ISSN No. 2454 – 1427

CDE
December 2015

Employment and Earnings in Rural India: 2004-2012

Shantanu Khanna
Email: shantanu@econdse.org
Department of Economics
Delhi School of Economics

Deepti Goel
Email: deepti@econdse.org
Department of Economics
Delhi School of Economics

René Morissette
Email: rene.morissette@statcan.gc.ca
Statistics Canada

Working Paper No. 250

http://www.cdedse.org/working-paper-frameset.htm

CENTRE FOR DEVELOPMENT ECONOMICS
DELHI SCHOOL OF ECONOMICS
DELHI 110007

Employment and Earnings in Rural India: 2004-2012*

Shantanu Khanna
University of Delhi
(Email: shantanu@econdse.org)

Deepti Goel
Delhi School of Economics
(Email: deepti@econdse.org)

René Morissette
Statistics Canada
(Email: rene.morissette@statcan.gc.ca)

December 2015

* The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 290752.
The views expressed in this paper do not reflect the views of Statistics Canada.

Abstract: This paper analyzes the changes in employment and earnings of paid workers in rural India from 2004/05 to 2011/12. While the employment rate of adults remained stable at 51 percent during this period, it increased for men and fell for women. Real earnings of wage earners increased at all percentiles, and the percentage increase was larger at the lower end. Consequently, earnings inequality declined. Unconditional quantile regressions reveal that education contributed positively to the increase in earnings, but it also increased inequality. Recentered Influence Function decompositions reveal that throughout the earnings distribution, except at the very top, both changes in 'characteristics' and in 'returns to these characteristics' tended to increase earnings, with changes in returns playing a bigger role. Decomposing inequality measures reveals that in spite of the change in characteristics having had an inequality increasing effect, inequality fell mainly because workers at lower deciles experienced a greater improvement in returns to their characteristics than those at the top.

JEL Codes: J30, J31, O53

Keywords: Earnings, Inequality, Wage Distribution, Rural India

1 Introduction

In their discussion of India's economic growth during 1980-2004, Kotwal et al (2011) point to the existence of two Indias: "One of educated managers and engineers who have been able to take advantage of the opportunities made available through globalization and the other—a huge mass of undereducated people who are making a living in low productivity jobs in the informal sector—the largest of which is still agriculture." This paper is about this second India, which mainly resides in its rural parts. Another interesting feature the authors highlight that distinguishes India from other Asian countries is that even though its non-farm sector has shown high growth rates, the share of agriculture in total labor force has declined very slowly. Agriculture continues to employ the largest share of the workforce, but its contribution to gross value added is much smaller; it is also the slowest growing sector of the economy.[1] Given these facts, the concern about whether the benefits of high overall GDP growth have trickled down to those at the bottom is perhaps even more pertinent for rural India. Hasan and Gupta (2014) also conjecture that one of the reasons for slow poverty reduction in the post-reform period[2] is that growth in rural India, which accounts for the majority of India's poor, has been considerably slower than growth in urban India. In this paper, we therefore focus on rural India and examine how some aspects of the rural labor market have evolved over the seven-year period between 2004/05 and 2011/12.

We first document some broad changes in employment, and then examine how the individual earnings (in real terms) distribution has evolved over time. Several studies have documented that along with the high growth rates of GDP that have characterized the Indian economy since the 1980s, there has been an increase in inequality.[3] However, the studies that have tracked inequality have either focused on consumption expenditure (Sen and Himanshu 2004; Cain et al 2010; Motiram and Vakulabharanam 2013; Subramanian and Jayaraj 2015), or on earnings of paid workers in urban India (Kijima 2006; Azam 2012). We fill the gap by looking at how inequality in the earnings of paid workers has evolved in rural India.[4]

We document changes in real weekly earnings not just at the mean but also at various percentiles. It is important to do so because several studies have found that earnings inequality is increasingly concentrated at the upper quantiles. For India, Azam (2012) and Kijima (2006) find this for urban wage earners, and Banerjee and Piketty (2005) find it for income tax payers. We also use unconditional quantile regressions (as opposed to conditional quantile regressions) to account for the impact of changes in

[1] According to the Economic Survey, in 2011, the employment shares of agriculture, industry, and services, were 48.9, 24.3 and 26.9 percent, respectively, whereas their shares in Gross Value Added (at factor cost, current prices) were 18.9, 32.9, and 48.2 percent, respectively (GOI 2015). Also, between 2004/05 and 2011/12, Gross Domestic Product (at factor cost, 2004/05 prices) in these sectors grew at 4.2, 8.5 and 9.6 percent per annum, respectively (authors calculations based on RBI 2015).

[2] India initiated a number of measures toward trade and foreign direct investment (FDI) liberalization starting in 1991. Post-reform refers to the period since the early 1990s.

[3] A notable exception is Dutta (2005). For the period, 1983-99, at the all-India level she finds an increase in wage rate inequality among regular salaried workers, but a *decrease* among casual labor.

[4] Compared to earnings, there may be some advantages in looking at consumption data (Goldberg and Pavcnik 2007). The latter may be a better measure of lifetime wellbeing because they encompass intertemporal smoothing, and because they suffer from fewer reporting errors. For the US, while Krueger and Perri (2006) argue that consumption inequality is lower than income inequality, Aguiar and Bils (2015) apply a correction to consumption data and show that once corrected the two are much closer. For urban India, Goldberg and Pavcnik (2007) note that consumption and wage inequality have moved in the same direction. We feel that it is important to juxtapose the two to get a complete picture.

workers' characteristics on the unconditional quantiles of the earnings distribution. Finally, we use the Recentered Influence Function (RIF) Decomposition developed by Firpo, Fortin, and Lemieux (2009) to divide the overall change in earnings inequality into a structure effect (due to change in returns to personal characteristics) and a composition effect (due to change in the distribution of personal characteristics).

We find that during the period from 2004 to 2012, real earnings among wage earners increased for all percentiles, and in percentage terms they increased by more at the lower end. Consequently, earnings inequality declined in rural India. Unconditional quantile regressions reveal that education contributed positively to the increase in earnings, but it had an inequality increasing effect. The RIF decompositions reveal that throughout the earnings distribution, except at the very top, both the composition effect and the structure effect tended to increase earnings, with changes in the latter having played a bigger role. Decomposing various inequality measures reveals that in spite of the composition effect having had an inequality-increasing role, inequality fell mainly because workers in the lower deciles experienced a greater improvement in returns to their characteristics than those at the top.

The rest of the paper is organized as follows. Section 2 discusses the methodology used to analyze the change in earnings. Section 3 describes the data and the analysis sample. Section 4 presents the results, first documenting broad employment patterns in rural India and then analyzing the earnings distribution of paid workers/wage earners. Section 5 concludes.

2 Methodology

Below we briefly explain the methodology used to analyse the change in earnings over time. For a detailed exposition of this and other decomposition techniques, see Fortin et al. (2011).

2.1 Unconditional Quantile Regressions

Unconditional quantile regressions (UQR, Firpo et al., 2009) help us examine the marginal effects of covariates on the unconditional quantiles of an outcome variable. UQR differ from the traditional quantile regressions (Koenker and Bassett, 1978) in that the latter examine the marginal effects of covariates on the *conditional* quantiles. For instance, if we observe that the quantile regression coefficients for college education are larger as we move from the first to the ninth decile, we can say that college education increases earnings dispersion within a group of individuals having the same vector of covariate values. However, in order to claim that college education increases overall earnings dispersion among all individuals (irrespective of their covariates), we need to rely on unconditional quantile regressions. To understand UQRs we begin with the concept of an Influence Function (IF).

The IF of any distributional statistic represents the influence of an observation on that statistic. The Recentered Influence Function (RIF) is obtained by adding back the statistic to the IF. Specifically, let w denote earnings, then in case of $q_\theta = Q_\theta(w)$, the θ^{th} quantile of the unconditional earnings distribution,

$$IF(w, q_\theta) = (\theta - \mathbb{I}\{w \leq q_\theta\})/f_w(q_\theta) \quad \textbf{\{1\}}$$

where $\mathbb{I}\{.\}$ is an indicator function and f_w is the density of the marginal distribution of earnings. The RIF in this case is given by:

$$RIF(w, q_\theta) = q_\theta + IF(w, q_\theta) = q_\theta + (\theta - \mathbb{I}\{w \leq q_\theta\})/f_w(q_\theta) \quad \textbf{\{2\}}$$

Note that the expected value of the RIF will be q_θ itself.

The conditional expectation of the RIF modelled as a function of the explanatory variables, $\boldsymbol{X}$, gives us the unconditional quantile regression (or RIF regression) model:

$$E[RIF(w, q_\theta)|\boldsymbol{X}] = m_\theta(\boldsymbol{X}) \qquad \textbf{\{3\}}$$

In its simplest form,

$$E[RIF(w, q_\theta)|\boldsymbol{X}] = \boldsymbol{X\beta} + \varepsilon \qquad \textbf{\{4\}}$$

where $\boldsymbol{\beta}$ represents the marginal effect of $\boldsymbol{X}$ on the quantile, q_θ, and can be estimated by Ordinary Least Squares (OLS). Thus, in UQR or RIF-regressions, the dependent variable is replaced by the corresponding RIF of the statistic of interest.

2.2 RIF Decomposition

The RIF decomposition divides the overall change in any distributional statistic into a structure effect (due to the changes in returns to characteristics/covariates), and a composition effect (due to the changes in the distribution of covariates). Compared to other decomposition methods such as the Machado-Mata (Machado and Mata 2005), the RIF decomposition has the added advantage of further dividing the structure and composition effects into the contribution of each covariate. In this way, it is closest in spirit to the decomposition method proposed by Blinder (1973) and Oaxaca (1973).

To see this, note that in the case of quantiles, the RIF is estimated based on the sample quantile $\widehat{q_\theta}$, and estimating the density $f_w(\widehat{q_\theta})$ using kernel methods (Firpo et al., 2009). Plugging these values in equation **{2}** gives $\widehat{RIF}(w, q_\theta)$. The RIF regressions are estimated replacing the usual dependent variable values with $\widehat{RIF}(w, q_\theta)$. The resulting RIF coefficients for each year (T) would be:

$$\hat{\beta}_{T,\theta} = (\textstyle\sum_{i \in T} X_i \,.\, X_i')^{-1} \sum_{i \in T} \widehat{RIF}(w_{Ti}, q_{T\theta}) \,.\, X_i \,, \; T\text{=1, 2} \qquad \textbf{\{5\}}$$

The decomposition for any unconditional quantile θ would be:

$$\widehat{\Delta}^{\theta}_{Total} = \underbrace{\bar{X}_2(\hat{\beta}_{2,\theta} - \hat{\beta}_{1,\theta})}_{\widehat{\Delta}^{\theta}_{Structure}} + \underbrace{(\bar{X}_2 - \bar{X}_1)\hat{\beta}_{1,\theta}}_{\widehat{\Delta}^{\theta}_{Composition}} \qquad \textbf{\{6\}}$$

Further, to examine the contribution of each covariate in the overall composition effect, the latter can be written as follows:

$$\widehat{\Delta}^{\theta}_{Composition} = \textstyle\sum_{k=1}^{K} (\bar{X}_{2k} - \bar{X}_{1k})\hat{\beta}_{1k,\theta} \qquad \textbf{\{7\}}$$

Though the above discussion on RIF-Decomposition focused on quantiles, it is also applicable to any other distributional statistic. We present the RIF decomposition for quantiles as well as the Gini.

3 Data

We use two rounds of the Employment Unemployment Survey (EUS) conducted by the National Sample Survey Organization (NSSO) for the years 2004/05 and 2011/12. Our target population is individuals between the ages of 15 and 64, living in rural areas of 23 major states of India.[5] We exclude those who were either disabled or enrolled full time in educational institutions.

[5] The 23 states are Andhra Pradesh, Assam, Bihar, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Manipur, Meghalaya, Orissa, Punjab,

In the second part of the analysis, we focus on wage earners.[6] Nominal earnings are converted into real terms (2004/05 prices) using consumer price indices provided by the Labour Bureau, Government of India.[7] We also trim the real earnings distribution of each year by dropping 0.1 percent of observations from the top and the bottom.[8] Ultimately, our analysis sample consists of approximately, 44,500 workers in 2004/05 and 36,000 workers in 2011/12. This corresponds to about 103 million workers in 2004/05 and about 117 million in 2011/12.

4 Results

We first present broad changes in employment and then the findings related to the evolution of the earnings distribution.

4.1 Changes in Employment

Table 1 shows, for each survey year, the percentage of the working age (15-64) population of rural India that is employed as wage earners or self-employed.[9] The data are presented separately by age, gender, marital status, religion, caste and education.

Before turning to changes in employment, we note three features of the rural labour market in India. First, there is a big difference in the employment rates of men and women: In 2011/12, 82 percent of the men were employed, whereas this figure is only 21 percent for women. Second, employment rates are higher among Scheduled Tribes (ST) and Scheduled Castes (SC) than among Other Backward Classes (OBC) and individuals in the 'Others' category.[10] Third, employment rates are highest among the most educated.[11] In 2011/12, 62 percent of those with a college or beyond education were employed, whereas this figure was only 45 percent for the illiterates. Also, in the same year, in the college and beyond category, 39 percent were wage earners, while only 26 percent of illiterates were paid workers. These large differences in employment rates between educational groups are indicative of underlying inequality in access to earnings opportunities.

Rajasthan, Tamil Nadu, Tripura, Uttar Pradesh, Uttaranchal, and West Bengal. In 2004/05, there were 28 states and 7 union territories in India. We excluded the states and union territories for which there were no deflators.

[6] The main reason for restricting to wage earners is that EUS does not collect earnings data for self-employed individuals. In her analysis of earnings inequality for urban India, Kijima (2006) imputes the earnings of the self-employed using Mincerian equations estimated on the sample of regular wage/salaried workers. We refrain from this imputation as it imposes identical returns to covariates for both sets of workers, an assumption that may not be true.

[7] We use the Consumer Price Index for Rural Labourers (CPI-RL), the relevant price index for rural areas.

[8] We do this in order to remove outliers and possible data entry errors.

[9] The EUS records an individual's main activity during the year (principal activity status) as well as during the week (weekly activity status), preceding the survey. We use the weekly activity status to classify individuals as self-employed or wage earners.

[10] Data on caste are available by broad administrative categories: Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Classes (OBC) and 'Others'. The first three groups are beneficiaries of affirmative action, awarded to them by the state on account of accumulated disadvantage, and in the case of SC and ST, added stigmatization on account of their caste/tribe status. The 'Others' are a heterogeneous residual category, a rough proxy for upper castes (Deshpande 2011).

[11] We use five education categories throughout the paper: illiterates, primary and middle school, secondary (grade 10 graduates), higher secondary (grades 12 graduates), and college and beyond (college graduates and post-graduates).

Turning to changes over time, overall there was no change in the employment rate over the seven-year period: it stood at 51 percent in both years.[12] There was a marginal increase (about one percentage point) in the share of wage earners, and a corresponding decrease in the share of self-employed. Next we try to identify particular sub-groups that may have seen drastic changes in their employment shares.

There was a 3.3 percentage point increase in the employment rate of men, whereas the employment rate of women decreased by 3.5 percentage points. For men the change was driven by an increase in paid work, and for women it was due to a decrease in both paid work and self-employment. Employment fell among the ST (by 3.2 percentage points) and this was entirely due to a decrease in paid work. We see that paid work among Muslims increased (by 3.9 percentage points). Finally, among the two highest education categories (those with higher secondary education or above), the share of self-employment declined (between 4 to 5 percentage points).

Table 1: Employment Status by Sub-categories, Rural India

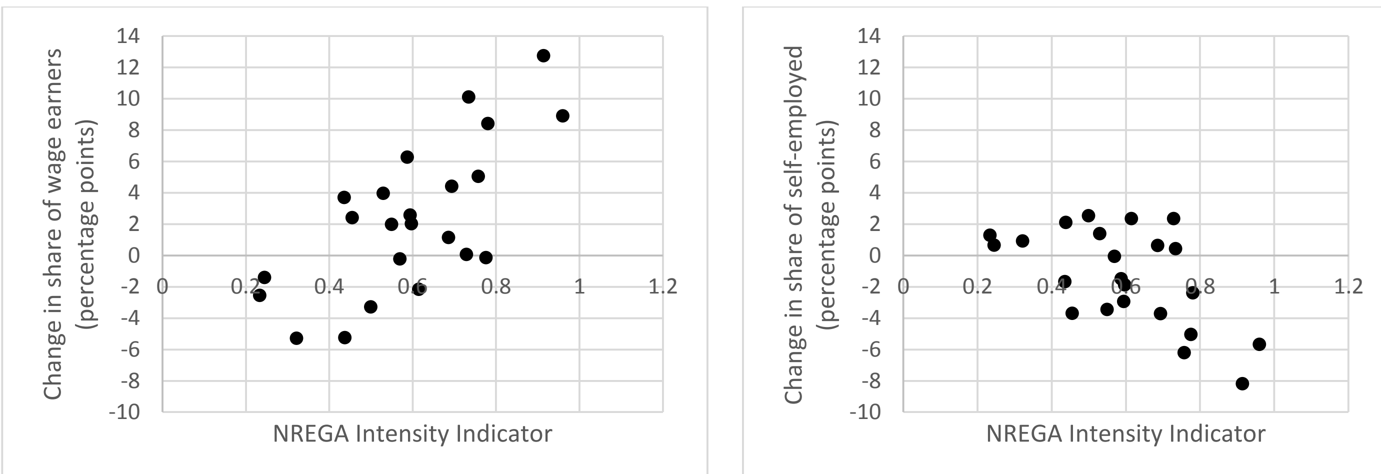
		2004/05			2011/12		
		Wage Earners	Self-Employed	Employed	Wage Earners	Self-Employed	Employed
All Individuals		27.0	24.2	51.2	28.2	22.9	51.1
Age	15-24	26.9	8.9	35.8	27.0	7.6	34.6
	25-54	28.5	28.3	56.8	29.8	25.7	55.5
	54-64	17.6	34.6	52.2	20.8	31.7	52.5
Gender	Women	15.9	8.7	24.6	13.8	7.3	21.1
	Men	38.6	40.5	79.1	43.3	39.1	82.4
Marital Status	Married	25.5	26.9	52.4	26.5	25.2	51.7
	Not Married	32.3	14.7	47.0	35.3	12.8	48.1
Religion	Muslims	21.1	25.0	46.1	25.0	23.4	48.4
	Non-Muslims	27.7	24.1	51.8	28.6	22.8	51.4
Caste	ST	33.9	22.2	56.1	30.7	22.2	52.9
	SC	39.3	18.2	57.5	39.6	16.9	56.5
	OBC	24.0	25.6	49.6	26.2	23.3	49.5
	Others	19.2	27.8	47.0	20.9	27.5	48.4
Education	Illiterate	27.3	20.2	47.5	26.4	18.9	45.3
	Primary and Middle	26.8	27.3	54.1	29.7	24.7	54.4
	Secondary	23.1	29.4	52.5	25.4	28.3	53.7
	Higher Secondary	23.2	29.8	53.0	24.9	25.7	50.6
	College and Beyond	36.6	28.4	65.0	39.1	23.1	62.2

In 2006, the central government introduced an employment guarantee scheme, the Mahatma Gandhi National Rural Employment Guarantee Scheme (NREGA) under which every rural household is guaranteed

[12] In 2011/12, among those not employed, 2.2 percent was unemployed and 46.8 percent was not in the labour force. These shares are similar in 2004/05 as well.

a hundred days of unskilled manual work in a year at state minimum wages. Using data from the EUS itself, we rank the 23 states according to the intensity of implementation seen in 2011/12. The NREGA intensity indicator is calculated as the share of persons within a state who got work under the scheme in the year prior to the survey. In Figure 1 we present a scatter plot with changes in shares for each employment category (wage earners and self-employed) over the period plotted against the NREGA intensity indicator for all 23 states. In the left panel we see a positive association between NREGA intensity and the change in share of wage earners, while in the right panel we see a negative relationship between NREGA intensity and the change in the share of self-employed. This could be indicative of a shift out of self-employment into wage employment in states where the NREGA was functioning better, though the link may or may not be causal.

Figure 1: Change in Employment Shares versus NREGA Intensity in 2011/12, All States



Next, we wanted to see if there was any relationship between changes in work status and how rich/poor a state was on average. Figure 2 presents a scatter plot with changes in work status over the period plotted against the average monthly per capita expenditures (MPCE) in 2004/05 for the 23 states. We do not observe any discernible pattern from either panel.

Figure 2: Changes in Employment Shares versus Average Monthly per Capita Expenditure in 2004/05, All States

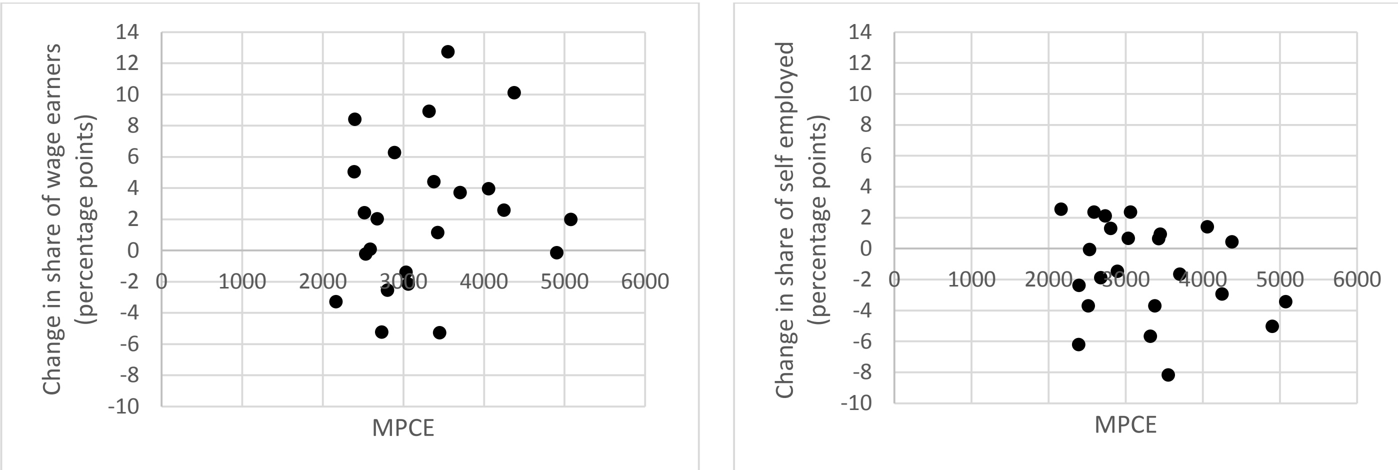


Table 2 presents the results of a multinomial logit model with the outcome variable as the three labour force participation categories – wage earners, self-employed and not employed. We control for all available personal characteristics, namely, age, gender, marital status, religion, caste, education and state of residence, and include a year of survey dummy.[13] Since we are interested in analyzing changes over time, we are mainly interested in the marginal effect of year of survey on each category. This tells us the transition probabilities over time that the changes in the distribution of personal characteristics cannot explain. As seen in Table 2, controlling for various personal characteristics, a worker is 2.8 percentage points less likely to be self-employed and 2.2 percentage points more likely to be a wage earner in 2011/12 relative to 2004/05.[14] Thus, over time there was some movement away from self-employment toward paid work, though the magnitude of this trend is small.

Table 2: Transition Probabilities (2004/05 to 2011/12), Multinomial Logit

Participation Status	Share in 2004/05 (%)	Marginal Effect of Time	Standard Error
Wage Earners	26.4	0.022	0.002
Self-Employed	25.0	-0.028	0.002
Not Employed	48.5	0.006	0.002

4.2 Changes in the Distribution of Earnings from Paid Work

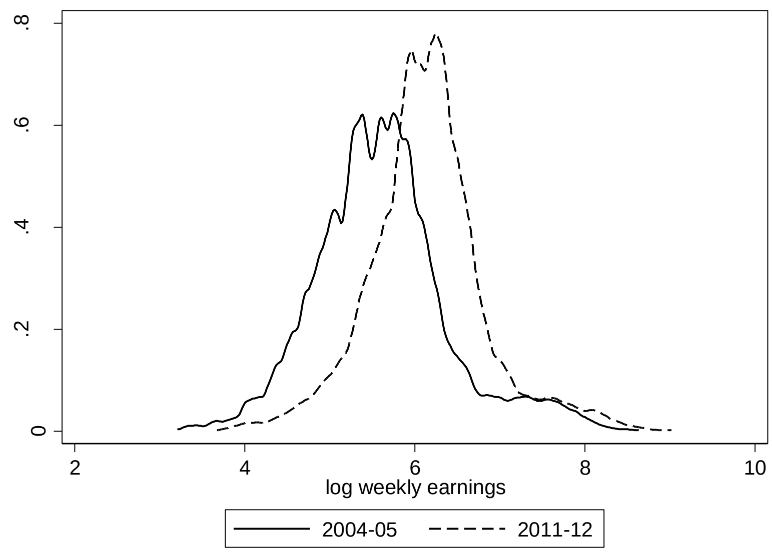
In the rest of the paper, we focus solely on wage earners. As seen in Table 1 they constituted 27 and 28 percent of the working age population in rural India in 2004/05 and 2011/12, respectively. Figure 3 presents the kernel density estimates of the log of real weekly earnings for 2004/05 and 2011/12. The earnings density for each year is skewed to the right wherein the median is less than the mean. Over the

[13] These covariates are the same as the ones used later in the paper for the Mincerian wage regressions and decompositions.

[14] These results hold even when we categorize individuals into the three categories using their principal activity status instead of the weekly activity status.

seven-year period the earnings density shifted to the right and became more peaked (less dispersed). The mean real weekly earnings increased from 392 to about 605 rupees, while median real weekly earnings increased from 265 to 458 rupees. The all-India rural poverty line (defined in terms of minimum consumption expenditure needed to meet a specified nutritional and living standard) for 2004/05 was 447 rupees per capita per month (Planning Commission 2014). Multiplying weekly earnings by four, the mean (median) real monthly earnings in 2004/05 was 3.5 (2.4) times the poverty line, and in 2011/12 it was 5.4 (4.1) times this value.

Figure 3: Earnings Densities, 2004/05 and 2011/12



4.2.1 Changes in Earnings Inequality

Figure 4 plots the real weekly earnings (in rupees) at each percentile for 2004/05 and 2011/12. At each percentile, earnings were higher in 2011/12 than in 2004/05. The gap between the two curves reveals that the increase in earnings was, in absolute terms (i.e. measured in rupees), greater for higher percentiles. For instance, real weekly earnings increased by 98 rupees at the first decile, 193 rupees at the median and 307 rupees at the ninth decile.

Yet the percentage increase in earnings was much greater at the lower end of the distribution than at the top.[15] For instance, earnings increased by 89 percent at the first decile, 72 percent at the median, and by 44 percent at the ninth decile (Figure 5). This suggests that earnings inequality—which is defined in relative terms rather than absolute terms—declined over the seven-year period.

---

[15] Using consumption expenditure data (also collected by the NSSO), for the period between 2004/05 and 2009/10, Subramanian and Jayaraj (2015) find the same pattern of an increase in real consumption expenditures at all deciles for rural India, with the highest growth occurring at the third and fourth deciles. For urban India also, they find an increase at all deciles, but the growth rates monotonically increase from lower to upper deciles.

Figure 4: Real Weekly Earnings, by percentile, 2004/05 and 2011/12

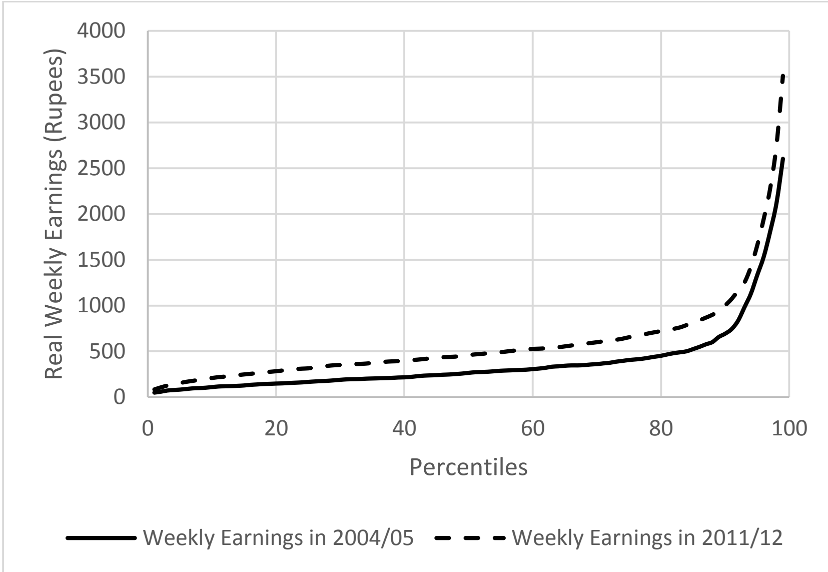


Figure 5: Change in Log Real Weekly Earnings, by percentile, 2004/05 to 2011/2012

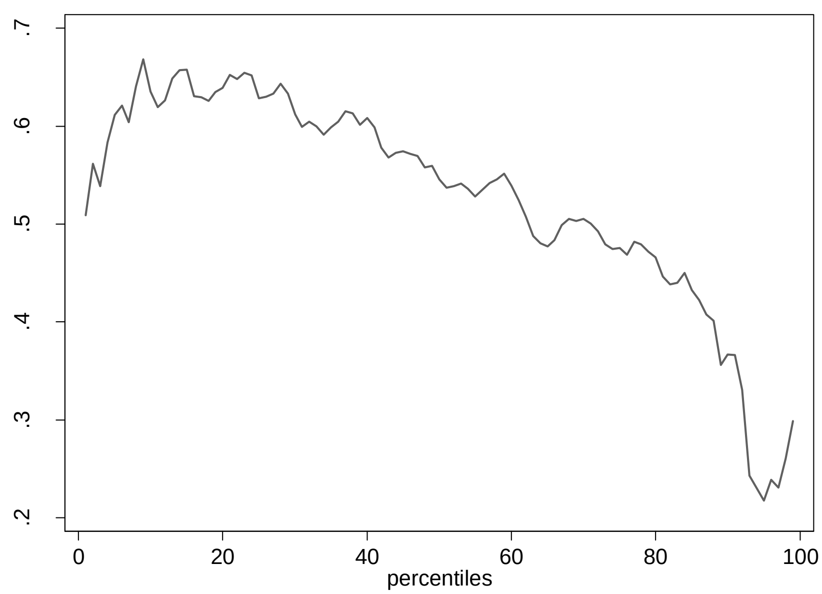


Figure 6 confirms this: it shows that the Lorenz curve of weekly earnings for 2011/12 lies above the one for 2004/05, unambiguously indicating a fall in inequality.

Figure 6: Lorenz Curves of Real Weekly Earnings, 2004/05 and 2011/12

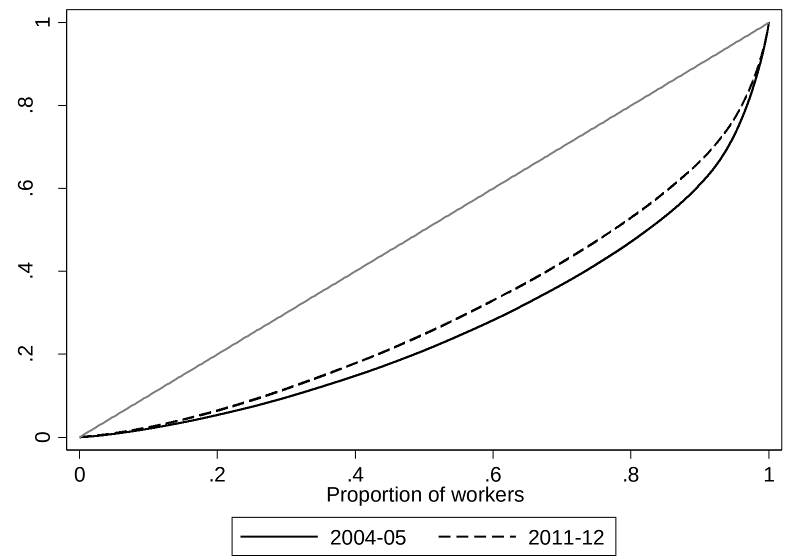


Table 3 supplements Figures 4, 5 and 6 and shows how various inequality measures changed over time. The ratio of the (raw) earnings at the twenty-fifth to the tenth percentile was steady at about 1.52. At the middle of the distribution, there was some decrease in inequality as measured by the ratio at the sixtieth to the fortieth percentile. In contrast, the ratio at the ninetieth to the seventy-fifth percentile fell very sharply from 1.71 to 1.53. Thus, the decrease in inequality was more due to the changes at the top and middle of the distribution than at the bottom.

Table 3: Inequality Measures for Real Weekly Earnings from Paid Work

	2004/05	20011/12
25-10	1.52	1.51
60-40	1.41	1.32
90-75	1.71	1.53
Variance of log Earnings	0.61	0.48
Gini	0.461	0.395

The decrease in inequality is also reflected in the variance of log earnings and in the Gini coefficients. The Gini of real weekly earnings from paid work fell from 0.461 to 0.395.[16] Subramanian and Jayaraj (2015) use consumption expenditure data (also from the NSSO) and find that between 2004/05 and 2009/10, the Gini declined from 0.305 to 0.299 in rural India. For urban India, it increased from 0.376 to 0.393. It is noteworthy that the direction of change in inequality that they find using consumption expenditure is the same as ours even though their period of study is shorter.[17]

[16] If we consider daily wage rates instead of real weekly earnings, the Gini fell from 0.398 to 0.358. This indicates that it is wage rates, and not so much the time spent working, that is driving the decrease in earnings inequality. We study this in detail in the next sub-section.

[17] Inequality measures using NSSO data are underestimates as the surveys are not equipped to capture adequate numbers of the very rich (Banerjee and Piketty 2005). Nonetheless, it is reasonable to use these data to make

4.2.2 Wage Rates or Days Worked: Decomposition of the Variance in Log Earnings

So far our analysis has been about weekly earnings. The EUS also collects data on the number of half-days worked within the corresponding week to acquire these earnings. The following equations illustrate the decomposition of the variance of log earnings to study the separate contributions of wage rates and time spent working.

$$Weekly\ Earnings\ (E) = Average\ Daily\ Wage\ Rate(W) * Number\ of\ days\ worked\ (D)$$

$$E = W * D$$

$$\ln(E) = \ln(W) + \ln(D)$$

$$\underbrace{Var[\ln(E)]}_{1} = \underbrace{Var[\ln(W)]}_{2} + \underbrace{Var[\ln(D)]}_{3} + \underbrace{2 * Covariance[\ln(W), \ln(D)]}_{4}$$

Thus, the variance of log earnings can tell us how much of the weekly earnings dispersion (1) can be accounted by the dispersion in wage rates (2), the dispersion in workdays (3), and the co-movement of wage rates and workdays (4). We implement this decomposition for both survey years, and then calculate the difference between them to assess how much of the change in inequality in weekly earnings (where inequality is measured by variance of log weekly earnings), can be accounted for by:

- change in inequality in wage rates
- change in inequality in workdays
- change in co-movement of wage rates and workdays[18]

We present the results of the decomposition in Table 4.

Table 4: Decomposition of Variance of Log Earnings

Year	$Var[\ln(E)]$	$Var[\ln(W)]$	$Var[\ln(D)]$	$2 * Cov[\ln(W), \ln(D)]$
2004/05	0.61	0.43	0.13	0.06
2011/12	0.48	0.36	0.09	0.03
Change over time	-0.14	-0.07	-0.04	-0.03

In both years, in an accounting sense, the inequality in log earnings is mainly because of inequality in daily wages rates rather than inequality in days worked or because highly paid workers worked for longer time: In both years, over 70 percent of inequality in log earnings is due to inequality in wage rates.[19] The covariance between wage rates and days worked is positive, which means that it is indeed the case that highly paid workers also tend to work more days in the week. Inequality declined over time as seen in the decrease in the variance in log earnings. About 50 percent of this decline is due to a decline in inequality of wage rates. The rest is due to a decrease in inequality of days worked, and a weaker relationship between highly paid workers working more number of days.

---

conclusions about trends assuming that the nature of bias remains the same across surveys separated by short periods of time.

[18] Although the variance of log weekly earnings allows us to quantify a 'wage rate' effect, a 'workdays' effect, and a 'covariance effect', it does not necessarily fall when one rupee is transferred from a rich worker to a poor one. However, this limitation is inconsequential since we have shown (using the Lorenz curves) that inequality has unambiguously fallen over time.

[19] There is both a lower and an upper limit to the number of days worked (from half a day to seven days in a week), and that may contribute to the variance of log days worked being relatively low.

4.3 Earnings Regressions

We estimate earnings regressions (both OLS and RIF) with the log of real weekly earnings as the dependent variable separately for the years 2004/05 and 2011/12. The explanatory variables include all the personal characteristics included in the multinomial logit discussed earlier in section 4.1. Age enters the model in a quadratic form as a proxy for work experience. 'Others' and illiterates are the omitted categories for caste and education, respectively.

Before moving to the regression results, we present some descriptive statistics. From Table 5, we note that mean (log) weekly earnings increased over the period. The average age among wage earners also increased, perhaps an indication of later entry into the labour market as more people acquire higher education. There was also an increase in the share of males, married and Muslim. The proportion of those belonging to ST and SC declined. Education levels rose significantly: The proportion of illiterates decreased by around 11 percentage points, and the proportion in each schooling level and with a college education increased.

Table 5: Descriptive Statistics, Wage Earners in Rural India

	2004/05	2011/12
Number of Observations	44,477	35,946
Mean log Real Weekly Earnings (Std. dev.)	5.61 (0.78)	6.14 (0.69)
Mean Age (Std. dev.)	34.1 (11.72)	35.9 (11.69)
Male (%)	69.9	75.0
Married (%)	74.3	76.3
Muslim (%)	8.4	10.9
Caste Categories (%)		
ST	12.9	12.0
SC	30.8	28.9
OBC	37.9	41.5
Others	18.3	17.7
Education Categories (%)		
Illiterate	47.0	35.7
Primary and Middle	39.4	43.9
Secondary	6.1	9.4
Higher Secondary	2.9	4.6
College and Beyond	4.6	6.4

Next, we turn to OLS, and RIF regressions at the nine deciles, estimated separately for each year. Figures 7 and 8 plot these regression coefficients. The left column of plots are for 2004/05, and the right for 2011/12. For each selected covariate, RIF regression coefficients are plotted against the corresponding deciles. The dashed lines represent the 95 percent confidence interval of the coefficients. The solid horizontal lines are the OLS coefficients.[20]

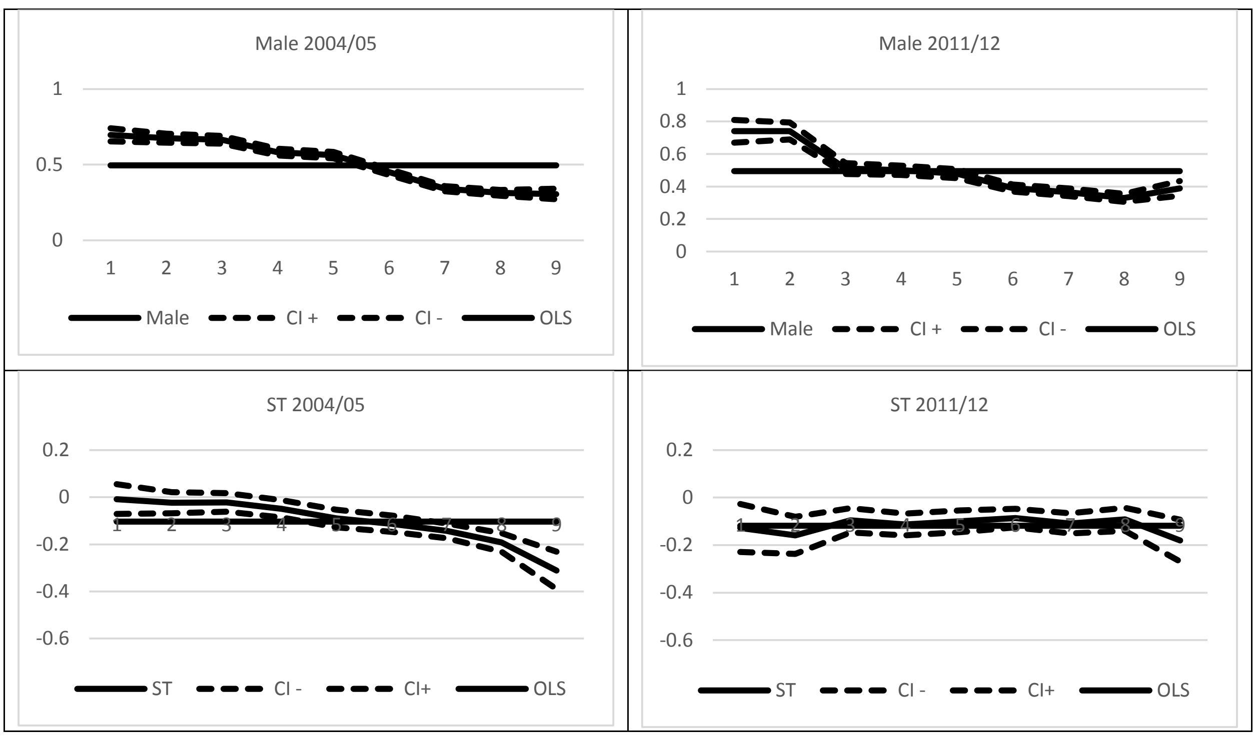
[20] The detailed tables for the OLS and RIF regressions are available from the authors on request

The first row of Figure 7 shows that the OLS and the nine RIF coefficients for the male dummy are positive and significant, implying the presence of a gender earnings gap. It is interesting to note that the RIF male coefficients are decreasing across deciles: in 2011/12, the male coefficient value is 0.74 at the first decile, 0.48 at the median, and 0.39 at the ninth decile. This 'sticky floor' effect implies that while men earn more than what women do throughout the distribution, the effect is more pronounced at the bottom of the distribution. Deshpande et al. (2015) also find a sticky floor for 1999/2000 and 2009/10 among regular salaried workers in India. The decreasing RIF coefficients also mean that—assuming that returns to observed characteristics remain unchanged—having a greater proportion of men would reduce earnings inequality among wage earners. This is unambiguously true for 2004/05 (as coefficients decline monotonically across all percentiles), and it is true for the lower part of the 2011/12 distribution.[21]

The second through fourth rows of Figure 7 show the presence of caste earnings gaps, though we do not see such gaps in all parts of the distribution.[22] On average, SCs, STs and OBCs earn significantly less than 'Others' do. The RIF regression coefficients for these covariates indicate that there is an earnings penalty for SC and OBC at the top deciles but not at the bottom. Thus, earnings inequality among wage earners would decrease if the shares of these caste categories were to increase.

The fifth row of Figure 7 indicates that there are positive returns to being married at the upper deciles. Thus, if the proportion of married were to increase earnings inequality among wage earners would increase.

Figure 7: RIF Regression Coefficients for selected Personal Characteristics, 2004/05 and 20011/12



[21] Note that RIF regressions assume that there are no general equilibrium effects.

[22] The 'Others' group includes, but is not confined to, the Hindu upper castes. The EUS data do not allow us to isolate Hindu upper castes and therefore this four-way division understates the gaps between the Hindu upper castes and the most marginalized ST and SC groups (Deshpande 2011).

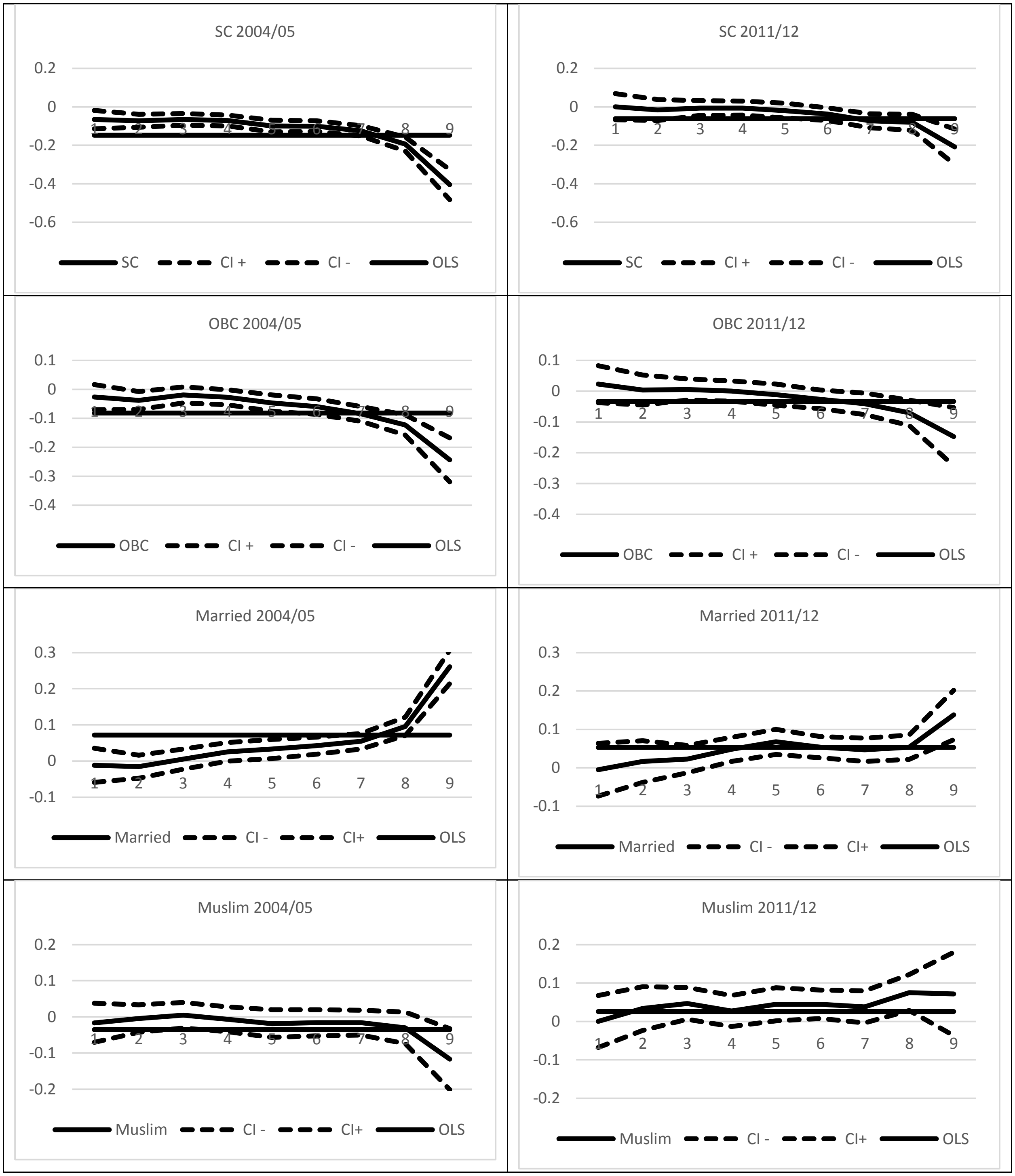
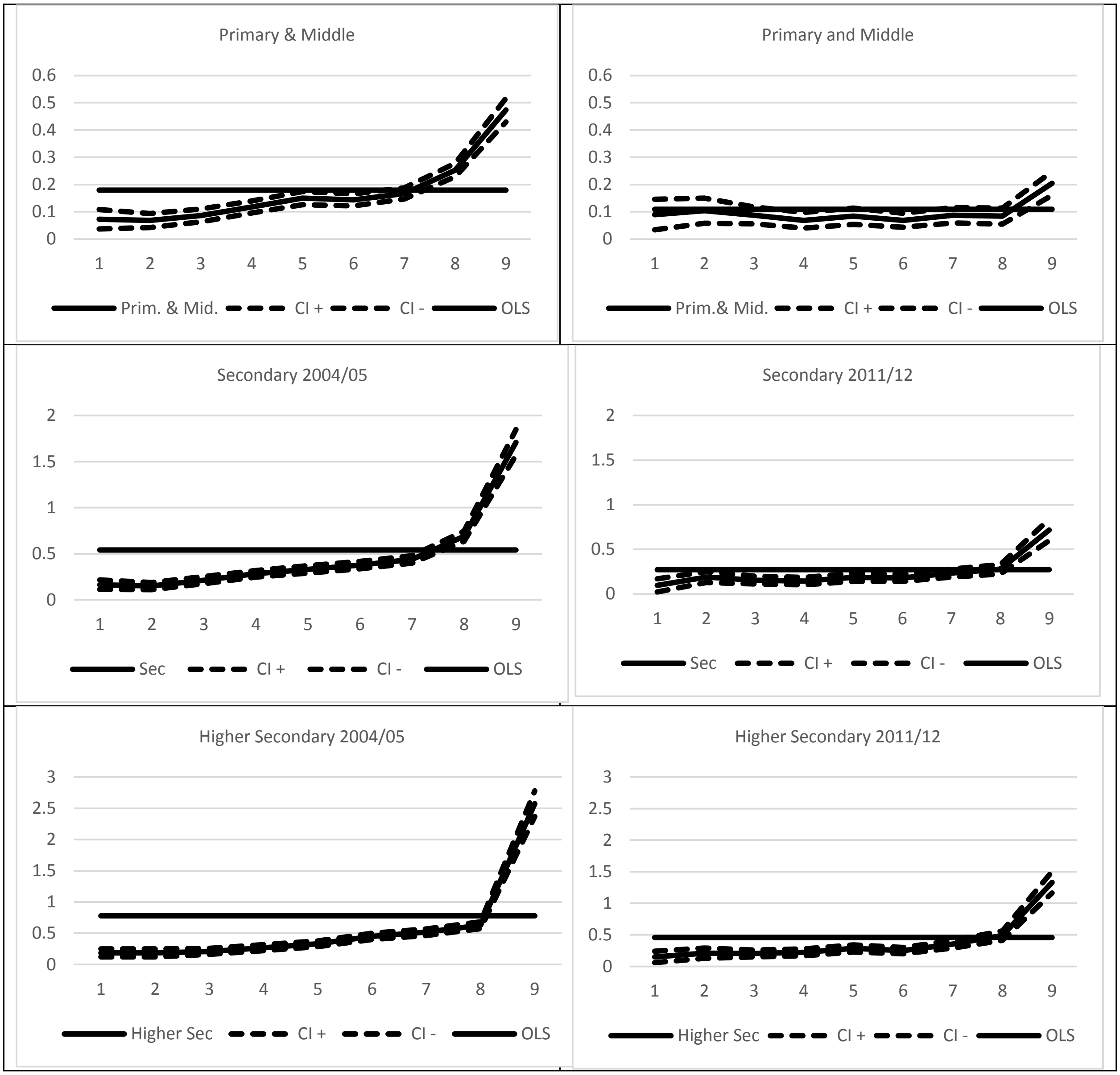
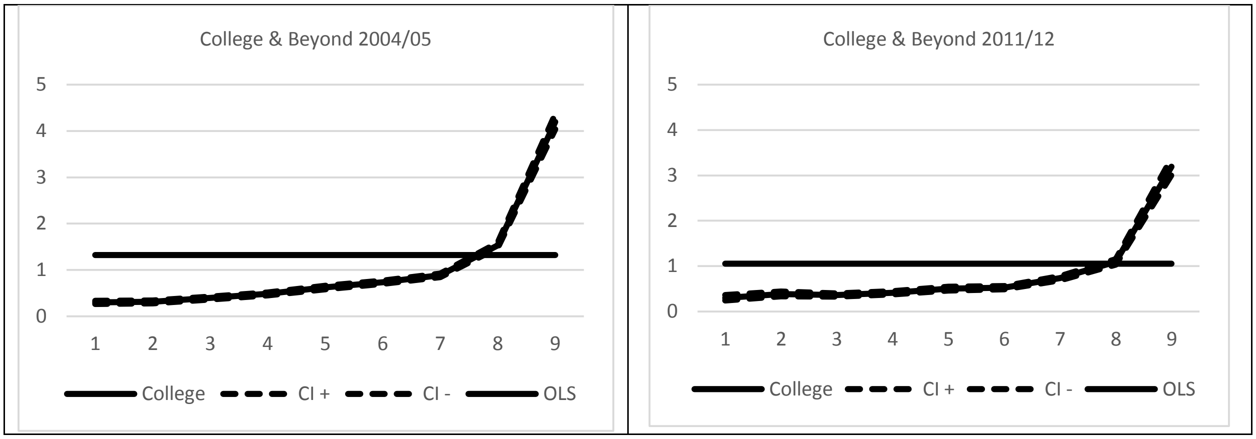


Figure 8 examines RIF coefficients for various education categories vis-à-vis the illiterates. First, there is clear evidence of positive returns to education. Perhaps the most striking feature is that, in both years, for each education category, there is a monotonic increase in returns as we move up the earnings distribution, with a sharp increase at the ninth decile. For instance, in 2011/12, the coefficient of 'college and beyond' is 0.43 at the first decile, 0.54 at the median and 3.37 at the ninth decile. Thus, educating the illiterate population increases earnings dispersion. Education increases both the level and the dispersion of earnings. This finding for rural India is similar to the evidence presented in Azam 2012a for regular salaried workers in urban India. Using conditional quantile regressions on EUS data for 1983, 1993/94 and 2004/05, he finds that returns to secondary and tertiary education have increased over time and are larger at higher quantiles.

Figure 8 also reveals how the impact of education on earnings dispersion has changed over time. The profile of RIF coefficients across deciles is flatter in 2011/12 than what it was in 2004/05 revealing that the inequality enhancing effect of education has weakened over the period.

Figure 8: RIF Regression Coefficients for Education Categories, 2004/05 and 2011/12





4.4 RIF Decomposition Results

Next we turn to RIF decompositions to analyse the changes in the real earnings distribution. We carry out the aggregate decompositions and the detailed decomposition of the composition effect.

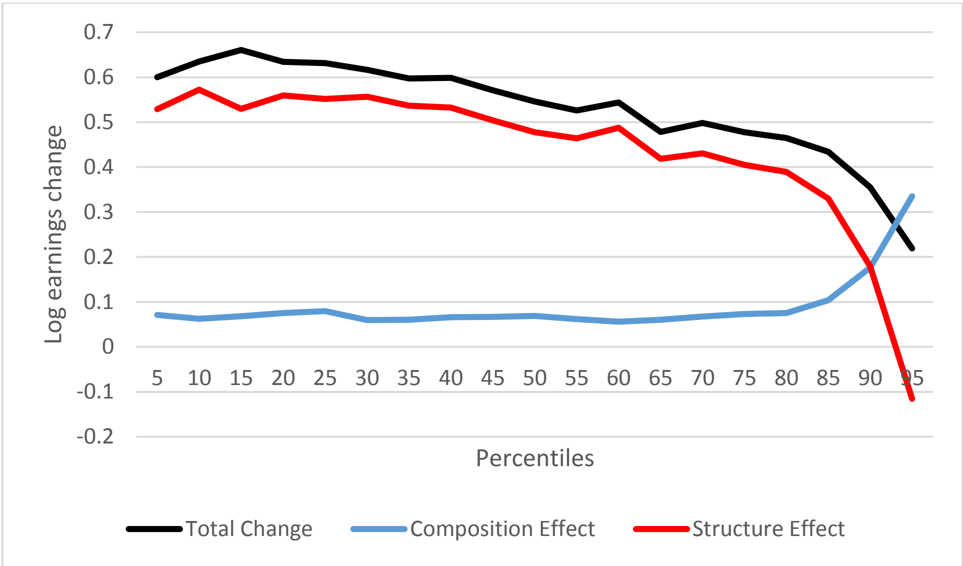
4.4.1 The Aggregate Decomposition of Change in Earnings

The RIF method is used to decompose the change in the (log) earnings distribution at different quantiles into the composition effect (the contribution of change in worker characteristics) and the structure effect (the contribution of change in returns to those characteristics). Figure 9 shows the results of the aggregate decomposition. We consider the decomposition based on the counterfactual that relies on the characteristics distribution of 2004/05 and returns of 2011/12.[23] For each vigintile, the total difference in log real earnings over the period is plotted in black. The downward slope of the total difference graph once again shows that the lower quantiles experienced a larger percentage increase in earnings than the higher quantiles.

The total difference is decomposed into the structure (red) and the composition effect (blue). Both components made significant contributions to the overall increase in log real earnings over the seven-year period. The only exception to this is at the nineteenth vigintile (95th percentile), where the structure effect is negative. Thus, the contribution of the structure effect to the overall increase in log earnings was positive and much larger than the composition effect at all but the top two vigintiles.[24]

[23] The results based on the counterfactual relying on the characteristics distribution of 2011/12 and returns of 2004/05 are similar and are available on request.
[24] We also implemented the aggregate decomposition using the Melly's refinement (Melly 2006) of the Machado-Mata Decomposition (Machado and Mata 2005) and found similar results.

Figure 9: The RIF Aggregate Decomposition



An important conclusion from the decomposition is that most of the decline in inequality occurred because the returns to personal characteristics improved a lot more for the lower percentiles than for higher ones. In fact, it is clear that while changing characteristics did lead to an improvement in real earnings throughout the distribution, it had an inequality increasing effect. The composition effect increased sharply after the eighth decile, implying that had 'returns to characteristics' been held constant over the period, earnings inequality would have risen.

Table 6 confirms this by decomposing several measures of inequality. The first column shows the difference between the log of real weekly earnings at the 90th and the 10th percentiles. Similarly, the second and the third column present the 50-10 and 90-50 differences, and the final column gives the Gini values for real weekly earnings. The difference between the values of these inequality measures over the period is the total change to be decomposed. Table 6 confirms that the structure effect had an inequality decreasing effect, while the composition effect had an inequality increasing effect seen mainly at the top of the wage distribution (as measured by the effect on the 90-50 measure), and is virtually absent at the bottom of the distribution (as measure by the effect on the 50-10 measure). In fact, had labor market characteristics remained the same in 2011/12 as they were in 2004/05, earnings inequality as measured by the Gini coefficient would have dropped from 0.461 to 0.366 (instead of the observed Gini of 0.395 in 2011/12). In summary, the decomposition of all inequality measures reveals that the decline in inequality came exclusively from the structure effect.

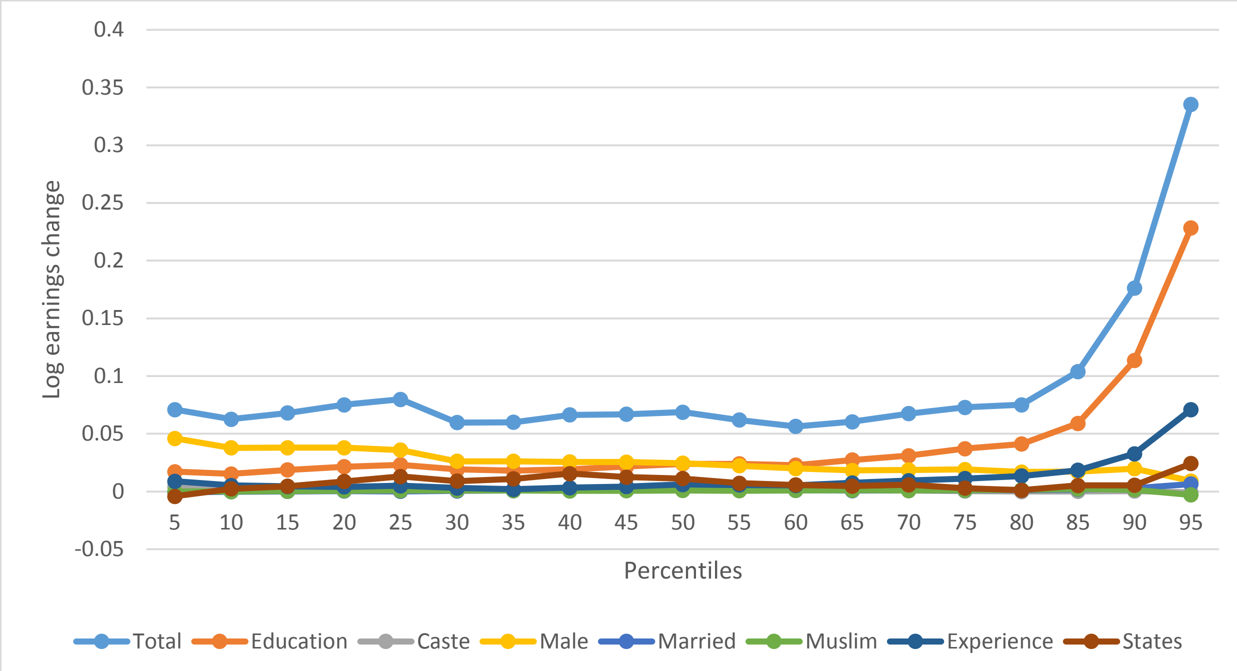
Table 6: Decomposition of Changes in Inequality Measures from 2004/05 to 2011/12				
	90-10	**50-10**	**90-50**	**Gini**
Value in 2004/05	1.855	0.881	0.974	0.461
Value in 2011/12	1.575	0.792	0.783	0.395
Total Change	-0.280	-0.089	-0.191	-0.066
Aggregate Decomposition of Total Change				
Structure Effect	-0.393	-0.095	-0.298	-0.095
Composition Effect	0.113	0.006	0.107	0.029
Detailed Decomposition of the Composition Effect				
Education	0.098	0.009	0.090	0.024
Experience	0.027	0.001	0.027	0.008
Male	-0.018	-0.013	-0.005	-0.006
Married	0.003	0.001	0.001	0.001
States	0.003	0.009	-0.006	0.002
Muslim	0.001	0.001	0.001	0.000
Caste	-0.002	-0.001	0.000	-0.001

4.4.2 Detailed Decomposition of the Composition Effect

In Figure 10 and the bottom panel of Table 6, we present the detailed decomposition of the composition effect to ascertain which set of covariates were important in driving the total composition effect. The inequality increasing effect was mainly driven by changes in the distribution of education, and to a lesser extent of experience. Changes in the distribution of marital status, state of residence, religion, and caste did not have a major impact on inequality. A greater proportion of men, however, contributed to the decline in inequality.

Although the detailed decomposition of the structure effect would have been instructive, it has an important limitation. The choice of the omitted or reference group for various categorical variables (such as caste and education) can influence the contribution of each covariate in the structure effect. Since the choice of the reference categories is arbitrary, results of the detailed decomposition can vary widely. For an excellent discussion of this (and other) limitations of decomposition techniques, see Fortin et al. 2011.

Figure 10: Detailed Decomposition of the Composition Effect



We conjecture that some part of the structure effect in explaining both the overall increase in rural earnings and the decline in earnings inequality, may be attributed to the implementation of the NREGA. The NREGA, first introduced in 2006 in the poorest districts of the country, was the flagship anti-poverty program of the erstwhile UPA government. By 2009, it was rolled out to remaining districts. Other studies (Azam 2012b; Berg et al. 2015; Imbert and Papp 2015) have shown that the scheme did result in a rise in rural wage rates for casual labor.

5 Conclusions

Using nationally representative data from the Employment Unemployment Survey we examined the changes in employment and in real weekly earnings from paid work for rural India from 2004/05 to 2011/12. First we summarize our findings and then place them in wider context.

The overall employment rate remained stable over this seven-year period and stood at 51 percent of the working age population. However, in spite of an already low base, the employment share among women declined from 24.6 percent to 21.1 percent.[25] The bulk of the paper focused on wage earners, who constituted 28 percent of the rural working age population in 2011/12. Real earnings for these workers increased at all percentiles. Expressed in rupee terms, the increase was larger at higher percentiles, while in percentage terms it was larger at lower percentiles. Thus, while earnings differences across percentiles widened, earnings inequality decreased. The decline in inequality was reflected in the drop in the Gini of real weekly earnings: It fell from 0.461 to 0.395. An accounting exercise revealed that about half the

[25] India has much lower female labour force participation rates even when compared to other developing countries. According to the World Bank, in 2013, overall female labour force participation rates for Brazil, China, India and South Africa were 59, 64, 27 and 45 percent, respectively. Data accessed from http://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS/countries/1W?display=default Last accessed on December 11, 2015.

decrease in inequality over time (measured by the decrease in variance of log real weekly earnings) was due to decrease in inequality of wage rates. The rest was due to decrease in inequality of days worked, and a weakening of the positive relationship between wage rates and days worked.

Unconditional Quantile Regressions revealed that an increase in the proportion of married and educated individuals contributes to an increase in earnings inequality. On the other hand, an increase in the proportion of male workers, and workers from Scheduled Castes and Other Backward Communities, contributes to a decrease in earnings inequality. The aggregate RIF decomposition showed that, except at the top end of the earnings distribution, the increase in real earnings was largely due to the structure effect i.e. an improvement in the returns to productive characteristics. The composition effect, namely, a change in the distribution of productive characteristics in the population, also contributed to an increase in real earnings, but compared to the structure effect it played a smaller role. At the top of the distribution, the structure effect was negative, although real earnings improved even at the top due to a positive composition effect. A decomposition of various inequality measures showed that the decrease in inequality over time was purely due to the structure effect: The returns to productive characteristics improved more for people at the lower percentiles. Had it not been for the structure effect inequality would have increased on account of the composition effect, especially the effect from an increased number of highly educated individuals among wage earners.

As mentioned above, we found that real earnings for wage earners in rural India increased at all percentiles. Using consumption expenditure data, other studies (Kotwal et al. 2011, for all-India during 1983-2004/05; Subramanian and Jayaraj 2015, for rural and urban separately during 2004/05-2009/10) have also documented an improvement in all parts of the distribution. Together, this provides evidence that economic growth in the post-reform period (after the early 1990s) has been accompanied by a reduction in poverty. However, Kotwal and Chaudhuri (2013) remark that in the post-reform period the decline in poverty has not been commensurate with the high per capita GDP growth.[26] According to official estimates, between 2009/10 and 2011/12, the share of persons below the poverty line[27] decreased from 39.6 to 30.9 percent in rural India, and from 35.1 to 26.4 percent in urban India. These figures still represent very large numbers of people living below a minimum acceptable standard of living and therefore leave no room for complacency.[28]

[26] One of the channels that they highlight in explaining the slow reduction in poverty is that growth in India is mainly driven by services sector, and the derived demand from services for goods produced in the informal sector (family owned farms and shops, trades people, business establishments employing less than 10 and typically 1 or 2 persons) is low.

[27] The poverty line is based on a normatively determined minimum level of consumption expenditure per person considered as adequate for nourishment, clothing, house rent, conveyance and education, and a behaviourally determined level of other non-food expenses (Planning Commission 2014). The poverty line in 2011/12 was 972 rupees per capita per month in rural, and 1407 in urban.

[28] In 2011/12 these poverty ratios implied about 261 and 102 million poor persons in rural and urban India, respectively. The severe deprivation in India is also reflected in the fact that 48 percent of children under five are stunted, 20 percent are wasted and 43 percent are underweight. The percentage underweight is almost twice as high as the average percentage of underweight children in sub-Saharan African countries (NFHS 2009).

According to World Bank estimates, India has the lowest Gini value among the BRICS countries.[29] However, this masks the extreme deprivations and inequities in access to health care, education and physical infrastructure such as safe water and sanitation (Drèze and Sen 2013). These inequities are reflected in the fact that India suffers from the simultaneous occurrence of over nutrition and under nutrition.[30] Although, in this paper we have documented an increase in real earnings for rural India and a decrease in rural earnings inequality, other studies on India have found that poverty hasn't declined as fast as it should, that there has been an increase in urban inequality, and rural urban divide has widened (Motiram and Vakulabharanam 2012).

[29] The Gini values are as follows: Brazil-53.9 (2009); Russia-39.7 (2009); India-33.9 (2009); China-42.1 (2010) and South Africa-63 (2008). These are available at http://data.worldbank.org/indicator/SI.POV.GINI. Last accessed on November 18, 2015.

[30] 36 percent of women and 34 percent of men are undernourished, while 13 percent of women and 9 percent of men are overweight or obese (NFHS 2009).

6 References

- Aguiar, M. and M. Bils. 2015. “Has Consumption Inequality Mirrored Income Inequality?” American Economic Review; 105(9): 2725-56.
- Azam, M. 2012a. “Changes in Wage Structure in Urban India, 1983–2004: A Quantile Regression Decomposition.” World Development; 40(6): 1135-1150.
- Azam, M. 2012b. “The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment” IZA Discussion Papers; IZA DP No. 6548.
- Banerjee, A. and T. Piketty. 2005. “Top Indian Incomes, 1922-2000.” World Bank Economic Review; 19(1): 1-20.
- Berg, E., S. Bhattacharyya, D. Rajasekhar, and R. Manjula. 2015. “Can Public Works Increase Equilibrium Wages? Evidence from India’s National Rural Employment Guarantee.” http://www.erlendberg.info/agwages.pdf.
- Blinder, A. 1973. “Wage Discrimination: Reduced Form and Structural Estimates.” Journal of Human Resources; 8:436-455.
- Cain, J. S., R. Hasan, R. Magsombol, and A. Tandon. 2010. “Accounting for Inequality in India: Evidence from Household Expenditures.” World Development; 38(3): 282-297.
- Deshpande, A. 2011. “The Grammar of Caste: Economic Discrimination in Contemporary India.” Oxford University Press, New Delhi.
- Deshpande, A., D. Goel, and S. Khanna. 2015. “Bad Karma or Discrimination? Male-Female Wage Gaps among Salaried Workers in India.” IZA Discussion Papers, IZA DP No. 9485.
- Drèze, J. and A. Sen. 2013. “An Uncertain Glory: India and its Contradictions.” Princeton University Press.
- Dutta, P. V. 2005. “Accounting for Wage Inequality in India.” Poverty Research Unit at Sussex, PRUS Working Paper No. 29.
- Firpo, S., N. M. Fortin, and T. Lemieux. 2009. “Unconditional Quantile Regressions.” Econometrica; 77: 953–973. doi: 10.3982/ECTA6822
- Fortin, N. M., T. Lemieux, and S. Firpo. 2011. “Decomposition Methods in Economics.” Handbook of Labor Economics (edited by Ashenfelter and Card.); Volume 4A, Chapter 1.
- GOI. 2015. “Economic Survey 2014-15.” Government of India, Ministry of Finance.
- Goldberg, P. K. and N. Pavcnik. 2007. “Distributional Effects of Globalization in Developing Countries.” Journal of Economic Literature; Vol XLV: 39-82.
- Hasan, R. and A. S. Gupta. 2014. “Growth, structural change, and poverty reduction: Evidence from India.” Ideas for India, http://www.ideasforindia.in/article.aspx?article_id=266.
- Imbert, C. and J. Papp. 2015. “Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee.” American Economic Journal: Applied Economics; 7(2):233-263.
- Kijima, Y. 2006. “Why did wage inequality increase? Evidence from urban India 1983–99.” Journal of Development Economics; 81: 97-117.
- Koenker, R. and G. Bassett. 1978. “Regression Quantiles.” Econometrica; 46: 33- 50.
- Kotwal, A. and A. R. Chaudhuri. 2013. “Why is Poverty Declining so Slowly in India?” Ideas for India, http://www.ideasforindia.in/article.aspx?article_id=110
- Kotwal A., B. Ramaswami, and W. Wadhwa. 2011. “Economic Liberalization and Indian Economic Growth: What's the Evidence?” Journal of Economic Literature; 49(4): 1152-1199.

- Krueger, D. and F. Perri. 2006. “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory.” Review of Economic Studies; 73(1): 163–93.
- Machado, J. F., and J. Mata. 2005. “Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression.” Journal of Applied Econometrics; 20: 445–465.
- Melly, B. 2006. “Estimation of Counterfactual Distributions using Quantile Regression.” University of St. Gallen, Discussion Paper.
- Motiram, S. and V. Vakulabharanam. 2012. “Indian Inequality: Patterns and Changes, 1993 – 2010.” India Development Report, Oxford University Press, Vol 7: 224-232.
- NFHS. 2009. “Nutrition in India. National Family Health Survey (NFHS-3) India 2005-06.” Ministry of Health and Family Welfare, Government of India.
- Oaxaca, R. L. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” International Economic Review; 14: 693-709.
- Planning Commission. 2014. “Report of the Expert Group to Review the Methodology for Measurement of Poverty.” Planning Commission, Government of India.
- RBI. 2015. “Handbook of Statistics on the Indian Economy 2014-15.” Reserve Bank of India.
- Sen, A. and Himanshu. 2004. “Poverty and Inequality in India: II: Widening Disparities during the 1990s.” Economic and Political Weekly; 39(39): 4361-4375.
- Subramanian, S. and D. Jayaraj. 2015. “Growth and Inequality in the Distribution of India’s Consumption Expenditure.” WIDER Working Paper 2015/025.