ELSEVIER
www.elsevier.com/locate/worlddev

World Development Vol. xx, pp. xxx–xxx, 2016

http://dx.doi.org/10.1016/j.worlddev.2016.08.004

Measuring Segregation of the Poor: Evidence from India

SHATAKSHEE DHONGDE*
Georgia Institute of Technology, Atlanta, USA

Summary. — Compared to the extensive literature on measurement of poverty, the question of how the poor are distributed regionally has received less attention. This paper fills the gap in the literature by providing a conceptual framework to measure inequality in the distribution of the poor. A poverty segregation curve is used to compare a region's share of the poor population with its share in the overall population. A unique contribution of the paper is formulating a generalized version of the poverty segregation curve. Unlike the segregation curve, the generalized segregation curve also takes average poverty rates into account while ranking distributions. Thus the generalized segregation curve may rank a distribution with substantially lower poverty rates above a distribution with higher poverty rates, when differences between their segregation curves are relatively small. The segregation curves are used to analyze changes in the distribution of the poor in India since the economic reforms in the early 1990s. Poverty rates and shares among all states, territories, and districts in India are estimated using data from the National Sample Survey Organization in 1994, 2004, and 2010. In the decades following the reforms India witnessed high growth rates and declining poverty rates. Despite the reduction in poverty, our analysis is the first to reveal that there was a significant rise in segregation of the poor over time. Some states had disproportionately high share of the poor compared to their share in the total population. Within states, the extent of segregation was lower among some of the poorest states and higher in less poor states. The generalized segregation curves show that the substantial decline in poverty rates since the reforms was not adequate to compensate the rise in segregation of the poor. Increase in segregation was also evident when the poor are identified as those who lacked education, health, and access to basic services. Reducing the segregation of the poor is important if India is to attain regionally balanced economic growth.

Key words — distribution, India, inequality, poverty, segregation

1. INTRODUCTION

There is a large and diverse body of literature studying several aspects of poverty, including definition, measurement, data availability, estimation methods, and its causes and effects. However the question of how the poor are distributed regionally has received less attention. When poverty is unevenly distributed across a country, targeting regions with higher poverty shares is more effective in lowering average poverty levels. There have been studies which undertake poverty mapping to display spatial clusters of poverty. This technique generates estimates of poverty for small geographic areas by combining survey data with census data when micro level data are not available (Elbers, Lanjouw, & Lanjouw, 2003). The poverty mapping methodology however does not provide a conceptual framework to measure segregation of the poor. This paper fills the gap in the literature by using an axiomatic approach to measure inequality in poverty distribution.

The paper makes three significant contributions to the field. First and foremost, it addresses the much ignored topic of regional inequality in the distribution of the poor by plotting poverty segregation curves. The segregation curve has been previously used in a variety of contexts. Early on, it was used by Duncan and Duncan (1955) to measure residential segregation, by Taeuber and James (1982) to measure racial segregation in schools, and to measure occupational segregation by gender (Hutchens, 1991) or by age (Dygalo, 2007). Jayaraj and Subramanian (2000, 2007) used segregation curves to measure the femaleness and the agedness of the population respectively. However we are aware of only Fujii (2008) who used the curve in the context of poverty. His concentration curve is similar to the poverty segregation curve. He used the curve to assess the potential gains from geographic targeting of certain welfare programs in Cambodia. [1] The poverty segregation curve compares a region's share of the poor population with its share in the overall population. The poor are regionally segregated when regions' share of the poor does not resemble their share in the overall population. Thus perfect integration (zero segregation) implies that each region has the same share in the poor and the overall population (poor and non-poor combined).

The paper's second and most important contribution is the formulation of a generalized version of the segregation curve. Although the segregation curve shares many similarities with the Lorenz curve, it has not yet been generalized like the latter (Shorrocks, 1983). The segregation curve compares the relative distribution of the poor while ignoring the average poverty levels. We propose a generalized segregation curve that takes into account the average poverty ratio. Thus a generalized segregation curve may rank a distribution with substantially lower poverty rates above a distribution with higher poverty rates, when differences between the distributions are relatively small and the segregation curves intersect. The generalized curve thus reveals the extent to which greater inequality in the distribution of the poor can be compensated by lower poverty levels.

Finally, the paper conducts a detailed analysis of segregation of poverty in India. Since the adoption of broad ranging economic reforms, India experienced rapid economic growth

* I am grateful to the International Association for Research in Income and Wealth (IARIW) for awarding the Nancy and Richard Ruggles Memorial Fund Prize for an early version of this paper at the IARIW Conference, Boston, 2012. Comments received from Dustin Chambers, Robert Hutchens, Prasanta Pattanaik, and Jacques Silber were useful. Finally, the editor, Arun Agrawal and three anonymous referees provided constructive comments which helped revise the manuscript. All remaining errors are my own. Final revision accepted: August 2, 2016.

Please cite this article in press as: Dhongde, S. Measuring Segregation of the Poor: Evidence from India, *World Development* (2016), http://dx.doi.org/10.1016/j.worlddev.2016.08.004

and reduction in poverty levels; though there is much debate about the precise estimates of poverty. Differences in poverty estimates largely originate from differences in methodological assumptions, use of poverty lines, and adjustments made to survey data (Deaton & Kozel, 2005). Despite the voluminous literature estimating poverty in India, relatively less attention has been paid to the regional distribution of the poor. Panagaria and Mukim (2014) estimated poverty in Indian states and among social and religious groups.[2] Jha and Sharma (2003) found that there is remarkable stability in the rankings of regions by poverty and that these rankings have not changed significantly since economic reforms. Dhongde (2006) concluded that spatial differences in average incomes rather than differences in relative distribution of income explained to a greater extent the spatial differences in poverty rates in Indian states. Krishna and Shariff (2011) used panel data on household incomes to examine the movement of households in and out of poverty among geographic regions. Datt and Ravallion (2011) found much stronger evidence of a feedback effect from urban economic growth to rural poverty reduction since the economic reforms. However this paper goes beyond estimating poverty rates in different states in India. It contributes to this literature by measuring segregation of the poor between and within the states in India. Despite the decline in poverty as evident in Panagaria and Mukim (2014) and other studies, the analysis in the paper for the first time reveals that there was a significant rise in the regional segregation of the poor over time.

The extent of segregation in India is measured by comparing plots of the segregation curve, and estimating the Gini index and the dissimilarity index of segregation. Changes in the regional distribution of the poor are measured in the two decades following the economic reforms. Segregation is measured across *all* states and territories, and districts within each state using data from the National Sample Survey (NSS) for the years 1993–94, 2004–05, and 2009–10. The NSS is the largest survey conducted and used to estimate official poverty statistics in India. The NSS data are used to plot segregation curves which compare states' share in the total poor versus their share in the total population. The dominance of the segregation curves is tested statistically by using non-parametric tests proposed by Barrett, Donald, and Bhattacharya (2014). Our results reveal that segregation of the poor in India increased significantly over a period of time. The substantial decline in poverty rates since the reforms was not adequate to compensate the rise in segregation of the poor. The rise in segregation was also evident when the poor were identified as those who lacked education, health, and access to basic services. Within states, the extent of segregation was lower in poorer states and higher in less poor states. In order to effectively reduce poverty levels in India, it is important to target states and districts with disproportionately high poor population.

The remainder of the paper is organized as follows. A poverty segregation curve is defined and its properties are discussed in Section 2. In Section 3, a generalized segregation curve is proposed. Data and background information on the Indian economy are summarized in Section 4. Section 5 contains discussion of segregation of the poor in India. Section 6 concludes. An Appendix to the paper contains discussion on alternative formulation of the generalized segregation curve.

2. A POVERTY SEGREGATION CURVE

Several notions of segregation have been proposed in the literature.[3] The poverty segregation curve is based on a nuanced notion of "evenness" in the distribution of the poor. The poverty segregation curve is not based on the idea of concentration or exposure of the poor. Unlike concentration, centralization or clustering, evenness does not conceive segregation as geographic isolation but as disproportionate shares of population groups. The poverty segregation curve does not argue for "equal" but for "proportional" numbers of poor across regions.[4] The poverty segregation curve uses the distribution of the overall population among regions as a benchmark and compares the extent to which the distribution of the poor deviates from this benchmark distribution.[5] Choosing the overall population is more meaningful and intuitive since standard poverty indices such as the headcount ratio compare the poor as a proportion of the total population.

(a) *Notation*

Let a country's total population T, at any given time, be divided into two mutually exclusive groups, namely, the poor T^P, and the non-poor T^N, so that $T = T^P + T^N$, and T^P, $T^N > 0$. We are interested in the distribution of the population across R different regions in the country, denoted by $i = 1, 2, \ldots R, R \geqslant 2$. In any region i, let x_i^P denote the number of poor, x_i^N denote the number of non-poor and x_i denote the region's population, so that $x_i = x_i^P + x_i^N$. Thus, the total poor in the country are given by $T^P = \sum_{i=1}^{R} x_i^P$, similarly the total non-poor in the country are given by $T^N = \sum_{i=1}^{R} x_i^N$ and the population is given by $T = \sum_{i=1}^{R} x_i$. A typical distribution X, of the poor relative to the total population across different regions can be summarized as: $X = \begin{bmatrix} x_1^P & x_2^P & \ldots & x_R^P \\ x_1 & x_2 & & x_R \end{bmatrix}$. The set of all possible distributions of the poor across regions, for a fixed value of R is given by D_R, and the union set for different R values is given by $D = \cup_{R \in I} D_R$ where $I \geqslant 1$ is a set of positive integers.

(b) *Definition*

A poverty segregation curve for $X \in D$, relates the cumulative proportion of the poor population $\left(\sum_{i=1}^{m} \frac{x_i^P}{T^P}\right)$ as a function of the cumulative proportion of the total population $\left(\sum_{i=1}^{m} \frac{x_i}{T}\right)$ in m regions combined, when regions are ordered in an ascending value of their poverty rate $\left(\frac{x_i^P}{x_i}\right)$.

There is no segregation if every region's share of the poor population is equal to its share of the total population i.e., $\left(\frac{x_i^P}{T^P} = \frac{x_i}{T}\right)$, implying each region's poverty ratio is exactly equal to the national poverty ratio, $\left(\frac{x_i^P}{x_i} = \frac{x_j^P}{x_j} = \frac{T^P}{T}\right) \forall i,\ j = 1, 2 \ldots R$. Conversely, if the poor are completely segregated then a region, say j's, entire population is poor $(x_j^P = x_j = T^P)$ and there are no poor in any other regions, $\forall i \neq j$, $(x_i^N = x_i)$ and $\left(\sum_{i \neq j} x_i^N = T^N\right)$ resulting in a reverse L-shaped segregation curve. The curve is piecewise linear, positively sloped, and convex in shape. It is bounded between $[0, 0]$ and $[1, 1]$. The greater the distribution of the poor resembles the distribution of the overall population, the closer the segregation curve is to the diagonal line. For any two distributions$(X, Y \in D)$, the poverty segregation curve of X '*dominates*' that of Y, denoted by $(X \geqslant_S Y)$, if and only if Y's segregation curve lies at some point below and at no point above the curve formed by X. If $(Y =_S X)$, then both the distributions are exactly alike and the segregation curves overlap.

Please cite this article in press as: Dhongde, S. Measuring Segregation of the Poor: Evidence from India, *World Development* (2016), http://dx.doi.org/10.1016/j.worlddev.2016.08.004