. Schmidt, *The Measurement of Productive Efficiency and Productivity Growth* (pp. 92-250). Oxford: Oxford University Press.
38. Hader, J., & Russel, W. R. (1969) "Rules for Ordering Uncertain Prospects". *American Economic Review* **59**, 25-34.
39. Hutchens, R. (2004) "One Measure of Segregation". *International Economic Review* **45 (2)**, 555-578.

40. Jensen, U., Gartner, H., & Rassler, S. (2006) "Measuring Overeducation with Earnings Frontiers and Multiply Imputed Censored Income Data". IAB Discussion Paper 11.
41. 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.
42. Kumbhakar, S. C., & Lovell, C. (2003) "*Stochastic Frontier Analysis".* Cambridge: Cambridge University Press.
43. Lab4. (2000) "Stochastic Dominance Testing :Short Course on Poverty & Development for Nordic Ph.D. Students". *University of Copenhagen* .
44. Landeau, S. S. (2003) "Chilean Labor Market Efficiency: An Earnings Frontier Approach". *Estudios de Economica* **30 (1)**, 87-102.
45. Lovell, C. A. (1993) "Production Frontier and Productive Efficiency". In H. O. Fried, C. A. Lovell, & S. S. Schmidt (Eds.). Oxford: Oxford University Press.
46. Madden, D. (2000) "Poverty in Ireland, 1987-1994: A Stochastic Dominance Approach". *The Economic and Social Review* **31 (3)**, 187-214.
47. Medeiros, M. (2006) "The Rich and the Poor: The Construction of an Affluence Line from The Poverty Line". *Social Indicators Research* **78**, 1-18.
48. Mincer, J. (1958) "Investment in Human Capital and the Personal Income Distribution". *Journal of Political Economy* **66 (4)**, 281-302.
49. 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.
50. Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets". *International Economic Review* **14**, 693-709.
51. Peichl, A. (2010) "Measuring Richness and Poverty: A Micro Data Application to Europe and Germany". *Review of Income and Wealth* **57 (3)**, 597-619.
52. Quirk, J., & Saposnik, R. (1962) "Admissibility and Measurable Utility Functions". *Review of Economic Studies* **29**, 140-146.

53. Ramos, X., & Silber, J. (2005) "On The Application of Efficiency Analysis to the Study of the Dimensions of Human Development". *Review of Income and Wealth* **51 (2)**, 285-309.
54. 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.
55. Ravallion, M. (1992) "Poverty Comparisons : A Guide to Concepts and Methods". Living Standards Measurement Study Working Paper 88.
56. Ravallion, M. (1994) "*Poverty Comparisons"*. Switzerland: Harwood Academic Publishers.
57. Richmond, J. (1974) "Estimating the Efficiency of Production". *International Economic Review* **15**, 515-521.
58. Rothschild, M., & Stiglitz, J. (1970) "Increasing Risks: A Definition". *Journal of Economic Enquiry* **2**, 225-243.
59. 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 .
60. Schultz, T. W. (1965) "*Transforming Traditional Agriculture"*. New Haven: Yale University Press.
61. Smith, A. (1759) "The Theory of Moral Sentiments".
62. Tax, S. (1953) "Penny Capitalism: A Guatemalan Indian Economy". *Smithsonian Institution Institute of Social Anthropology* **16**.
63. Tong, Y. K. (2009) "A Study of E-Recruitment Technology Adoption in Malaysia". *Industrial Management & Data Systems* **109 (2)**, 281-300.
64. Truett, D., & Roberts, B. (1973) "Classical Production Function, Technical Optimality and Scale Adjustments of the Firm". *American Economic Review* **63**, 975-982.
65. 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.
66. World Bank. (2003) "*Poverty Reduction Strategy Sourcebook"*.
67. Zadeh, L. A. (1965) "Fuzzy Set". *Information and Control* **8**, 338-353.

Tables

Table 1: Estimation of Earnings Frontier

Inputs	Estimates	t-values
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 2: Distribution of Poor and Non Poor over Efficiency Classes (Poverty Line: Rs 382.82)

Efficiency class	Proportion-poor	Proportion-non poor
0.1725 (1st percentile)	0.1002	0.8998
0.1752(2nd percentile)	0.1050	0.8950
0.1788(3rd percentile)	0.1164	0.8836
0.1904(4th percentile)	0.1370	0.8630
0.2133(5th percentile)	0.1461	0.8539
0.2498(6th percentile)	0.1690	0.8311
0.3000(7th percentile)	0.2420	0.7580
0.3436(8th percentile)	0.2968	0.7032
0.4920(9th percentile)	0.3676	0.6324
1.0000	0.2831	0.7169

Table 3: Distribution of Poor and Non Poor over Efficient and Inefficient Groups (Poverty Line: Rs 382.82)

Bench mark efficiency score	Efficient (E)			Inefficient (I)		
	FGT0	FGT1	FGT2	FGT0	FGT1	FGT2
Mean Efficiency Score	0.26565	0.04807	0.01305	0.11965	0.01462	0.00272
75th Percentile	0.31798	0.06222	0.01766	0.14234	0.01886	0.00392
95th Percentile	0.19027	0.05971	0.02386	0.19663	0.03068	0.00729

Table 4.Dominance analysis for region E and I (bench mark efficiency score =0.23*)

Poverty line	$D^1{}_E$	$D^1{}_I$	t-statistic
226.5179	0.0272	0.0013	-8.49
234.1937	0.0335	0.0022	-9.25
241.8695	0.0417	0.0046	-9.73
249.5453	0.0588	0.007	-11.48
257.2212	0.0731	0.01	-12.56
264.897	0.0836	0.0122	-13.33
272.5728	0.0935	0.017	-13.42
280.2486	0.1088	0.0192	-14.71
287.9244	0.1216	0.0225	-15.43
295.6003	0.1364	0.025	-16.51
303.2761	0.1507	0.0326	-16.61
310.9519	0.17	0.0401	-17.27
318.6277	0.1877	0.0515	-17.18
326.3036	0.2087	0.0628	-17.51
333.9794	0.227	0.0721	-17.91
341.6552	0.2475	0.0863	-17.9
349.331	0.2686	0.094	-18.82
357.0069	0.2856	0.1036	-19.14
364.6827	0.3099	0.1204	-19.27
372.3585	0.326	0.1361	-18.86
380.0343	0.3447	0.1519	-18.73
387.7102	0.366	0.1705	-18.56
395.386	0.3821	0.1892	-17.98
403.0618	0.3957	0.2082	-17.21
410.7376	0.4101	0.227	-16.56

418.4135	0.4261	0.2459	-16.1
426.0893	0.4396	0.2673	-15.21
433.7651	0.4521	0.2877	-14.37
441.4409	0.4646	0.3069	-13.68
449.1168	0.4777	0.3294	-12.75
456.7926	0.4894	0.3475	-12.14
464.4684	0.4989	0.3645	-11.45
472.1442	0.5093	0.389	-10.2
479.8201	0.5201	0.4062	-9.63
487.4959	0.5267	0.4253	-8.56
495.1717	0.5386	0.4498	-7.48
502.8475	0.5443	0.4701	-6.24
510.5234	0.5526	0.4895	-5.31
518.1992	0.5599	0.5053	-4.6
525.875	0.5639	0.5239	-3.37

*Bench mark efficiency score=mean efficiency score

Dominance is achieved at order 1.

Table 5. Dominance analysis for region E and I (bench mark efficiency score =0.25*)

Poverty line	$D^1{}_E$	$D^1{}_I$	t-statistic
224.2468	0.0295	0.0024	-7.46
231.9809	0.0384	0.0034	-8.45
239.7149	0.0449	0.0055	-8.76
247.449	0.0633	0.0079	-10.47
255.183	0.0821	0.0123	-11.64
262.9171	0.0973	0.0148	-12.72
270.6511	0.1085	0.0188	-13.12
278.3852	0.1255	0.0224	-14.13
286.1192	0.136	0.0238	-14.86
293.8533	0.1521	0.0294	-15.45
301.5874	0.1715	0.0363	-16.13
309.3214	0.1939	0.0458	-16.74
317.0555	0.2079	0.0576	-16.39
324.7895	0.2288	0.0681	-16.83
332.5236	0.2459	0.083	-16.49
340.2576	0.2659	0.0965	-16.59
347.9917	0.2855	0.1056	-17.19

355.7258	0.3046	0.1177	-17.44
363.4598	0.3251	0.1347	-17.32
371.1939	0.341	0.1526	-16.81
378.9279	0.3565	0.1691	-16.44
386.662	0.3801	0.1878	-16.55
394.396	0.3949	0.2079	-15.86
402.1301	0.4104	0.2262	-15.44
409.8642	0.4203	0.2462	-14.45
417.5982	0.4347	0.2647	-13.97
425.3323	0.4455	0.2867	-12.95
433.0663	0.4556	0.3072	-12.01
440.8004	0.466	0.3254	-11.32
448.5344	0.4783	0.3458	-10.61
456.2685	0.488	0.3663	-9.7
464.0025	0.4956	0.3832	-8.94
471.7366	0.5085	0.4033	-8.35
479.4707	0.5173	0.4222	-7.53
487.2047	0.5237	0.4405	-6.58
494.9388	0.5323	0.464	-5.39
502.6728	0.5372	0.4837	-4.23
510.4069	0.5448	0.501	-3.47
518.1409	0.551	0.5169	-2.69
525.875	0.5556	0.5332	-1.77

*Bench mark efficiency score=75th percentile value of state efficiency scores

Dominance is not achieved at order 1.

⇓

Table 6. Dominance analysis for region E and I with (bench mark efficiency score =0.25)

Poverty line	${D^2}_E$	${D^2}_I$	t-statistic
224.2468	0.7296	0.0372	-4.99
231.9809	0.9956	0.058	-5.92
239.7149	1.312	0.092	-6.72
247.449	1.7259	0.1451	-7.61
255.183	2.2926	0.2241	-8.71
262.9171	2.9967	0.3303	-9.82
270.6511	3.7961	0.4559	-10.79
278.3852	4.692	0.6163	-11.61
286.1192	5.7074	0.7941	-12.42
293.8533	6.8354	1.006	-13.17
301.5874	8.0955	1.2555	-13.9
309.3214	9.5099	1.5735	-14.59

317.0555	11.064	1.9621	-15.22
324.7895	12.7395	2.4599	-15.71
332.5236	14.5859	3.0507	-16.19
340.2576	16.5556	3.7451	-16.58
347.9917	18.7021	4.5309	-16.99
355.7258	20.9855	5.4013	-17.36
363.4598	23.4086	6.3717	-17.7
371.1939	25.9863	7.4871	-17.99
378.9279	28.6813	8.7386	-18.19
386.662	31.541	10.1175	-18.4
394.396	34.5285	11.6542	-18.54
402.1301	37.6351	13.3285	-18.64
409.8642	40.8428	15.1545	-18.67
417.5982	44.1503	17.1277	-18.66
425.3323	47.554	19.2643	-18.6
433.0663	51.0365	21.5504	-18.49
440.8004	54.5968	23.9964	-18.33
448.5344	58.2582	26.5933	-18.15
456.2685	62.0005	29.3594	-17.93
464.0025	65.8062	32.2638	-17.69
471.7366	69.7041	35.3051	-17.44
479.4707	73.6718	38.5057	-17.17
487.2047	77.6927	41.8442	-16.87
494.9388	81.7763	45.3437	-16.55
502.6728	85.9142	49.0076	-16.19
510.4069	90.0961	52.8133	-15.82
518.1409	94.3402	56.7474	-15.45
525.875	98.6206	60.8082	-15.06

Dominance is achieved at order 2.

Table 7. Estimation of (Treatment) Effect of Efficiency on Status of Poverty

Bench mark Efficiency*	Outcome ($status\ of\ poverty$)		Average Treatment effect (ATT)	Std. Error (ATT)	T value (ATT)
	Treated (T=1)	Control (T=0)			
50th Percentile	0.190	0.194	-0.004	0.007	-0.520
75th Percentile	0.180	0.178	0.002	0.009	0.180
90th Percentile	0.157	0.150	0.007	0.012	0.570
95th Percentile	0.135	0.140	-0.005	0.016	-0.320

*percentiles have been computed on state level efficiency scores.

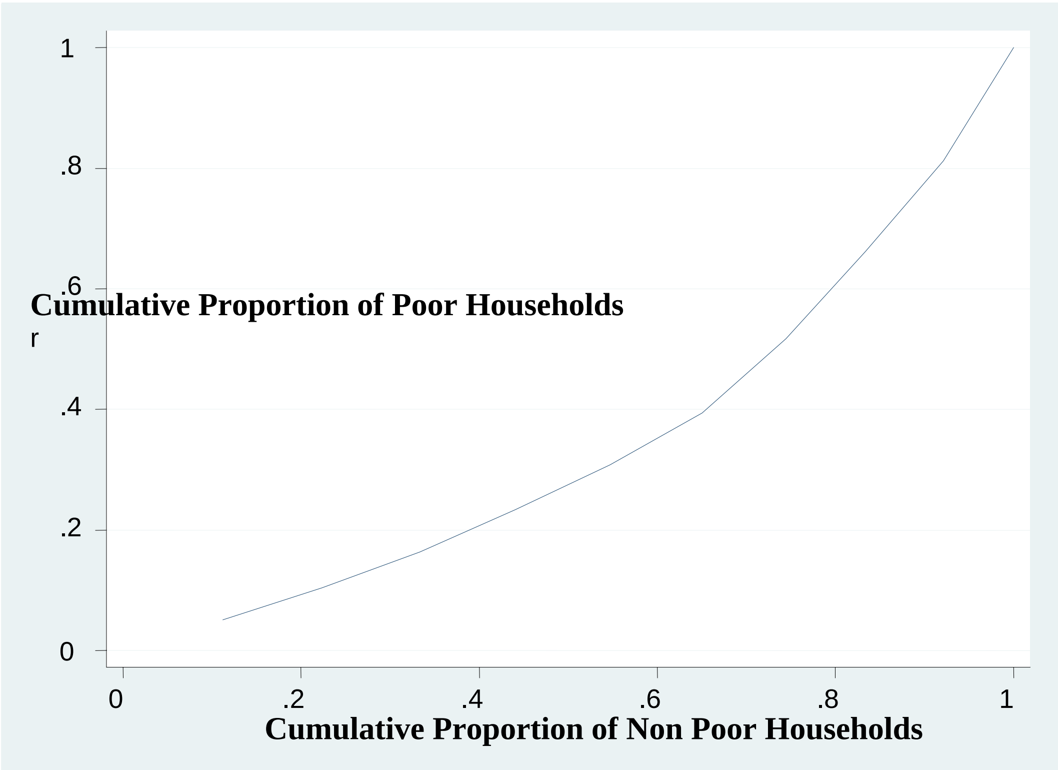

Figure 1: Poverty Segregation Curve