SUDIPTA GHOSH*
University of British Columbia

VIKTORIA HNATKOVSKA†
University of British Columbia

AMARTYA LAHIRI‡
University of British Columbia

Rural–Urban Disparities in India in the Time of Growth§

ABSTRACT The period since the 1990s has seen an aggregate growth takeoff in India during which the rural agricultural sector has gradually ceded space in both employment and output share to non-agricultural sectors. How have agriculture-dependent rural workers and households fared during this episode relative to their urban counterparts? Using household-level survey data, we find that rural–urban education gaps have declined significantly between 1983 and 2012. Moreover, occupation choices in the two sectors have become more aligned with an expansion of non-farm occupations in rural India. Consumption gaps between rural and urban households also declined between 1983 and 2005 for the bottom 45th percentile of the distribution. The period since 2005 has, however, witnessed a rise in consumption gaps. Using state-level data, we show that per capita income levels and growth rates are positive correlates of rural–urban consumption gaps, while the education gap is a negative correlate.

Keywords: *Rural–Urban Disparity, Consumption Gaps, Education*

JEL Classification: *E2, O1, R2*

1. Introduction

Periods, when macroeconomic growth takes off in countries, are often accompanied by an underlying microeconomic churn. Such periods are typically characterized by a shrinking of the rural agricultural sector and

* *gsudipta01@gmail.com*
† *hnatkovs@gmail.com*
‡ *amartyalahiri@gmail.com*
§ The authors would like to thank, without implicating, the discussants Rohini Somanathan and Jeff Hammer, the editor Karthik Muralidharan, and participants at the IPF 2019 Conference at NCAER, New Delhi, for their comments and suggestions.

an expansion of the urban non-agricultural sectors. How do these episodes impact the fortunes of rural and urban households? The question is often also at the heart of discussions regarding the evolution of inequality in developing economies. Not surprisingly, in a recent cross-country study on a sample of 65 countries, Young (2013) finds that around 40 percent of the average inequality in consumption is due to urban–rural gaps.

In this paper, we examine the fortunes of rural and urban workers in India between 1983 and 2012. We provide a comprehensive empirical documentation of the trends in rural and urban disparities in India since 1983 in education, occupation distributions, and consumption. In addition, the paper examines state-level evidence to both corroborate the aggregate evidence and uncover potential driving forces for the evolving pattern of rural–urban disparity.

Using seven thick rounds of the National Sample Survey (NSS) of households in India between 1983 and 2012, we analyze the evolution over time of education attainment, occupation choices, and consumption levels of rural and urban workers. Our analysis yields several results.

First, while the educational attainment rates of both rural and urban individuals have been rising, the gap between them has been shrinking dramatically over time in terms of both the years of schooling and the relative distribution of workers in different education categories. Second, the shares of non-farm jobs (both white- and blue-collar) have expanded dramatically in rural areas, leading to a reduction in the dissimilarity between the occupation choices of rural and urban households.

Third, mean per capita consumption differences between urban and rural households have narrowed up to the bottom 15 percentile of the consumption distribution. For higher percentiles, however, the consumption gaps have widened between 1983 and 2012. We find that this widening of rural–urban consumption gaps is a relatively recent phenomenon that has set in after 2004–05 with a sharp widening of the gap post 2009–10.

Fourth, we examine the disaggregated state-level data to find the key determinants of the evolving rural–urban gaps in consumption using the panel structure of the state-level data for identification. We find that the rural–urban consumption gaps, both in mean and quantiles, are significantly greater in states with higher levels and growth rates of per capita Net State Domestic Product (NSDP). In addition, states with lower rural–urban education gaps tend to have significantly lower consumption gaps.

One curious feature of our findings is that consumption gaps between rural and urban households reversed their two-decade-long narrowing trend and began widening after 2004–05. This is particularly puzzling since

India introduced one of the world's largest public works and employment programs in the form of the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) in 2006. MGNREGA guaranteed 100 days of work to all rural workers. This should, in theory, have provided rural households with a consumption boost post-2006.

Our interest in rural–urban gaps is probably closest in spirit to the work of Young (2013), who examined rural–urban consumption expenditure gaps in 65 countries. He found that only a small fraction of the rural–urban inequality can be accounted for by individual characteristics, like education differences. He attributed the remaining gaps to the competitive sorting of workers to rural and urban areas based on their unobserved skills.

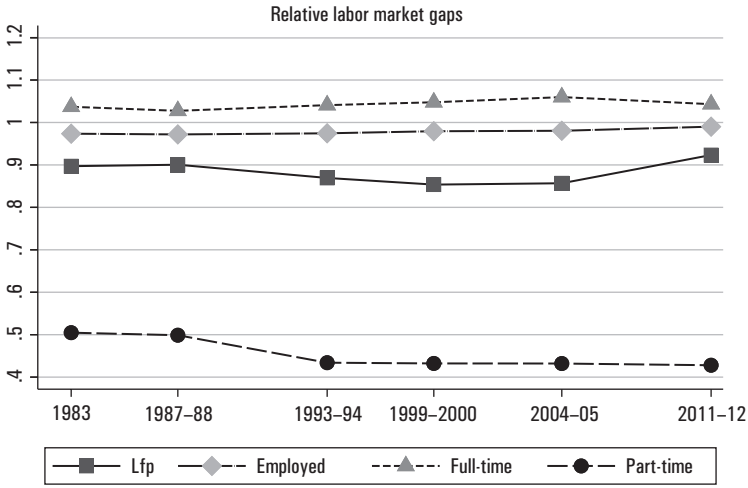
Our work is also related to empirical literature studying rural–urban gaps in different countries (see, for instance, Nguyen et al. [2007] for Vietnam; Qu and Zhao [2008] and Wu and Perloff [2005] for China). These papers generally employ household survey data and relate changes in rural–urban inequality to individual and household characteristics. Our study is the first to conduct a similar analysis for India for multiple years, as well as to extend the analysis to the state level.

The rest of the paper is organized as follows: Section 2 presents the data and some motivating statistics. Section 3 delineates the main results on the evolution of the rural–urban gaps as well as the analysis of the extent to which these changes occurred due to changes in the individual characteristics of workers. Section 4 presents the results of urban–rural gaps across states in India. Section 5 investigates the possible factors driving the dynamics of rural–urban consumption gaps. Section 6 contains concluding thoughts.

2. Data

Our data comes from successive rounds of the NSS of households in India for employment and consumption. The survey rounds that we include in the study are 1983 (Round 38), 1987–88 (Round 43), 1993–94 (Round 50), 1999–00 (Round 55), 2004–05 (Round 61), 2009–10 (Round 66), and 2011–12 (Round 68). Since our focus is on determining the trends in outcomes of the workforce in rural and urban India, we restrict the sample to individuals in the working age group of 16–65 years, who are not enrolled in any educational institution, and for whom we have educational information as well as information on their weekly activity status (unemployed, regular salaried worker, casual worker, or self-employed). This restriction excludes individuals who report home production as their primary weekly

FIGURE 1. Ratio of Urban to Rural Labor Force Participation and Employment Rates



Source: Authors' calculations from NSS data (see text for details).
Notes: "Lfp" = ratio of urban to rural labor force participation rates; "employed"= ratio of urban to rural employment rates; "full-time" = ratio of urban to rural full-time employment rate, and "part-time" = ratio of urban to rural part-time employment rates.

activity. We further restrict the sample to individuals who belong to male-led households.[1] These restrictions leave us with a sample size that varies between 159,000 and 221,000 individuals per survey round.

Figure 1 plots the urban to rural ratios in labor force participation rates, overall employment rates, as well as full-time and part-time employment rates. As can be seen from the figure, there was some increase in the relative rural part-time work incidence between 1987 and 2012. Apart from this, all other trends were basically flat. Details on our data are provided in Appendix A.1.

We summarize demographic characteristics in our sample across the rounds in Table 1. The table breaks down the overall patterns by individuals and households and by rural and urban locations. Clearly, the sample is overwhelmingly rural with about 75 percent of households, on average, being residents in rural areas. Rural residents are less likely to be male, slightly more likely to be married, and belong to marginally larger households than their urban counterparts. Lastly, rural areas have more members

1. Male-led households are the norm in India.

of backward castes as measured by the proportion of Scheduled Castes (SCs) and Scheduled Tribes (STs).

The panel labeled "Urban–Rural Difference" in Table 1 reports the differences in individual and household characteristics between urban and rural areas for all our survey rounds. Clearly, the share of the rural labor force has declined over time. The average age of individuals in both urban and rural areas has increased over time. Families have also become smaller in both locations, but the decline has been more rapid in urban areas, leading to a

TABLE 1. Summary Statistics for NSS Data from 1983 to 2011–12 for Individuals in the Working Age Group 16–65 used in this Paper

	(a) Individuals			*(b) Households*		
		Proportion		*Proportion*		*Household*
Urban	*Age*	*Male*	*Married*	*Urban*	*SC/ST*	*Size*
1983	35.57	0.83	0.73	0.25	0.16	6.36
	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
1987–88	35.70	0.83	0.73	0.24	0.16	6.15
	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
1993–94	36.16	0.81	0.75	0.26	0.16	5.70
	(0.06)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
1999–00	36.52	0.82	0.74	0.27	0.19	5.79
	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
2004–05	36.69	0.81	0.74	0.27	0.18	5.57
	(0.08)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
2011–12	37.94	0.82	0.77	0.31	0.18	5.14
	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
Rural	*Age*	*Male*	*Married*	*Rural*	*SC/ST*	*HH Size*
1983	35.86	0.71	0.79	0.75	0.29	6.52
	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
1987–88	35.79	0.72	0.79	0.76	0.30	6.42
	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
1993–94	35.99	0.69	0.80	0.74	0.31	6.08
	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
1999–00	36.28	0.69	0.80	0.73	0.33	6.23
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
2004–05	36.71	0.68	0.80	0.73	0.32	6.03
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
2011–12	38.21	0.74	0.81	0.69	0.32	5.52
	(0.08)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)

(Table 1 Contd.)

(Table 1 Contd.)

Urban–Rural Difference	*Age*	*Male*	*Married*	*Rural*	*SC/ST*	*HH Size*
1983	–0.29***	0.12***	–0.06***	–0.50***	–0.13***	–0.16***
	(0.08)	(0.00)	(0.00)	(0.03)	(0.00)	(0.03)
1987–88	–0.09	0.11***	–0.06***	–0 52***	–0.14***	–0.27***
	(0.08)	(0.00)	(0.00)	(0.03)	(0.00)	(0.03)
1993–94	0.17	0.12***	–0.05***	–0.48***	–0.15***	–0.38***
	(0.08)	(0.00)	(0.00)	(0.03)	(0.00)	(0.02)
1999–00	0.27*	0.13***	–0.06***	–0.46***	–0.14***	–0.44***
	(0.08)	(0.00)	(0.00)	(0.03)	(0.00)	(0.03)
2004–05	–0.02***	0.13***	–0.06***	–0.46***	–0.14***	–0.46***
	(0.10)	(0.00)	(0.00)	(0.04)	(0.00)	(0.03)
2011–12	–0.27***	0.08***	–0.04***	–0.38***	–0.14***	–0.38***
	(0.12)	(0.00)	(0.00)	(0.04)	(0.00)	(0.03)

Source: Authors' calculations. See text for details of data being used of 16–65-year-old working age individuals.

Notes: Columns (a) give statistics at the individual level, while Columns (b) give statistics at the household level. The section labeled "urban–rural difference" reports the difference in characteristics between urban individuals/households and rural individuals/households. Standard errors are reported in parentheses: *p value $\leq$ 0.10, **p value $\leq$ 0.05, ***p value $\leq$ 0.01.

large differential in this characteristic between the two areas. The shares of male workers remained stable in urban areas but showed a sharp increase in rural areas in the last survey round.

Education in the NSS data is presented as a category variable with the survey listing the highest education-attainment level in terms of categories such as primary and middle. In order to ease the presentation, we proceed in two ways. First, we construct a variable for the years of education. We do so by assigning the years of education to each category based on a simple mapping: not literate = 0 years; literate but below primary = 2 years; primary = 5 years; middle = 8 years; secondary and higher secondary = 10 years; graduate = 15 years, and post-graduate = 17 years. Diplomas are treated similarly, depending on the specifics of the attainment level.[2] Second, we use the reported education categories but aggregate them into five broad groups as follows: 1 for illiterate, 2 for some education but below primary school, 3 for primary school, 4 for middle, and 5 for secondary and above. The results from the two approaches are similar. While we use the second

2. We are forced to combine secondary and higher secondary into a combined group of 10 years because the higher-secondary classification is missing in the 38th and 43rd Rounds. The only way to retain comparability across rounds, then, is to combine the two categories.

method for our econometric specifications as these are the actual reported data (as opposed to the years series that was constructed by us), we also show results from the first approach below.

For our analysis of occupation choices, we aggregate the reported three-digit occupation categories in the survey into three broad occupation categories: *white-collar* occupations such as administrators, executives, managers, professionals, technical, and clerical workers; *blue-collar* occupations, like sales workers, service workers, and production workers, and *agrarian* occupations including farmers, fishermen, loggers, hunters.

We obtain consumption as the monthly per capita expenditure (MPCE) of rural and urban households. We convert nominal MPCE into real terms using state-level poverty lines that differ for the rural and urban sectors. We express all consumption figures in the 1983 rural Maharashtra poverty lines.[3]

3. Empirical Findings

How did urban and rural workers fare during our sample period? We characterize differences in education attainments, occupations, labor income, and consumption of the rural and urban workforce to answer this question.

3.1. Education

Table 2 shows the average number of years of education of the urban and rural workforce across the six rounds in our sample. The two features that emerge from the table are: (a) education attainment rates as measured by the number of years of education were rising in both urban and rural sectors during this period, and (b) the rural–urban education gap shrank monotonically over this period.

The average number of years of education for the urban worker was 170 percent higher than that for the typical rural worker in 1983 (5.56 years to 2.06 years). This advantage declined to 78 percent by 2011–12 (8.34 years

3. In 2004–05, the Planning Commission of India changed the methodology for estimation of poverty lines. Among other changes, it switched from anchoring the poverty lines to a calorie intake norm towards consumer expenditures more generally. This led to a change in the consumption basket underlying poverty lines calculations. To retain comparability across rounds, we convert the 2011–12 poverty lines obtained from the Planning Commission under the new methodology to the old basket using the 2004–05 adjustment factor. That factor was obtained from the poverty lines under the old and new methodologies available for the 2004–05 survey year. As a test, we used the same adjustment factor to obtain the implied "old" poverty lines for the 1993–94 survey round, for which the two sets of poverty lines are also available from the Planning Commission. We find that the actual old poverty lines and the implied "old" poverty lines are very similar, giving us confidence that our adjustment is valid.

TABLE 2. Education Gap: Average Years of Schooling

	Average Years of Education			*Relative Education Gap*
NSS Rounds	*Overall*	*Rural*	*Urban*	*Ratio of Urban/Rural*
1983	2.90	2.06	5.56	2.70***
	(0.01)	(0.01)	(0.03)	(0.02)
1987–88	3.14	2.29	5.89	2.57***
	(0.01)	(0.01)	(0.03)	(0.02)
1993–94	3.71	2.81	6.63	2.36***
	(0.01)	(0.01)	(0.03)	(0.03)
1999–00	4.27	3.30	7.21	2.18***
	(0.02)	(0.02)	(0.03)	(0.02)
2004–05	4.66	3.75	7.49	2.00***
	(0.02)	(0.02)	(0.04)	(0.01)
2011–12	5.77	4.69	8.34	1.78***
	(0.02)	(0.03)	(0.04)	(0.01)

Source: Authors' calculations. See text for details.

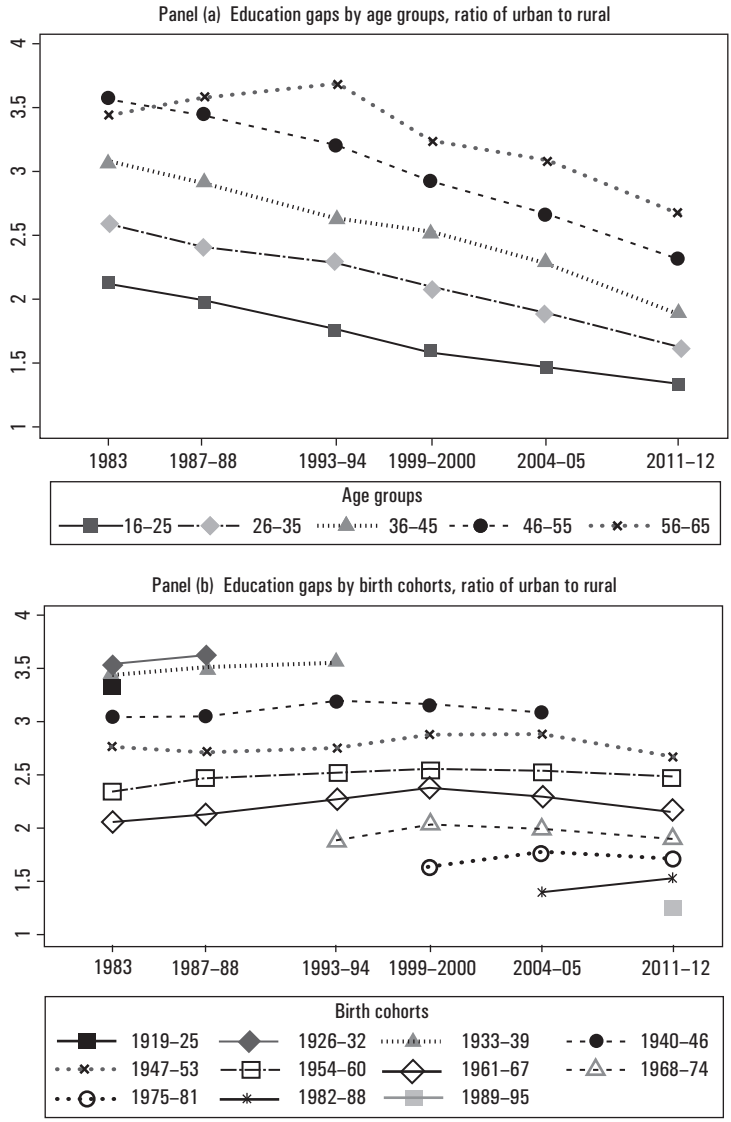
Notes: This table presents the average number of years of education for the overall sample and for urban and rural populations. The reported statistics are obtained for each NSS Round, which is shown in the first column. Standard errors are in parentheses. *p-value $\leq$ 0.10, **p-value $\leq$ 0.05, ***p-value $\leq$ 0.01.

to 4.69 years). To put these numbers in perspective, in 1983, the average urban worker had slightly more than primary education while the typical rural worker was literate but below primary level. By 2011–12, the average urban worker had about a middle school education, while the typical rural worker had almost attained primary education. While the overall numbers indicate the still dire state of literacy of the workforce in the country, the movements underneath do indicate improvements over time, with the rural workers improving faster.

Table 2, while revealing an improving trend for the average worker, nevertheless, masks potentially important underlying heterogeneity in education attainment by cohort, that is, variation by the age of the respondent. Panel (a) of Figure 2 shows the relative gap in the number of years of education between the typical urban and rural worker by age group. There are two key results to note here: (a) the gaps have been getting smaller over time for all age groups, and (b) the gaps are smaller for the younger age groups.

Is the education convergence taking place uniformly across all birth cohorts, or are the changes mainly being driven by aging effects? To disentangle the two, we compute relative education gaps for different birth cohorts for every survey year. These are plotted in Panel (b) of Figure 2. Clearly, almost all of the convergence in education attainments takes place through cross-cohort improvements, with the younger cohorts showing the smallest gaps. Aging effects are symmetric across all cohorts, except the eldest. Most strikingly, the

FIGURE 2. Education Gaps by (a) Age Groups and (b) Birth Cohorts: Ratio of Urban to Rural Average Years of Education



Source: Authors' calculations.

Notes: The panels in this figure show the ratio of the average number of years of education between the urban and rural workforces over time for different age groups and birth cohorts.

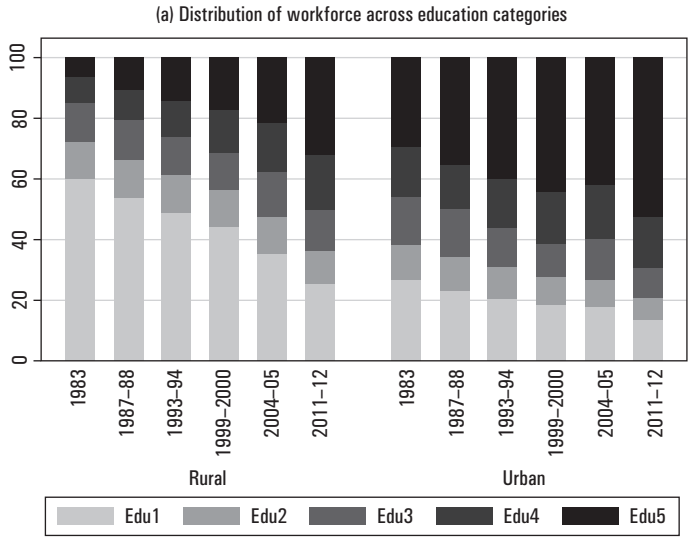
average gap in 2011–12 between urban and rural workers from the youngest birth cohort (born between 1989 and 1995) has almost disappeared, while the corresponding gap for those born between 1954 and 1960 stood at 150

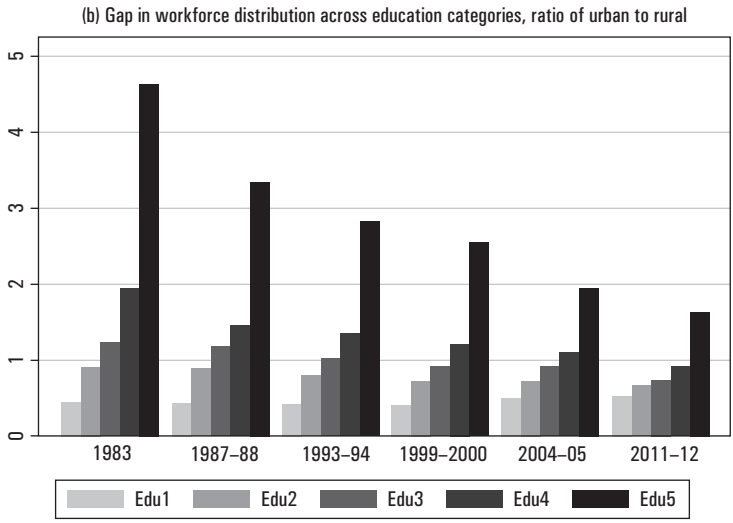
percent. Clearly, the declining rural–urban gaps are being driven by declining education gaps amongst the younger workers in the two sectors.

The time trends in the number of years of education potentially mask changes in the quality of education. In particular, they fail to reveal what kind of education is causing the rise in years: Is it people moving from middle school to secondary, or is it movement from illiteracy to some education? While both movements would add a similar number of years to the total, the impact on the quality of the workforce may be quite different. Further, we are also interested in determining whether the movements in urban and rural areas are being driven by very different movements in the category of education.

Panel (a) of Figure 3 shows the distribution of the urban and rural workforce by education category. Recall that education categories 1, 2, and 3 are "illiterate", "some education but below primary", and "primary education", respectively. Hence, in 1983, 55 percent of the urban labor force and over 85 percent of the rural labor force had primary or below education, reflecting the abysmal delivery of public services in education during the first 35 years of post-Independence India. By 2012, the primary and below category had come down to 30 percent for urban workers and 50 percent for rural workers. Simultaneously, the other notable trend during this period is the perceptible increase in the secondary and above category for workers in both sectors. For the urban sector, this category expanded from about 30 percent in 1983 to over 50 percent in 2012. Correspondingly, the share of secondary and

FIGURE 3. Education Distribution of Rural and Urban Samples and Ratio of Urban to Rural Distributions within Five Education Categories





Source: Authors' calculations.
Notes: The definition of our education categories is: Edu1 illiterate; Edu2 some education but below primary school; Edu3 primary school; Edu4 middle school; and Edu5 secondary and above. For more details see Section 2 of this paper. Panel (a) presents the distribution of the workforce across five education categories for different NSS rounds. Panel (b) presents relative gaps (ratio of urban to rural) in the distribution of urban relative to rural workers across the five education categories.

higher educated rural workers rose from just around 5 percent of the rural workforce in 1983 to about 30 percent in 2012. This, along with the decline in the proportion of rural illiterate workers from 60 percent to around 25 percent, represents the sharpest and most promising change since 1983.

Panel (b) of Figure 3 shows the changes in the relative education distributions of the urban and rural workforce. For each survey year, the figure shows the fraction of urban workers in each education category relative to the fraction of rural workers in that category. Thus, in 1983, urban workers were over-represented in the secondary and above category by a factor close to 5. Similarly, rural workers were over-represented in the education category 1 (illiterates) by a factor of 2. Clearly, the closer the height of the bars is to one, the more symmetric is the distribution of the two groups in that category, while the further away from one they are, the more skewed is the distribution. As the figure indicates, the biggest convergence in the education distribution between 1983 and 2012 was in categories 4 and 5 (middle, and secondary and above) where the bars shrank rapidly. The trends in the other three categories were more muted as compared to the convergence in categories 4 and 5.

While the visual impressions suggest convergence in education, are these trends statistically significant? We turn to this issue next by estimating

ordered multinomial probit regressions of education categories 1 to 5 on a constant and the rural dummy. The aim is to ascertain the significance of the difference between rural and urban areas in the probability of a worker belonging to each category as well as the significance of changes over time in these differences. Table 3 shows the results.

Panel (a) of Table 3 shows that the marginal effect of the rural dummy was significant for all rounds and all categories. The rural dummy significantly raised the probability of belonging to education categories 1 and 2 ("illiterate" and "some but below primary education", respectively) while

TABLE 3. Marginal Effects of a Rural Dummy in Ordered Probit Regressions for Education Categories

	Panel (a): Marginal Effects, Unconditional					
	1983	*1987–88*	*1993–94*	*1999–00*	*2004–05*	*2011–12*
Edu1	0.3489***	0.3391***	0.3215***	0.3026***	0.2757***	0.2242
	(0.0026)	(0.0022)	(0.0023)	(0.0025)	(0.0025)	(0.0031)
Edu2	–0.0021***	0.0051***	0.0151***	0.0231***	0.0319***	0.0421***
	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0006)	(0.0010)
Edu3	–0.0496***	–0.0393***	–0.0197***	–0.0037***	0.0072***	0.0245***
	(0.0006)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0008)
Edu4	–0.0889***	–0.0762***	–0.0656***	–0.0538***	–0.0480***	–0.0169***
	(0.0010)	(0.0008)	(0.0007)	(0.0007)	(0.0006)	(0.0007)
Edu5	–0.2082***	–0.2287***	–0.2514***	–0.2681***	–0.2668***	–0.2738***
	(0.0022)	(0.0020)	(0.0023)	(0.0028)	(0.0031)	(0.0040)
N	203,456	221,228	199,579	210,209	220,786	159,193

	Panel (b): Changes over Time			
	1983 to 1993–94	*1993 to 2004–05*	*2004 to 2011–12*	*1983 to 2011–12*
Edu1	–0.0274***	–0.0458***	–0.0515***	–0.1247***
	(0.0035)	(0.0034)	(0.0040)	(0.0040)
Edu2	0.0172***	0.0168***	0.0102***	0.0442***
	(0.0006)	(0.0007)	(0.0012)	(0.0011)
Edu3	0.0299***	0.0269***	0.0173***	0.0741***
	(0.0007)	(0.0006)	(0.0009)	(0.0010)
Edu4	0.0233***	0.0176***	0.0311***	0.0720***
	(0.0012)	(0.0009)	(0.0009)	(0.0012)
Edu5	–0.0432***	–0.0154***	–0.0070***	–0.0656***
	(0.0032)	(0.0039)	(0.0051)	(0.0046)

Source: Authors' calculations. See text for details.

Notes: Panel (a) of this table reports the marginal effects of the rural dummy in an ordered probit regression of Education Categories 1–5 on a constant and a rural dummy for each survey Round. Panel (b) of the table reports the change in the marginal effects over successive decades and the entire sample period. N refers to the number of observations. Standard errors are in parentheses. * p-value $\leq$ 0.10, ** p-value $\leq$ 0.05, *** p-value $\leq$ 0.01.

it significantly reduced the probability of belonging to categories 4–5. In category 3, the sign on the rural dummy had switched from negative to positive in 2004–05 and stayed that way in 2011–12.

Panel (b) of Table 3 shows that the changes over time in these marginal effects were also significant for all rounds and all categories. There are clearly significant convergent trends for education categories 1, 3, and 4. Category 1, where rural workers were over-represented in 1983, saw a declining marginal effect of the rural dummy. Categories 3 and 4 (primary and middle school, respectively), where rural workers were under-represented in 1983, saw a significant increase in the marginal effect of the rural status. Hence, the rural under-representation in these categories declined significantly. Categories 2 and 5 were, however, marked by a divergence in the distribution. Category 2, where rural workers were over-represented, saw an increase in the marginal effect of the rural dummy, while in category 5, where they were under-represented, the marginal effect of the rural dummy became even more negative. However, this divergence is not inconsistent with Figure 3. The figure shows trends in the relative gaps while the probit regressions show trends in the absolute gaps.

In summary, the overwhelming feature of the data on education attainment gaps suggests a strong and significant trend towards education convergence between the urban and rural workforce. This is evident in the comparison of the average number of years of education, the relative gaps by education category as well as the absolute gaps between the groups in most categories.

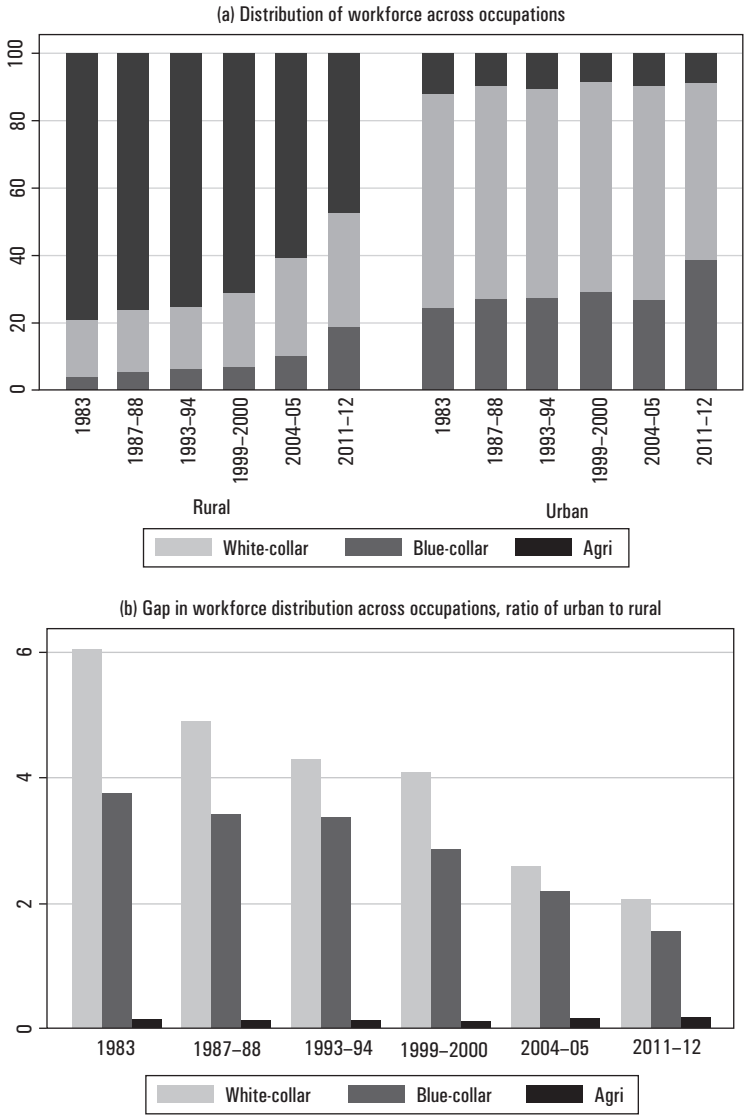
The convergence results on education come with the obvious caveat that education attainment rates, even when the education categories are examined, may not necessarily reflect the quality of education. Thus, a middle school education from rural areas might not imply the same degree of proficiency as a middle school education from urban areas. This is unfortunately a natural limitation of the data that we are using.

3.2. Occupational Choices

We now turn to occupational choices being made by the workforce in urban and rural areas. Figure 4 shows the distribution of agrarian, blue-collar, and white-collar occupations in urban and rural India across the survey rounds (Panel (a)) as well as the gap in these distributions between the sectors (Panel (b)).

The urban and rural occupational distributions have the obvious feature that urban areas have a much smaller fraction of the workforce in agrarian occupations, while rural areas have a minuscule share of people working in white-collar jobs. The crucial aspect, though, is the share of the workforce in

FIGURE 4. Occupational Distribution of Rural and Urban Samples and Ratio of Urban to Rural Distributions within Three Occupational Categories



Source: Authors' calculations.

Notes: Panel (a) presents the distribution of the workforce across three occupational categories, white-collar, blue-collar, and agriculture for different NSS rounds. Panel (b) presents relative gaps in the distribution of urban relative to rural workers across the three occupational categories. We construct the three occupation categories by combining the three-digit occupation categories in the surveys into three broad groups: *white-collar* occupations such as administrators, executives, managers, professionals, technical, and clerical workers; *blue-collar* occupations, such as sales workers, service workers, and production workers, and *agrarian* occupations including farmers, fishermen, loggers, and hunters. See Section 2 of this paper.

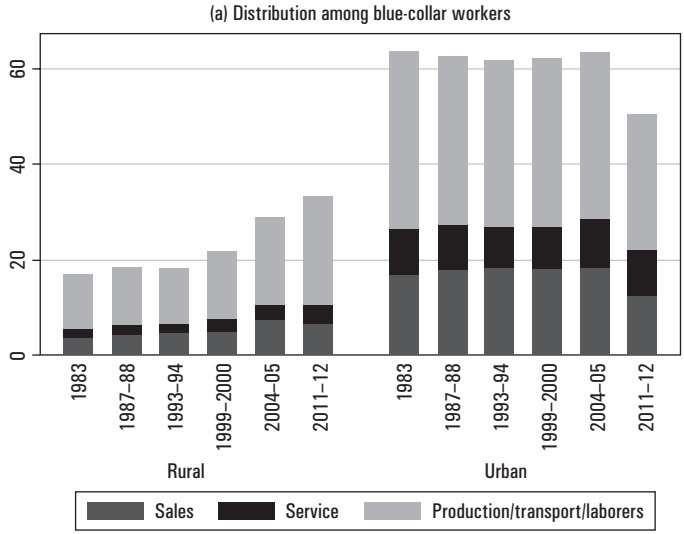
blue-collar jobs that pertain to both services and manufacturing. The urban sector clearly has dominance in these occupations. Importantly, however, the share of blue-collar jobs has been rising in rural areas. In fact, as Panel (b) of Figure 4 shows, the shares of both white-collar and blue-collar jobs in rural areas are rising faster than their corresponding shares in urban areas.

What are the non-farm occupations that are driving the convergence between rural and urban areas? We answer this question by considering disaggregated occupation categories within the white-collar and blue-collar categories. We start with the blue-collar jobs that have shown the most pronounced increase in rural areas.

Panel (a) of Figure 5 presents the breakdown of all blue-collar jobs into three types of occupations. The first group comprises *sales workers*, which include manufacturer's agents, retail and wholesale merchants and shopkeepers, salesmen working in trade, insurance, real estate, and securities, as well as various types of moneylenders. The second group comprises *service workers*, including hotel and restaurant staff, maintenance workers, barbers, policemen, firefighters. The third group consists of *production and transportation workers and laborers*. This group includes, among others, miners, quarry workers and various manufacturing workers.

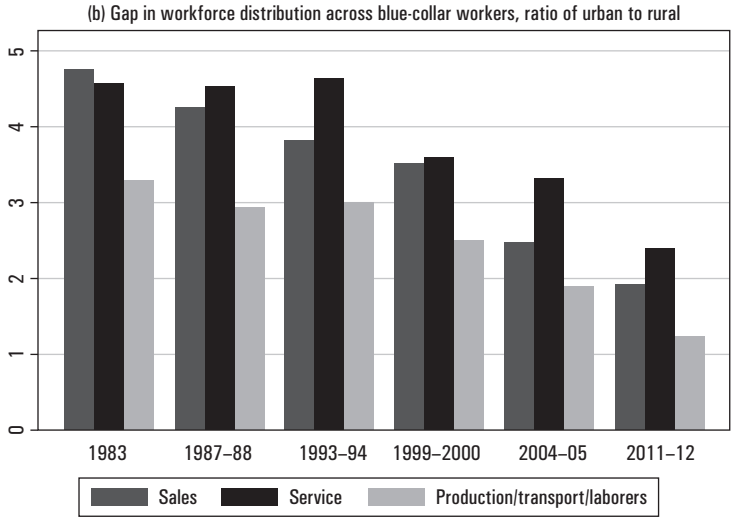
The main result that emerges from Panel (a) of Figure 5 is the rapid expansion of blue-collar jobs in the rural sector. The share of rural workers employed

FIGURE 5. Occupational Distribution of Rural and Urban Samples within Blue-collar Jobs



(Figure 5 Contd.)

(Figure 5 Contd.)



Source: Authors' calculations. See text for details.
Notes: Panel (a) of this figure presents the distribution of the workforce within blue-collar jobs for different NSS rounds. Panel (b) presents relative gaps in the distribution of urban relative to rural workers across different occupation categories among blue-collar workers.

in blue-collar jobs increased from under 18 percent to 37 percent between 1983 and 2012. This increase in the rural sector is in sharp contrast with the urban sector, where the share of blue-collar jobs remained roughly unchanged at around 60 percent during this period. Most of the increase in blue-collar jobs in the rural sector was accounted for by a twofold expansion in the share of production jobs (from 11 percent in 1983 to 27 percent in 2012). While sales and service jobs in the rural areas expanded as well, the increase was much less dramatic. In the urban sector, however, the trends have been quite different: While the shares of sales and service jobs have remained relatively unchanged, the share of production jobs has actually declined.

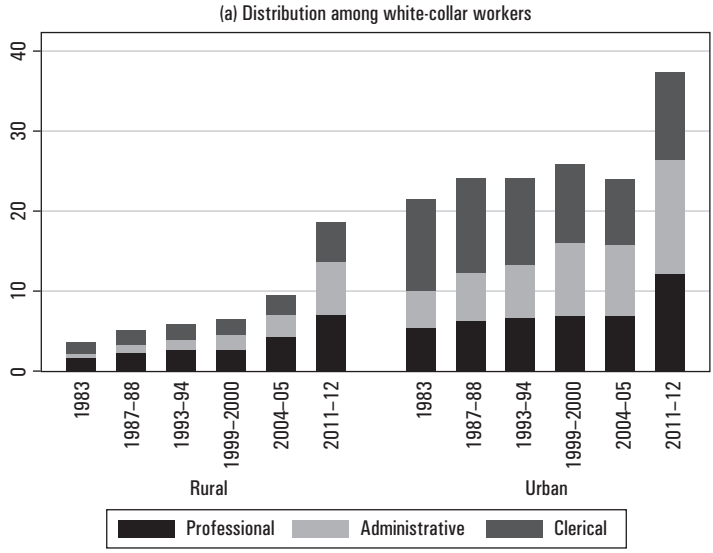
Clearly, such distributional changes should have led to a convergence in the rural and urban occupation distributions. To illustrate this, Panel (b) of Figure 5 presents the relative gaps in the workforce distribution across various blue-collar occupations. The largest gaps in the sectoral employment shares were observed in sales and service jobs, where the gap was more than four times in 1983. The distributional changes discussed above have led to a decline in the urban–rural gaps in these jobs. The more pronounced decline in the relative gap was in production occupations: from 3.25 in 1983 to almost parity in 2012.

Next, we turn to white-collar jobs. Panel (a) of Figure 6 presents the distribution of all white-collar jobs in each sector into three types of occupations. The first is *professional, technical, and related workers*. This group includes, for instance, chemists, engineers, agronomists, doctors and veterinarians, accountants, lawyers, and teachers. The second types of occupations comprise *administrative, executive, and managerial workers*, which include, for example, officials at various levels of the government, as well as proprietors, directors, and managers in various business and financial institutions. The third type of occupation consists of *clerical and related workers*. These include, for instance, village officials, bookkeepers, cashiers, various clerks, transport conductors and supervisors, mail distributors, and communications personnel.

Panel (a) of figure 6 shows that administrative jobs signify the fastest-growing occupation within the white-collar group in both rural and urban areas. It was the smallest category among all white-collar jobs in both sectors in 1983 but has expanded dramatically ever since to overtake clerical jobs as the second most popular occupation among white-collar jobs after professional occupations. Lastly, the share of professional jobs has also increased while the share of clerical and related jobs has shrunk in both the rural and urban sectors during the same time.

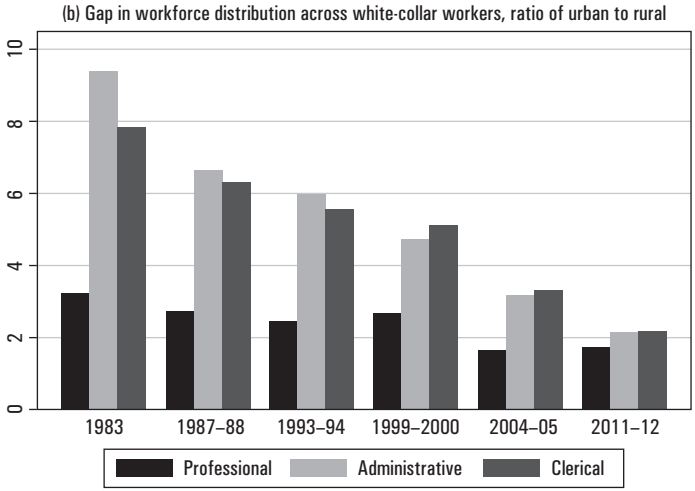
Have the expansions and contractions in various jobs been symmetric across rural and urban sectors? Panel (b) of Figure 6 presents relative gaps

FIGURE 6. Occupational Distribution of Rural and Urban Samples within White-collar Jobs



(Figure 6 Contd.)

(Figure 6 Contd.)



Source: Authors' calculations. See text for details.
Notes: Panel (a) of this figure presents the distribution of the workforce within white-collar jobs for different NSS rounds. Panel (b) presents relative gaps in the distribution of urban relative to rural workers across different occupational categories.

in the workforce distribution across various white-collar occupations. The biggest difference in occupation distribution between urban and rural sectors was in administrative jobs, but the gap has declined very sharply between 1983 and 2012. Similarly, the relative gap in clerical jobs has fallen, though the decline has been more muted. Lastly, the gap in professional jobs has halved during the same period.

Overall, these results suggest that the expansion of the rural non-farm sector has led to a convergence of rural–urban occupations, contrary to the popular belief that urban growth was deepening the rural–urban divide in India.

Is this visual image of sharp changes in the occupation distribution and convergent trends statistically significant? To examine this, we estimate a multinomial probit regression of occupation choices on a rural dummy and a constant for each survey round. The results for the marginal effects of the rural dummy are shown in Table 4. The rural dummy has a significantly negative marginal effect on the probability of being in white-collar and blue-collar jobs while having significant positive effects on the probability of being in agrarian jobs. However, as Panel (b) of the table indicates, between 1983 and 2012, the negative effect of the rural dummy in blue-collar occupations

declined (the marginal effect became less negative), while the positive effect on being in agrarian occupations became smaller, with both changes being highly significant. Since there was an initial under-representation of blue-collar occupations and over-representation of agrarian occupations in the rural sector, these results indicate an ongoing process of convergence across rural and urban areas in these two occupations.

At the same time, the gap in the shares of the workforce in white-collar jobs between urban and rural areas has widened. Note that this result is not inconsistent with Figure 4, which indicates convergence in the workforce distribution in white-collar jobs. The key difference is that Table 4 reports *absolute differences* in workforce distribution between the rural and urban workforce, whereas Figure 4 reports *relative differences* in that distribution. Crucially, blue-collar and agrarian jobs have shown convergence over time in both absolute and relative terms.

TABLE 4. Marginal Effect of Rural/Urban Dummy in Multinomial Probit Regressions for Occupations

Panel (a): Marginal Effects, Unconditional						
	1983	*1987–88*	*1993–94*	*1999–00*	*2004–05*	*2011–12*
White-Collar	–0.1900***	–0.2028***	–0.2042***	–0.2187***	–0.2153***	–0.2904***
	(0.0026)	(0.0024)	(0.0026)	(0.0031)	(0.0033)	(0.0044)
Blue-Collar	–0.4834***	–0.4568***	–0.4580***	–0.4381***	–0.4084***	–0.2692***
	(0.0031)	(0.0029)	(0.0030)	(0.0035)	(0.0038)	(0.0049)
Agrarian	0.6734***	0.6596***	0.6622***	0.6568***	0.6236***	0.5596***
	(0.0023)	(0.0021)	(0.0022)	(0.0023)	(0.0025)	(0.0034)
N	179,646	193,585	172,005	178,803	189,195	132,360

Panel (b): Changes				
	1983 to 1993–94	*1993 to 2004–05*	*2004 to 2011–12*	*1983 to 2011–12*
White-Collar	–0.0142***	–0.0111***	–0.0751***	–0.1004***
	(0.0052)	(0.0059)	(0.0055)	(0.0051)
Blue-Collar	0.0254***	0.0496***	0.1392***	0.2124***
	(0.0043)	(0.0048)	(0.0062)	(0.0058)
Agrarian	–0.0112***	–0.0386***	–0.064***	–0.1138***
	(0.0032)	(0.0033)	(0.0042)	(0.0041)

Source: Authors' calculations.

Notes: Panel (a) of this table reports the marginal effects of the rural/urban dummy from a multinomial probit regression of occupational choices on a constant and a rural dummy for each survey round. Panel (b) reports the change in the marginal effects of the rural dummy over successive decades and over the entire sample period. N refers to the number of observations. Agrarian jobs are the reference group in the regression. Standard errors are in parentheses: *p-value $\leq$ 0.10, **p-value $\leq$ 0.05, ***p-value $\leq$ 0.01.

3.3. Household Consumption

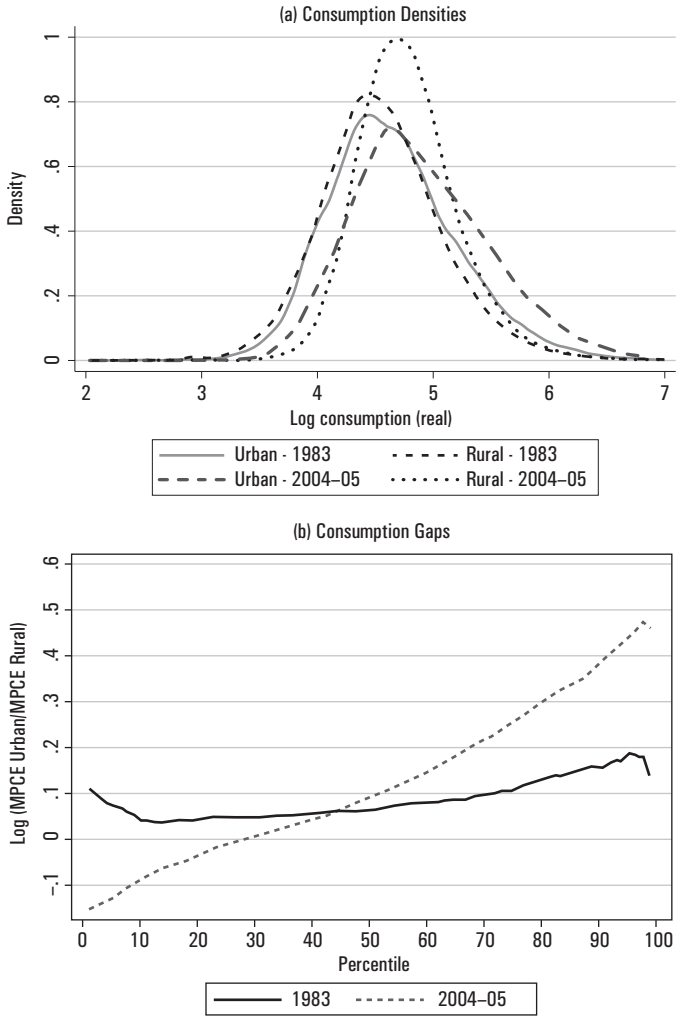
In studying urban–rural real consumption convergence, we are interested not just in the mean or median consumption gaps but rather in the behavior of the real consumption gap across the entire consumption distribution.

We start by taking a look at the distribution of log real MPCE for rural and urban households in our sample. In order to present the results, we break up our sample into two sub-periods: 1983 to 2004–05 and 2004–05 to 2011–12. We do this to distinguish long-run trends since 1983 from the potential effects of the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) that was introduced in 2005. MGNREGA provides a government guarantee of 100 days of wage employment in a financial year to all rural households whose adult members volunteer to do unskilled manual work. This Act could clearly have affected rural and urban wages. To control for the effects of this policy on real wages, we split our sample period into the pre- and post-MGNREGA periods.

We begin with the pre-MGNREGA period of 1983 to 2004–05. Panel (a) of Figure 7 plots the kernel densities of log MPCE for rural and urban households for the 1983 and 2004–05 survey rounds. The plot shows a very clear rightward shift of the consumption density function during this period for both rural and urban households. Panel (b) of Figure 7 presents the percentile (log) MPCE gaps between urban and rural workers for 1983 and 2004–05 household consumption densities functions in those two survey rounds. An upward sloping gap schedule indicates that consumption gaps are higher for higher consumption groups. A rightward shift in the schedule over time implies that the consumption gap has shrunk. The plot for 2004–05 lies to the right of that for 1983 till the 45th percentile, indicating that the gap between poorer urban and rural household consumption declined over this period. Interestingly, in 2004–05, rural consumption was actually higher than urban consumption for the bottom 30th percentile of the consumption distribution. This was in stark contrast to 1983 when urban consumption was higher than rural consumption for all the percentiles.

We now turn to an analysis of the post-MGNREGA consumption distributions. Panel (a) of Figure 8 shows the percentile consumption gaps between rural and urban households in 2004–05 and 2011–12. The figure shows that the urban–rural consumption convergence between the relatively poorer households that we uncovered for the 1983–2005 period reversed itself in the post-reform period. Panel (b) shows that as a result of the widening urban–rural consumption gaps between 2004 and 2012, the percentile consumption gaps in 2011–12 are higher than the corresponding gaps for 1983 for all except the bottom 15 percentiles. In fact, the median consumption premium of urban households increased from under 10 percent

FIGURE 7. Log Consumption Distributions of Urban and Rural Households for 1983 to 2004–05 and Consumption Gaps



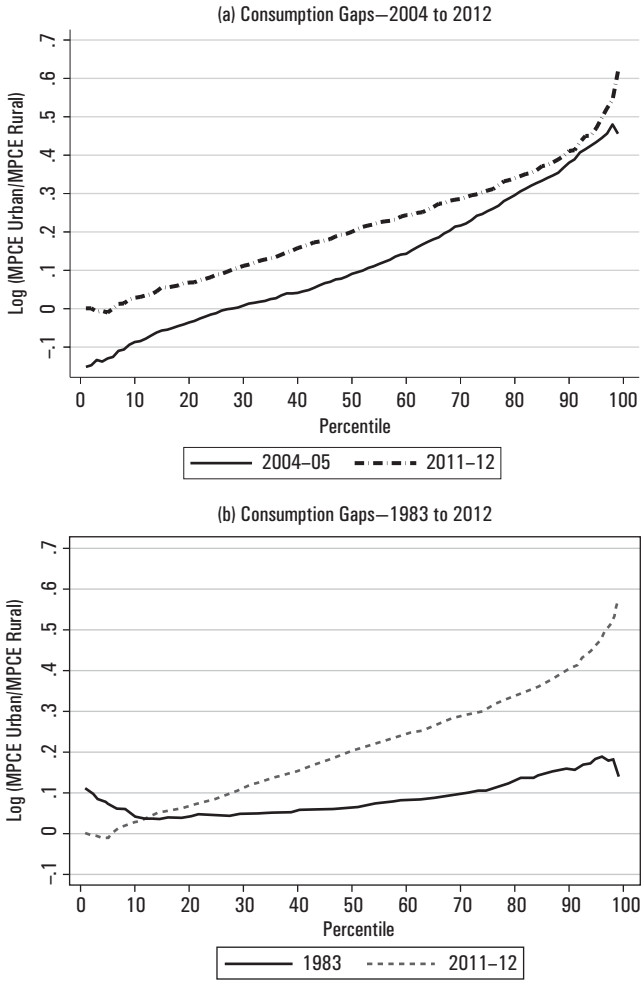
Source: Authors' calculations.

Notes: Panel (a) of this figure shows the estimated kernel densities of log real MPCE for urban and rural households, while Panel (b) shows the difference in percentiles of log MPCE between urban and rural households plotted against the percentile. The plots are for the 1983 and 2004–05 NSS rounds.

to close to 20 percent between 1983 and 2012 as a result of the widening rural–urban consumption dispersion since 2004–05.

To examine whether changes in the urban and rural consumption gaps are statistically significant, we estimate Recentered Influence Function (RIF) regressions developed by Firpo, Fortin, and Lemieux (2009) of the log real

FIGURE 8. Consumption Gaps between Urban and Rural Households, 2004–05 and 2011–12 and 1983 and 2011–12



Source: Authors' calculations.

Notes: Panel (a) of this figure shows the percentile gaps in log real consumption differences between urban and rural households in 2004–05 and 2011–12, while Panel (b) shows the corresponding percentile consumption gaps in 1983 and 2011–12.

consumption in our sample on a constant and a rural dummy for each survey Round. Our interest is in the coefficient on the rural dummy. We perform the analysis for different unconditional quantiles as well as the mean of the wage distribution.[4]

4. We use the RIF approach (developed by Firpo, Fortin, and Lemieux 2009) because we are interested in estimating the effect of the rural dummy for different points of the distribution, not

Panel (a) of Table 5 reports the estimated coefficient on the rural dummy for the 10th, 50th, and 90th percentiles as well as the mean for different survey rounds. The rural status significantly reduced household consumption for all percentiles of the distribution in 1983. Panel (b) of Table 5 shows that the size of the negative rural effect became significantly smaller over time for the 10th percentile between 1983 and 2012 but widened for all other quantiles during this period. These results corroborate the visual impression from Figure 7.

TABLE 5. Are Changes in Urban–Rural Consumption Gaps Significant?

Panel (a): Rural Dummy Coefficient						
	1983	*1987–88*	*1993–94*	*1999–00*	*2004–05*	*2011–12*
10th Percentile	−0.0491***	0.0563***	0.0118***	0.0240***	0.0833***	−0.0279***
	(0.0084)	(0.0073)	(0.0067)	(0.0080)	(0.0084)	(0.0096)
50th Percentile	−0.0637***	−0.0332***	−0.0871***	−0.0970***	−0.0731***	−0.1781***
	(0.0062)	(0.0053)	(0.0053)	(0.0060)	(0.0066)	(0.0082)
90th Percentile	−0.1641***	−0.1802***	−0.2862***	−0.3533***	−0.4245***	−0.4753***
	(0.0107)	(0.0098)	(0.0107)	(0.0117)	(0.0144)	(0.0164)
Mean	−0.0842***	−0.0514***	−0.1120***	−0.1397***	−0.1209***	−0.2088***
	(0.0056)	(0.0056)	(0.0049)	(0.0057)	(0.0064)	(0.0071)
N	97,844	103,079	94,236	98,256	100,229	81,420

Panel (b): Changes				
	1983 to 1993–94	*1993 to 2004–05*	*2004 to 2011–12*	*1983 to 2011–12*
10th Percentile	0.0609***	0.0715***	−0.1112***	0.0212***
	(0.0107)	(0.0107)	(0.0128)	(0.0128)
50th Percentile	−0.0234***	−0.0399***	−0.1050***	−0.1144***
	(0.0082)	(0.0085)	(0.0105)	(0.0103)
90th Percentile	−0.1221	−0.2443***	−0.0508***	−0.3112***
	(0.0151)	(0.0179)	(0.0196)	(0.0196)
Mean	−0.0278***	−0.0089***	−0.0879***	−0.1246***
	(0.0074)	(0.0080)	(0.0096)	(0.0090)

Source: Authors' calculations. See text for details.

Note: Panel (a) of this table reports the estimates of coefficients on rural dummy from RIF regressions of log MPCE on a rural dummy, age, age squared, and a constant. Results are reported for the 10th, 50th, and 90th percentiles. The row labeled "mean" reports the rural coefficient from the conditional mean regression. Panel (b) of this table reports the changes in the estimated coefficients over successive decades and the entire sample period. N refers to the number of observations. Standard errors are in parentheses. *p-value $\leq$ 0.10, **p-value $\leq$ 0.05, ***p-value $\leq$ 0.01.

just the mean. However, since the law of iterated expectations does not go through for quantiles, we cannot use standard mean regression methods to determine the unconditional effect of rural status on wages for different quantiles. The RIF methodology gets around this problem for quantiles. Details regarding this method can be found in Firpo, Fortin, and Lemieux (2009).

A couple of caveats regarding the consumption results are in order. First, the data does not allow us to distinguish between durable and non-durable expenditures of households. While durable expenditures are included in the household MPCE measure, we are unable to compute the flow of services from these durable purchases due to data limitations. Consequently, any systematic differences in durable expenditures between rural and urban households which impact the flow of consumption services over many years are not accounted for in our analysis.

Second, there may be heterogeneity in publicly provided consumption goods between rural and urban areas, which would impact the overall consumption differences between the locations. Since these are not accounted for in the MPCE measure, our analysis does not account for this. While this is an important issue, a careful accounting of this would take us well beyond the remit of this paper.

4. Rural–Urban Gaps in States

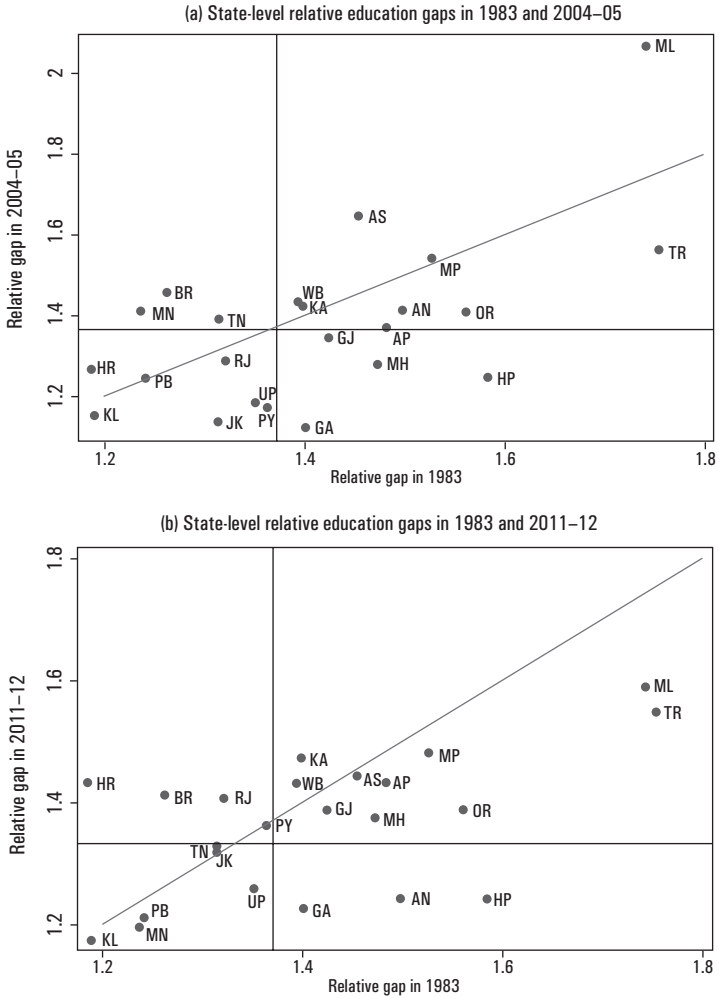
The aggregate patterns in rural–urban gaps in education, occupation, and consumption gaps suggest that there has been a trend towards the narrowing of the gaps since 1983. These patterns raise two important questions. First, do the aggregate trends signify a general phenomenon throughout the country or have they been driven mostly by changing trends in a few, possibly populous states? Second, what are the main factors driving these trends? Answering the first question naturally requires an examination of the trends in individual states. Answering the second question using only the aggregate time-series data is problematic due to the relatively limited length of time since 1983. Analyzing trends in a panel of individual states facilitates more robust identification.

4.1. State Education Gaps

We start with the state-level evidence on rural–urban education gaps. In order to examine the pattern of convergence in individual states over any given time interval, we plot the ratio of the average number of years of education of urban and rural workers in 1983 (the initial year of our sample) against the corresponding gap in the terminal year of our sample. A number greater than one indicates that urban workers enjoy an education premium, that is, they have more years of education than their rural counterparts. Panel (a) of Figure 9 shows the relative gap in the number of years of education of urban and rural workers in 1983 against 2004–05. Panel (b) of the figure shows the relative education gaps in 1983 against 2011–12.

The main takeaway from Figure 9 is that the urban–rural education gaps in 1983 were higher than the corresponding gaps in both 2004–05 (Panel (a) of the figure) and 2011–12 (Panel (b) of Figure 9). The solid diagonal lines in the figure are the 45-degree lines which indicate points of no change in the gaps. In both panels of the figure, most of the observations lie below the 45-degree line, indicating that the gaps in the terminal year were lower

FIGURE 9. Cross-State Educational Convergence, Ratio of Urban to Rural, 1983 compared to 2004–05 and 2011–12



Source: Authors' calculations. See text for details.

Notes: Panel (a) of this figure shows the relative urban–rural gaps in the number of years of education for each state in 1983 and 2004–05, while Panel (b) shows the corresponding education gaps in 1983 and 2011–12. The solid diagonal lines are 45-degree lines.

than in 1983. Importantly, there has been no change in the convergent trend after the introduction of MGNREGA in 2006.

4.2. State Occupation Gaps

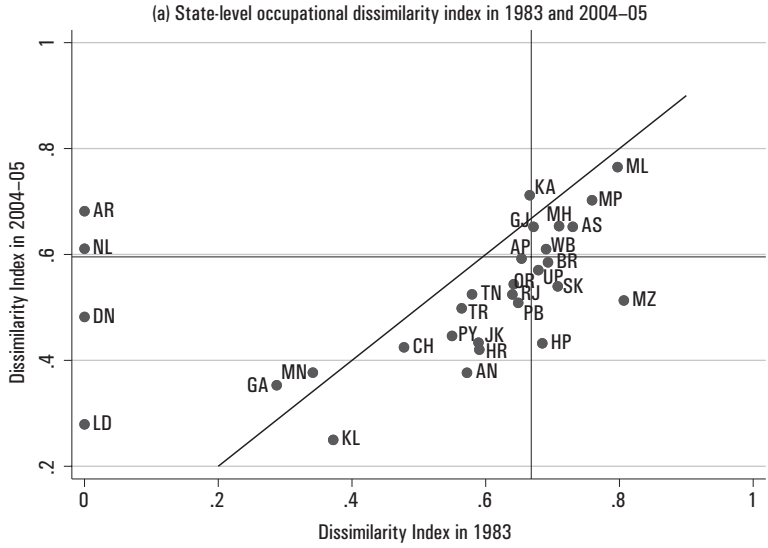
Our second variable of interest is the occupation distribution of rural and urban workers. Our specific interest lies in the evolution of the occupation distribution of rural and urban workers since 1983 in each state: Have they become more similar or dissimilar over time? To examine this issue, we compute an index of dissimilarity. We compute the Duncan Index of Occupational Dissimilarity:

$$D = \frac{1}{2}\sum_j \left| \frac{N_j^U}{N^U} - \frac{N_j^R}{N^R} \right|$$

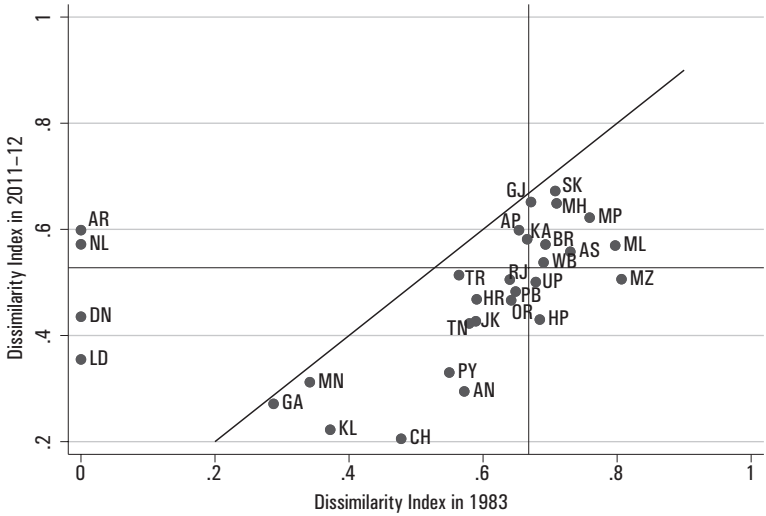
where N_j^k, $k = U, R$ indicates number of Type k workers in occupation j and N^k, $k = U, R$ is the total number of workers of Type k. The Duncan Index is bounded between 0, which is no dispersion, and 1, which indicates maximum dispersion as per this measure. Higher values of D indicate greater dissimilarity between urban and rural workers in their occupation choices.

Panel (a) of Figure 10 shows the Dissimilarity Index in 1983 and 2004–05 for each state, while Panel (b) shows the index for 1983 and 2011–12. Both

FIGURE 10. Cross-State Occupational Dissimilarity Index, 1983 compared to 2004–05 and 2011–12



(b) State-level occupational dissimilarity index in 1983 and 2011–12



Source: Authors' calculations. See text for details.

Notes: Panel (a) of this figure shows the Occupation Dissimilarity Index for each state in 1983 and 2004–05, while Panel (b) shows the corresponding Dissimilarity Index in 1983 and 2011–12. The solid diagonal lines are 45-degree lines.

panels reveal the same pattern: The occupational dissimilarity between urban and rural workers has declined since 1983 for almost all states since most points on the scatter lie below the 45-degree line.

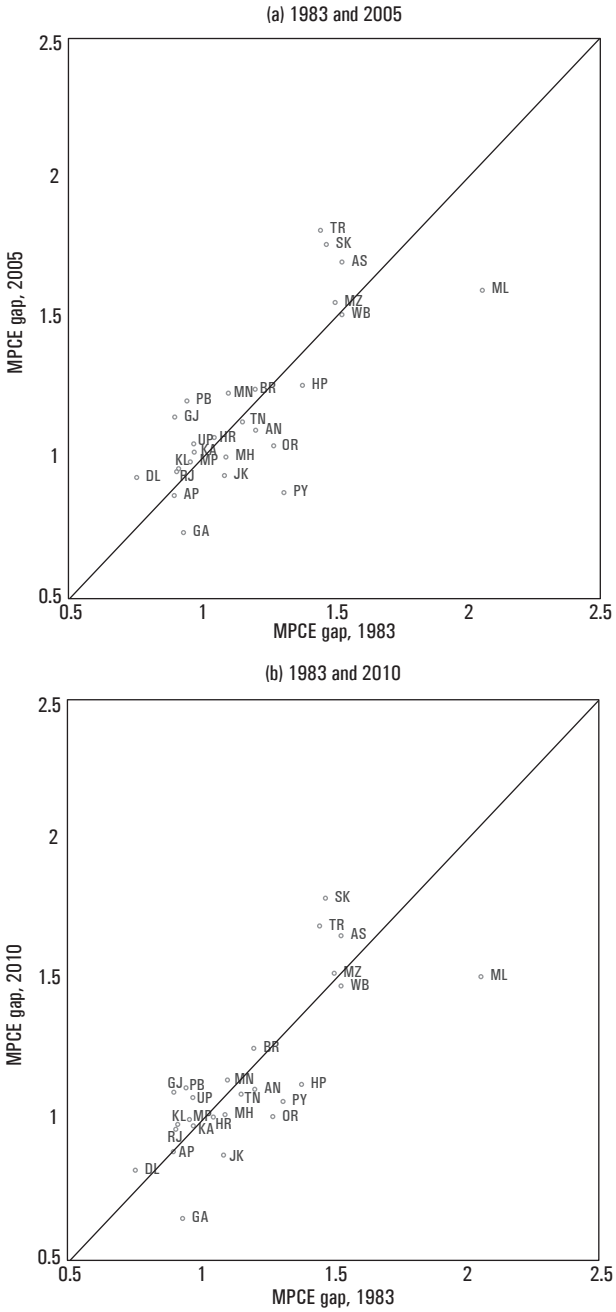
Panels (a) and (b) of Figure 10 also indicate that there was no reversal in the occupational dissimilarity trends till 2004–05 after the introduction of MGNREGA in 2006: The occupation distribution of urban and rural workers has continued to become similar over time.

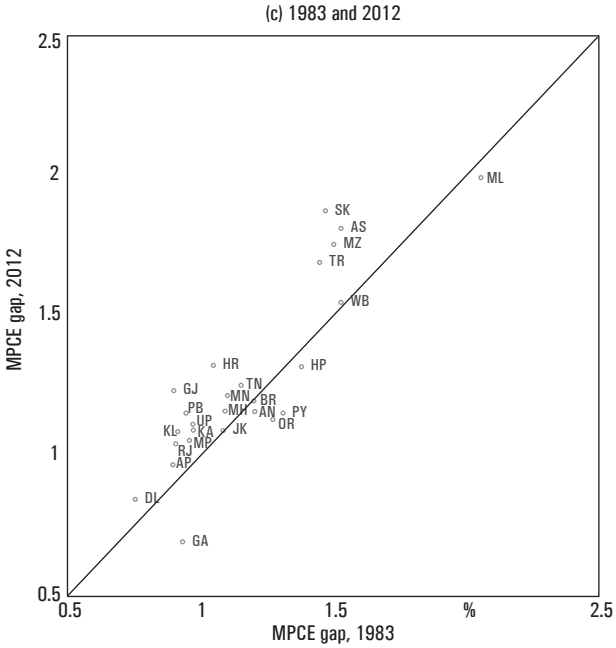
4.3. State Consumption Gaps

We now examine the trends in gaps in monthly per capita consumption expenditures of urban and rural households. Recall that our analysis of the aggregate NSS data showed that the urban–rural consumption gaps had contracted between 1983 and 2004–05 but widened between 2004–05 and 2011–12. Consequently, we examine the trends in the state-wise disaggregated data by breaking up the sample into the pre- and post-2005 trends.

Panel (a) of Figure 11 shows the relative gap in the mean consumption expenditure of urban and rural households in 1983 against the corresponding gap in 2004–05. Panel (b) of the figure shows the mean urban–rural relative consumption gap in 1983 against the gap in 2009–10, while Panel (c) of

FIGURE 11. State-Level Urban–Rural Consumptions Gaps, Ratio of Urban to Rural MPCE in 2005, 2010, and 2012 compared to 1983





Source: Authors' calculations. See text for details.

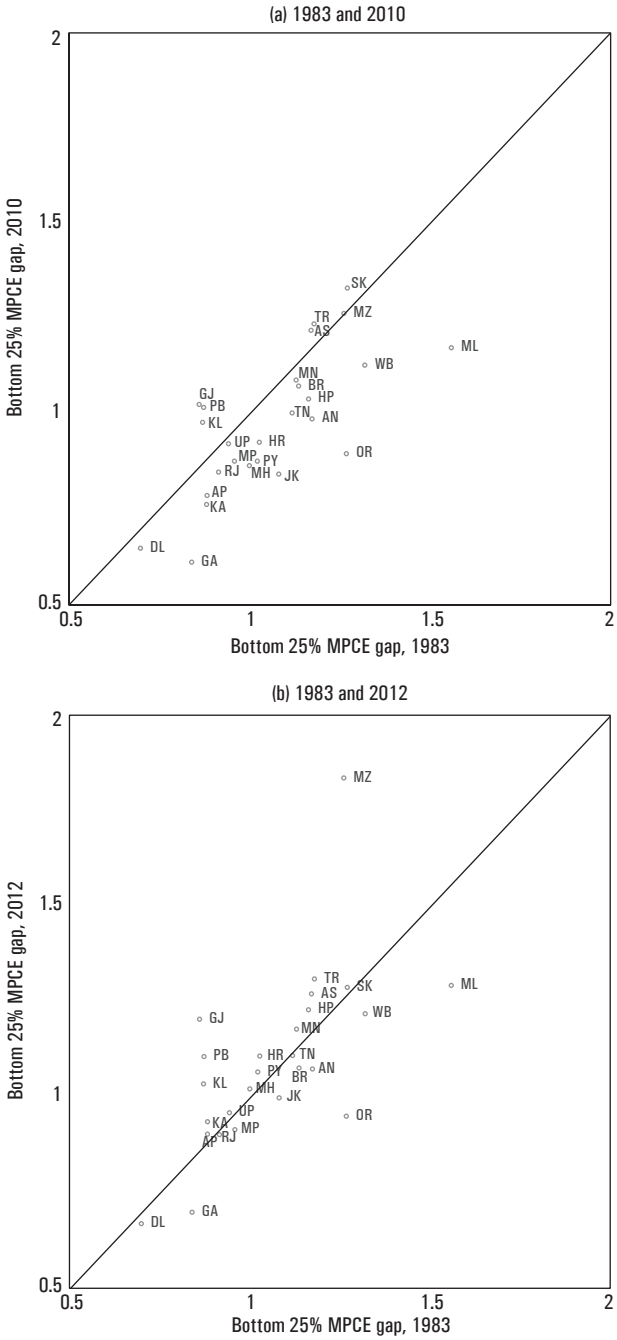
Notes: Panel (a) of this figure shows the urban–rural mean consumption gap for each state in 1983 and 2004–05; Panel (b) shows the gap in 1983 and 2009–10; Panel (c) shows the gap in 1983 and 2011–12. The solid diagonal lines are 45-degree lines.

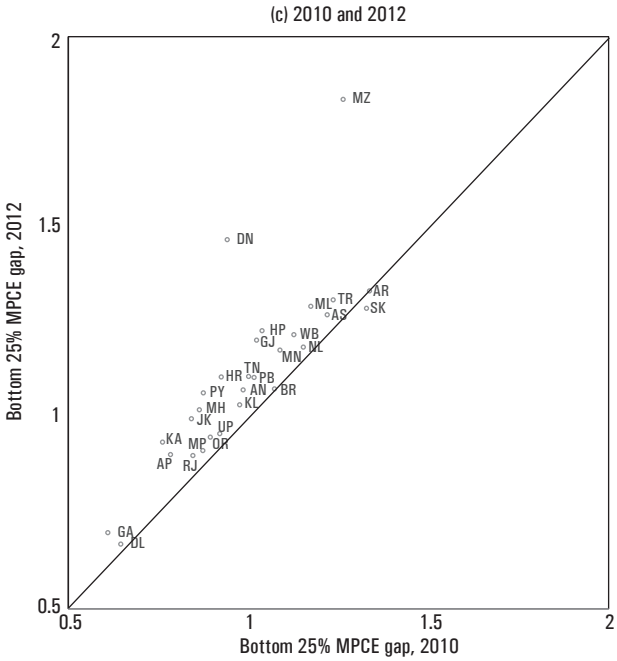
the figure shows the mean urban–rural relative consumption gap in 1983 against the gap in 2011–12.

Consistent with the aggregate pattern we saw in Figure 7, the urban–rural consumption gap declined in most states between 1983 and 2005 since the majority of the scatter points are below the 45-degree line. This pattern of convergence actually continued till the 2009–10 NSS round as indicated by the scatter of points below the 45-degree line in Panel (b) of the figure. The scatter of points, however, shifts up in Panel (c) of the figure, indicating that there has been a widening of the consumption gaps between urban and rural households since 2009–10.

The preceding figures showed the patterns in the urban–rural gaps in mean consumption expenditure. Do these patterns apply to the entire consumption distribution? The question is important from a distributional perspective. As we saw in the aggregate picture in Figure 7, the convergence trend was not uniform across the consumption distribution. We examine this in the state-level data in Figure 12, which shows the consumption gaps

FIGURE 12. State-Level Urban–Rural 25th Percentile Consumptions Gaps, Ratio of Urban to Rural MPCE, comparing across 1983, 2010, and 2012





Source: Authors' calculations.

Notes: Panel (a) of this figure shows the urban–rural 25th percentile consumption gap for each state in 1983 and 2009–10; Panel (b) shows the gap in 1983 and 2011–12; Panel (c) shows the gap in 2009–10 and 2011–12. The solid diagonal lines are 45-degree lines.

for the 25th percentile of the distribution, and Figure 13, which shows the evolution of the median consumption gaps.

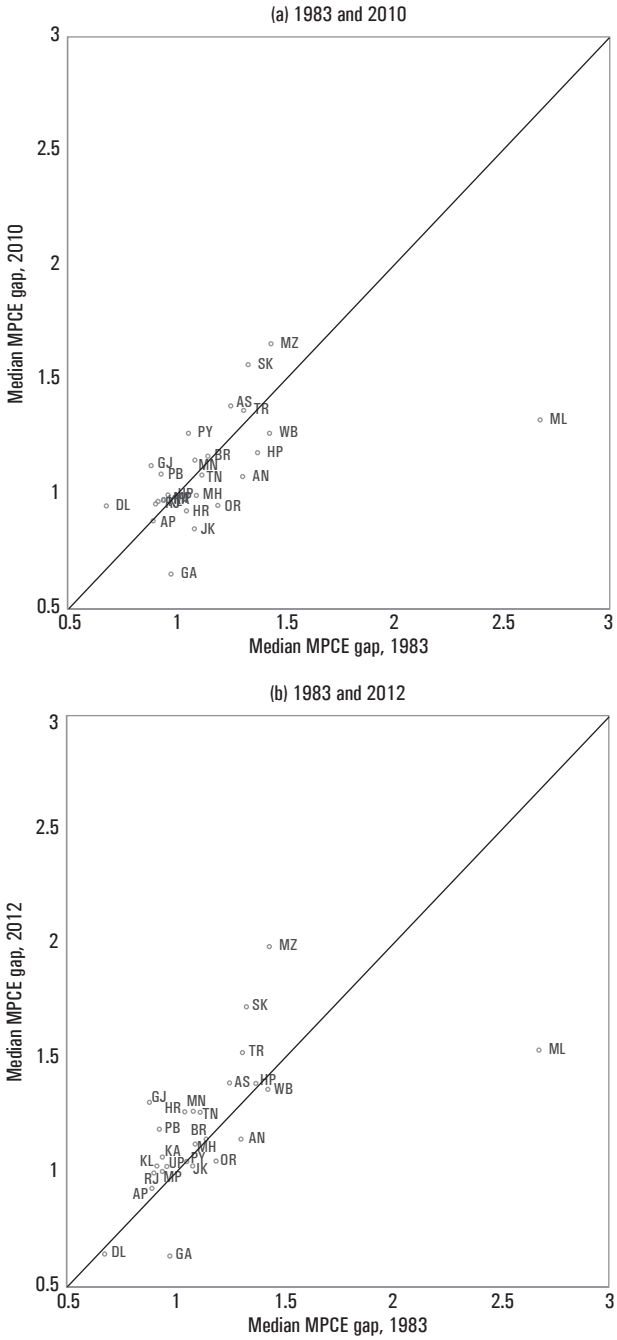
Both figures reveal a similar trend: narrowing consumption gaps across the consumption distribution from 1983 till 2004–05 and widening of the gap between 2009–10 and 2011–12. Between 2004–05 and 2009–10, the consumption gaps either stayed constant or marginally declined. These findings corroborate the patterns in the aggregate data that we saw previously.

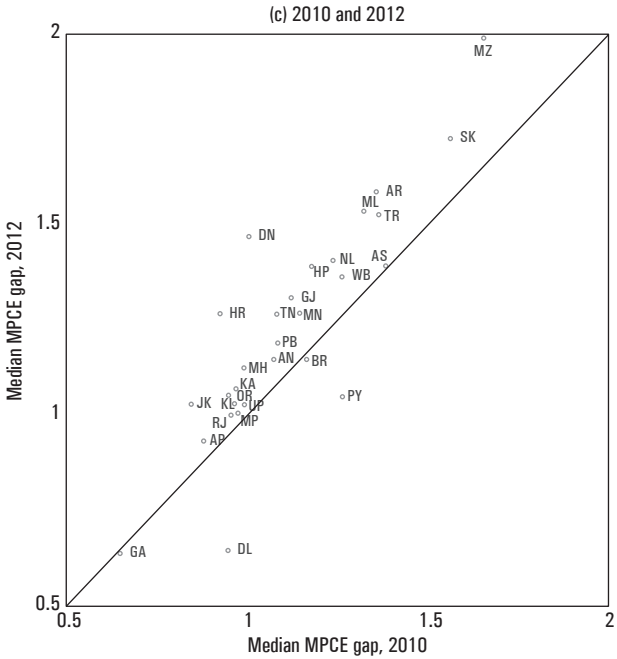
5. Explaining the Trends

What can explain the differences in state-level gaps in consumption expenditures? One hypothesis is that these gaps are influenced by differences in income levels across states and their different growth rates.

To investigate this hypothesis, we construct a panel of state-year observations on urban–rural consumption expenditures gaps and various state characteristics. The years correspond to the NSSO survey rounds. We then

FIGURE 13. State-Level Urban–Rural Median Consumptions Gaps, Ratio of Urban to Rural MPCE, comparing across 1983, 2010, and 2012





Source: Authors' calculations. See text for details.
Notes: Panel (a) of this figure shows the urban–rural median consumption gap for each state in 1983 and 2009–10; Panel (b) shows the gap in 1983 and 2011–12; Panel (c) shows the gap in 2009–10 and 2011–12. The solid lines are 45-degree lines.

estimate a panel regression of urban–rural consumption expenditure gaps on the initial level of per capita income, measured by the (log) per capita NSDP, and the growth rate of per capita NSDP in the period preceding the measured gap. To account for unobserved state-level characteristics, we include state fixed effects. In addition, since we showed that consumption gaps exhibit common trends across states, we also included survey round time fixed effects in the regressions. This specification is estimated for the mean gaps, median gaps, and gaps in the 25th and the 75th percentiles. The results are presented in Table 6.

We found that initial per capita income has a positive and significant effect on consumption expenditure gaps at different points of the distribution, that is, higher initial income is associated with higher gaps. Our results also imply that high per capita NSDP growth led to higher consumption expenditure gaps. While this may seem somewhat surprising, it suggests that high income and faster growth benefited urban areas more relative to rural areas. These results are confirmed for the median and percentile gaps.

TABLE 6. Consumption Expenditure Gaps

	(1) MPCE Mean	*(2) MPCE Median*	*(3) MPCE 25th*	*(4) MPCE 75th*
Per capita NSDP growth	1.127***	1.010**	1.572***	2.029**
	(0.409)	(0.446)	(0.420)	(0.814)
Log (initial per capita NSDP)	0.141**	0.161**	0.167**	0.126
	(0.069)	(0.075)	(0.070)	(0.137)
N	162	162	162	162
R-square	0.203	0.119	0.174	0.13

Source: Authors' calculations.
Note: The regressions include state and time (round) fixed effects. Standard errors in parentheses. *p ≤ 0.10, **p ≤ 0.05, ***p ≤ 0.01.

Next, we include additional controls in our regression specification. Specifically, we add education gaps (measured by the ratio of the number of years of education in urban to rural areas), and a measure of urbanization measured by the rural employment share. The results of these regressions are summarized in Table 7.

It is easy to see that the initial per capita NSDP and per capita NSDP growth rates retain their positive coefficients and significance. In addition, urban–rural education gaps exhibit a positive and significant effect on consumption gaps. Specifically, a one-unit increase in the number of years of the education gap leads to about 0.3–0.4 unit increase in the consumption expenditure gap. Lastly, an increase in the rural employment

TABLE 7. Consumption Expenditure Gaps, Extended Regressions

	(1) MPCE Mean	*(2) MPCE Median*	*(3) MPCE 25th*	*(4) MPCE 75th*
Per capita NSDP growth	0.818**	0.750*	1.246***	1.977**
	(0.379)	(0.432)	(0.398)	(0.794)
Log (initial NSDP)	0.105*	0.131*	0.127*	0.130
	(0.063)	(0.072)	(0.066)	(0.132)
Edu Gap	0.404***	0.340***	0.390***	0.223
	(0.080)	(0.092)	(0.084)	(0.168)
Initial mean rural employment share	0.175*	0.152	0.046	0.643***
	(0.089)	(0.101)	(0.093)	(0.186)
N	162	162	162	162
R-square	0.351	0.218	0.296	0.215

Source: Authors' calculations.
Note: The regressions include state and time (survey round) fixed effects. Standard errors in parentheses. *p ≤ 0.10, **p ≤ 0.05, ***p ≤ 0.01.

TABLE 8. Changes in Consumption Expenditure Gaps

	(1) *ΔMPCE Mean*	*(2)* *ΔMPCE Median*	*(3)* *ΔMPCE 25th*	*(4)* *ΔMPCE 75th*
Per capita NSDP growth	0.313	0.413	0.285	1.249
	(0.401)	(0.433)	(0.344)	(0.981)
ΔEdu gap	0.325***	0.175*	0.107	0.362*
	(0.084)	(0.091)	(0.072)	(0.205)
ΔMean rural employment share	−0.519*	−0.866***	−0.093	−1.266*
	(0.283)	(0.306)	(0.243)	(0.693)
N	159	159	159	159
R-square	0.131	0.098	0.026	0.068

Source: Authors' calculations.
Note: The regressions include state fixed effects. Standard errors in parentheses. *p ≤ 0.10, **p ≤ 0.05, ***p ≤ 0.01.

share is associated with an increase in the consumption gap, and this effect is significant for the mean and the top 25 percent of the consumption expenditure distribution.

Next, we turn to the changes in the consumption expenditure gaps over time (across rounds). We relate those changes to per capita NSDP growth, changes in the education gap between urban and rural areas, and changes in rural employment shares. The results of these regressions are presented in Table 8. We found that changes in the education gap map in a significant way into changes in the consumption gap, with the relationship between the two being positive. Changes in urbanization (measured by the change in the rural employment share) contribute to the reduction in the gap, with the effects being particularly pronounced for the mean and median gaps, as well as the 75th percentile of consumption distribution.

Overall, our results emphasize the importance of education in explaining both levels and changes in consumption gaps. Urbanization also plays a role, with greater urbanization reducing the levels of gaps but increasing them in changes.

6. Conclusion

The evolution of inequality in times of economic growth is a continuing area of applied interest to academics and policymakers. In this paper, we have examined the evolution of disparities between rural and urban workers in India between 1983 and 2012, a period that saw a sharp takeoff in growth in the country.

Our results suggest that rural–urban gaps in education and occupation choices have contracted sharply and significantly during this period. The evidence on consumption gaps between rural and urban households is more nuanced. There was a decline in the rural–urban consumption gaps for the bottom 45th percentile of households between 1983 and 2004–05. Some of this narrowing inequality has, however, reversed itself between 2004–05 and 2011–12. As a result, for the entire period 1983–2012, rural–urban consumption gaps have declined for only the bottom 15 percent of households. This widening of rural–urban consumption gaps since 2004–05 is a puzzle, particularly since the employment guarantee scheme for rural workers, MGNREGA, was introduced in 2006.

Our analysis of the data at the state level found that the aggregate patterns were general: The trends in rural–urban gaps in education, occupation, and consumption in most states were similar to the aggregate trends in India between 1983 and 2012. The state-level analysis, however, allowed us to identify the proximate determinants of rural–urban consumption gaps. We find that states with higher per capita income and higher per capita income growth are associated with higher consumption gaps, while states with lower education gaps tend to have smaller consumption gaps. Explaining changes in consumption gaps is more difficult, though changes in the education gap and changes in the rural labor force share do have significant explanatory power.

The results of rural–urban inequality also extend to the wage differences between rural and urban workers. In Hnatkovska and Lahiri (2016), we have shown that rural–urban wage disparities in India declined very sharply between 1983 and 2010. In fact, the size of the decline in the rural-urban wage gaps in India is a bit of a puzzle since it cannot be explained by standard worker covariates, such as education and demographics. Indeed, in Hnatkovska and Lahiri (2016), we show that the rural–urban wage gap dynamics in India stand in sharp contrast to China, where they have actually worsened since 1988.

These results on rural–urban inequality corroborate our findings of declining inequality across castes, falling inequality across genders, and the robustness of these patterns across states since 1983 that we have previously documented in Hnatkovska, Lahiri, and Paul (2012), Hnatkovska and Lahiri (2012), Hnatkovska, Lahiri, and Paul (2013), and Bhattacharjee, Hnatkovska, and Lahiri (2015). Clearly, the past three decades have seen a widespread decline in inter-group inequality for a large set of groups and across different socio-economic markers.

A couple of concerns regarding our results are worth clarifying. First, any discussion of rural–urban disparities is subject to issues surrounding the re-classification of rural areas into urban areas over time. India's urban

population growth was concentrated in large cities with populations exceeding one million. In 1981, there were just 12 cities in India with populations exceeding a million; they accounted for 26 percent of the urban population. By 2011, the number of million plus cities rose to 53, which collectively accounted for 43 percent of the urban population. Crucially, the average population density of the million-plus cities declined from 39,000/sq. km to 26,000/sq. km.[5] In effect, a number of rural areas adjoining large urban areas got absorbed into their proximate cities.

The effect of this form of re-classification on the rural–urban gaps is uncertain. It depends on the economic positions of the groups being re-classified, both with respect to the typical rural households and the typical urban households. As an example, if the re-classified rural area was amongst the higher economic groups in rural areas and amongst the lower economic groups in urban areas, then such a re-classification would reduce both rural and urban averages, leaving their effect on the relative gaps unclear. Consequently, we do not believe our results are trivially induced by such re-classifications.

A second issue is with respect to migrants. Our data does not have information on migrants across all rounds. A logical question that could be asked is whether rural consumption levels are being boosted by remittances from rural migrants to urban areas. This is certainly possible. However, this channel would imply that household consumption gaps between rural and urban areas should have declined faster than wage gaps between rural and urban workers. Our results on wage gaps in Hnatkovska and Lahiri (2016) suggest the opposite: Rural–urban wage gaps have shrunk much more sharply than consumption gaps.

Overall, our results suggest that periods of rapid economic growth are periods of declining inter-group inequality. In effect, growth tends to lift all boats.

Appendix on the Data

The National Sample Survey Office (NSSO), set up by the Government of India, conducts rounds of sample surveys to collect socio-economic data. Each round is earmarked for particular subject coverage. We use the latest six large or thick quinquennial Rounds—38 (January–December 1983), 43 (July 1987–June 1988), 50 (July 1993–June 1994), 55 (July 1999–June 2000), 61 (July 2004–June 2005), and 68 (July 2011–June 2012) on employment

5. Data on population and urban trends are derived from the Census of India (various rounds) and IIHS (2011).

and unemployment (Schedule 10). Rounds 38 and 55 also contain migration particulars of individuals.

The survey covers the whole country except for a few remote and inaccessible pockets. The NSS follows multi-stage stratified sampling with villages or urban blocks as first-stage units (FSUs) and households as ultimate stage units. The fieldwork in each round is conducted in several sub-rounds throughout the year so that seasonality is minimized. The sampling frame for the FSU is the list of villages (rural sector) or the NSS Urban Frame Survey blocks (urban sector) from the latest available census. The NSSO supplies household-level multipliers with the unit record data for each round to help minimize estimation errors on the part of researchers. The coding of the data changes from round to round. We recoded all changes to make the variables uniform and consistent over time.

In our data work, we only consider individuals that report their three-digit occupation code and education-attainment level. Occupation codes are drawn from the National Classification of Occupation (NCO) 1968. We use the "usual" occupation code reported by an individual for the usual principal activity over the previous year (relative to the survey year). The dataset does not contain information on the number of years of schooling for the individuals. Instead, it includes information on general education categories given as: (a) not literate-01, literate without formal schooling: EGS/NFEC/AEC-02, TLC-03,[6] others-04; (b) literate: below primary-05, primary-06, middle-07, secondary-08, higher secondary-10, diploma/certificate course-11, graduate-12, post-graduate, and above-13. We aggregate these into five similarly sized groups as discussed in the main text. We also convert these categories into the number of years of education. The mapping we have used has been discussed in the main text.

The NSS only reports activities undertaken by an individual over the previous week (relative to the survey week). Household members can undertake more than one activity in the reference week. For each activity, we know the "weekly" occupation code, number of days spent working in that activity, and the wage received from it. We identify the main activity for the individual as the one in which he/she spends the maximum number of days in a week. If there is more than one activity with an equal number of days worked, we consider the one with paid employment (wage is not zero or missing). Workers sometimes change their occupation due to seasonality or for other reasons. Lastly, we drop observations if the total number of days worked in the reference week is more than seven.

6. EGS is Education Guarantee Scheme, NFEC is Non-formal Education Courses, AEC is Adult Education Centres, and TLC is Total Literacy Campaign.

References

Bhattacharjee, S., V. Hnatkovska, and A. Lahiri. 2015. "The Evolution of Gender Gaps in India." *India Policy Forum 2014–15*, 11: 119–149. New Delhi: National Council of Applied Economic Research.

Firpo, S., N. M. Fortin, and T. Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica,* 77(3): 953–973.

Hnatkovska, V. and A. Lahiri. 2012. "The Post-Reform Narrowing of Inequality across Castes: Evidence from the States." In J.N. Bhagwati and A. Panagariya (eds.), *Reforms and Economic Transformation in India*, pp. 229–252. Oxford: Oxford University Press.

———. 2016. "Urbanization, Structural Transformation and Rural-Urban Disparities." *Working Paper*. Vancouver: University of British Columbia.

Hnatkovska, V., A. Lahiri, and S. Paul. 2012. "Castes and Labor Mobility." *American Economic Journal: Applied Economics,* 4(2): 274–307.

———. 2013. "Breaking the Caste Barrier: Intergenerational Mobility in India." *Journal of Human Resources,* 48(2): 435–473.

IIHS. 2011. "Urban India 2011: Evidence Report." Bengaluru: Indian Institute for Human Settlements.

Nguyen, B.T., J.W. Albrecht, S.B. Vroman, and M.D. Westbrook. 2007. "A Quantile Regression Decomposition of Urban–Rural Inequality in Vietnam." *Journal of Development Economics,* 83(2): 466–490.

Qu, Z.F., and Z. Zhao. 2008. "Urban–Rural Consumption Inequality in China from 1988 to 2002: Evidence from Quantile Regression Decomposition." *IZA Discussion Papers 3659*. Bonn: Institute for the Study of Labor (IZA).

Wu, X., and J.M. Perloff. 2005. "China's Income Distribution, 1985–2001." *The Review of Economics and Statistics,* 87(4): 763–775.

Young, A. 2013. "Inequality, the Urban–Rural Gap and Migration." *The Quarterly Journal of Economics,* 128(4): 1727–1785.

To view the entire video of this IPF session and the General Discussion that ended the session, please scan this QR code or use the following URL:
https://www.youtube.com/watch?v=vZOh93diE3c

Comments and Discussion*

Chair: **Kaushik Basu**
Cornell University

Rohini Somanathan
Delhi School of Economics

I think this is the perfect IPF paper because it provides us with the numbers and perspective we need for sensible discussions about policy. The paper uses seven thick rounds of the NSS between 1983 and 2012 and tracks rural–urban disparities in education, non-farm employment, and consumption. The analysis is done both at the national and at the state levels and the authors look beyond averages to the performance of particular quantiles in the distribution. They ask how growth is related to structural transformation in rural India—an enormously important policy question. Without these types of figures at hand, we are often discussing policy in a vacuum.

The following patterns emerge from the analysis: (a) average rural–urban education gaps are shrinking; (b) non-farm employment is rising; (c) rural–urban consumption gaps narrowed for the bottom 15 percent in the consumption distribution but widened for higher quantiles; and (d) at the state level, education gaps explained consumption gaps with consumption inequality being most prevalent in the richest and fastest-growing states.

My first set of comments relates to measurement issues. I then go on to alternative ways of decomposing the data to understand the mechanisms underlying the patterns in education and consumption that you observe.

My first comment on measurement relates to sample selection. I was a bit puzzled by what the authors kept out of the sample. They excluded female-headed households and also those not between the ages of 16 and 65 years. This may be a problem for measuring educational convergence because those above 16 years might still be enrolled in educational institutions.

So if, for example, secondary and higher education is much higher in urban areas, then convergence is going to be over-estimated because these

* To preserve the sense of the discussions at the India Policy Forum, these discussants' comments reflect the views expressed at the IPF and do not necessarily take into account revisions to the conference version of the paper in response to these and other comments in preparing the final, revised version published in this volume. The original conference version of the paper is available on www.ncaer.org.

people are not kept in the sample. So the authors might want to think of doing this with completed education, maybe look at 24–65-year olds.

Another measurement issue relates to the classification of employment. In periods of structural transformation in other Asian countries (China is a good, recent example), agricultural work was combined with other work and the non-farm sector was growing while agriculture remained the main activity for many people. The authors may, therefore, want to examine contributions to the non-farm sector for those who are primarily in agriculture by looking more closely at secondary activities. Another small point on occupation: The authors use "agrarian," "white-collar," and "blue-collar" as categories. What do we mean when thinking about white-collar? If someone opens a little kiosk for cellphone repair, are they white-collar or not? We see these small enterprises, selling a little bit in little shops dotted all over the country. Where do you place them?

A bigger question relates to the spatial distribution of urban growth. If one looks at the map of India for the Census years 2001 and 2011, the majority of urbanization is actually taking place in the West and the South. There is no growth in the rural population of Andhra Pradesh, Karnataka, and Kerala, while there is rapid rural population growth in the northern heartland. If we look at rural–urban gaps at a macro level, we are disproportionately capturing the West and the South for urban India and the North for rural India. Thus, the authors may want to look at state-wide differences, keeping this in mind.

Turning now to rural–urban gaps, we see the maximum movement from below primary to primary level in rural India and from primary to middle school level in urban India. One can think about decomposing this gap in terms of what households are doing differently given the facilities they have, and what is happening to the facility locations. There was a big expansion in primary schooling in the 1970s and 1980s, so now there is a primary school within easy access of most villages and within the village for 80 percent of villages. If this expansion in facilities was responsible for the closing of education gaps, then for the convergence to continue in future, we would need to expand rural facilities at higher levels. We still have high schools in only about one-fifth of Indian villages.

The authors might also want to exploit the NSS participation and expenditure survey in 2007–08 for more detailed information on household investments in education. These would allow for the mapping of quantiles in education with those of consumption expenditure. Broadly, it is observed that the bottom quintile went mainly to public schools for primary schooling, while 30 percent of the top quintile attended private schools. However,

when one goes from primary schools to high schools, things seem to change. So in the bottom quintile, 50 percent of the 12–18-year olds are enrolled in a public school and 5 percent are enrolled in a private school. For the top quintile, we have 58 percent enrolled in a public school and 20 percent in a private school. Public school expansion is, therefore, going to have quite different effects on educational inequality depending on the types of schools that are built.

Every time someone discusses this paper, they will see it through the lens of their own work. Because mine has been mainly on public goods, when I look at consumption convergence, I think about systematic measurement errors caused by the fact that public good access is not accounted for in income. If a household goes to a public school or hospital, it is not going to show up in their consumption expenditure, whereas if they go to a private hospital, it will. So households that don't have adequate public schooling and healthcare look richer in NSS consumption data than they really are because they are forced to go to private providers. How does this affect the convergence results in this paper?

The above comments suggest a few avenues for the authors to explore with their data. Overall, this is a very nice paper that has direct implications for policy and also opens up a number of interesting research questions on the spatial correlates of macroeconomic trends in education, consumption, and employment. I look forward to more of the authors' work in this area.

Jeff Hammer
NCAER

This paper is a straightforward comparison of the changes in consumption, education, and occupation between urban and rural India over the past few decades using repeated rounds of NSS. It finds that not only have all three changed (for the better, if we judge farm work to not be so great) but that differences between urban and rural areas have declined as well. I am sure it is true that there has been an overall improvement, but it is not possible to say honestly if the disparities between urban and rural areas have changed, and if so, in which direction. The reasons, unfortunately, lead us deep into the weeds of how these three variables are measured in the NSS surveys.

First off, on occupation, of course, there has been convergence. There has clearly been a big increase since the mid-1980s in off-farm employment in rural areas, which is found in the NSS surveys and many other sources. Whatever people do off-farm is likely to look much more like categories

appearing in urban areas, so the two regions will look more similar over time. It is nice to see this documented. A hypothesis that needs to be looked at is whether the liberalization of 1991 had anything to do with that trend and why.

In education, things are a bit more obscure. Yes, averages over years of schooling might be closer together between urban and rural areas. But two characteristics of "years of schooling" make the interpretation harder than it seems. As has been noted in the IPF Roundtable on the New Education Policy at this IPF, there is no clear connection between grades completed and any meaningful definition of education such as the ability to read, write, and do arithmetic at typical, age-appropriate levels. It is easy to get children into school and pass them from grade to grade, and that is apparently what has happened in both rural and urban areas. Since the rural areas in the 1980s were way behind the urban areas, simple measurements are going to favor increases starting from such a low base. Whether there was convergence in the real learning of reading, among other things, is anyone's guess.

The quantile regressions that the paper presents confirm this point. There seems to be a reasonable amount of explanatory power of education on consumption (say) at lower levels of both. At higher levels, the connection is much looser. At higher levels of education, the appropriate jobs actually require that knowledge has been gained in school. At lower levels, having been through school at all will lead to somewhat better jobs, making the connection between education and consumption clearer. This disappears at the higher end.

The second aspect of the education story is an artifact of the nature of its measurement. "Progress" or "development" or "consumption"—all concepts of a higher standards of living—are continuous and unbounded. But education is tightly bounded at 12+ years, and well before that, years tend to bunch up at the 5th, 8th, 10th, and 12th standards. So if there is general overall improvement, there will naturally be "convergence" since there is a limit to how high people tend to go through schooling, getting stuck at those specific cutoffs.

As for consumption itself, here is where the nature of the data raises most problems. The underlying theme of this IPF is the need to collect better data and handle it with care. While a workhorse of empirical research, "consumption per capita" is not exactly the same as the underlying concept of the standard of living. In the Indian case, the operational definition of consumption per capita is the particular module of the NSS, not necessarily a perfect match with that underlying concept. The intention is not to criticize the methods used in the NSS, since fixing at least some of its problems is very hard. However, there are at least four problems with its measurement that could very well differentially affect measurements in urban and rural

areas, causing even more problems in making conclusions about convergence between them.

Specifically, the NSS consumption module does not include: (a) housing (mostly, however, this has been corrected in the most recent NSS Round that has not been released), (b) consumer durables older than one year, (c) adult equivalent weights when making per capita calculations, and (d) the true market value of goods that are subsidized by the government. Each of these four differentially affects urban and rural sectors, so the gaps in true consumption may vary over time due to what is captured or not in the surveys.

If the economy were static, these left-out items would not vary much over time or between sectors. But we are in a period of rapid growth and there may be major discrepancies between changes in the true versus the measured concepts of well-being.

On durables, we have spending on new televisions but not the amortized value of ownership over the whole life of the appliance. Since there has been more rapid electrification in rural than urban areas, the former increasing from 20 percent to over 80 percent since the 1980s, we are likely to have many more new purchasers of things that use electricity in rural areas relative to urban areas. The NSS captures the flow of money for expenditure, but we do not get the flow value of consumption derived from the stock of durables. In a growing sector, the ratio of new purchases to the stock will be higher and the bias of measured consumption will be lower. If rural areas are getting electricity faster than urban areas, we will be overstating the improvement in consumption in rural areas. The true gap will be larger.

On adult equivalents, the problem is that all calculations are in pure per capita terms, that is, a household consisting of three adult brothers and a household of a couple and their infant daughter have the same denominator of three. Obviously, the baby does not eat as much as an adult nor does she demand as much furniture, living space, clothing, or anything as much as an adult. The couple did not actually get 33 percent worse off by having the baby. Some might argue they are better off. But in straight per-capita terms, no adjustment for household composition (or anything else related to scale economies at the household level) is made. This probably affects the comparison of urban and rural areas.

Table 1 of the paper has summary statistics showing family sizes getting smaller faster in cities. The questions are: How does the composition of the family change? What is the relative mismeasurement of the underlying concept? If the difference was all the number of children, then ignoring correction for adult equivalents will tend to push up our measure of the relative improvement of urban areas, that is, overstating urban improvement. The

denominator has fewer children in it (who should have been discounted in the first place).

As regards the market value of subsidized goods, this varies substantially between states and has changed over time. Also, if the relative incidence of subsidies differs between urban and rural areas, changes in the unmeasured subsidy will underestimate the value of consumption more in one than the other. In Tamil Nadu, there are free rice, free pots and pans, and free cooking fuel. The subsidies have increased substantially and differentially affect urban and rural areas. Other states are less generous, and the relative value of the subsidized goods has varied over time, meaning that state effects will not correct for such differences. Measures of consumption evaluated at PDS prices, say, will understate the growth of "real" consumption at market prices in rural areas relative to urban areas.

Housing is really complicated. If everybody paid rent, this would not be a problem. But formal rent payments are not that common, particularly in rural areas. And in urban areas, home ownership with no mortgage payments is missed entirely. Tenants could look much richer than their landlords because we measure the rent but not the quality of housing services for owners even if they have identical floors of a multi-family structure. So we capture some people's housing services and not others. In recent fieldwork, I have noticed that one way of saving or smoothing consumption over time is to add on to your house, in which case the adding on to your house looks like a consumer durable. If you did it this past year, we got you, but if you put in a *pucca* floor a year and a half ago, we miss you entirely. This is really a big problem. In the National Family Health Surveys between 1992 and 2006, households with *pucca* floors went from 15 percent to 85 percent in rural areas. Nothing like that happened in urban areas. It is possible that NSS consumption understates the growth in the true value of housing and, therefore, understates rural/urban convergence.

On the other hand, there has been substantial urbanization over the period. If that is accompanied by a big increase in the number of renters, this will make urban areas look like (i.e., in terms of measured consumption by NSS) they are getting richer faster than they really are in terms of changes in housing services. If there is a big increase in owners, then the bias goes the other way, and the increase in consumption in urban areas will be measured as being slower than it is. So for a large proportion of people's overall budget, we have arguments going in either direction. Correcting this measurement issue could lead to either more or less convergence between the sectors over time.

In any case, this is a good and clear paper. I just wonder how our view of things might be biased by the nature of the surveys used.

General Discussion

Initiating the discussion, Kaushik Basu, the chair, pointed out that there could be a natural explanation for the increase in the urban–rural gap in India after 2004, as highlighted in the paper. He urged the authors to examine the data to assess if poorer segments in urban areas were staying back or returning to their rural homes to participate in MGNREGA, or whether the rural workers who could have migrated to urban areas were not doing so for the same reason. The results could explain if the rural areas were becoming better off financially, or if the urban averages were improving. Both outcomes are entirely possible, and it would be very interesting to dig into the data to explore the reasons for the seemingly paradoxical findings of urban versus rural growth in the paper.

Another broad issue he raised was that conventionally one thinks of urbanization as people moving into urban areas or even new cities coming up, whereas the data used in the paper hints at a novel conceptual category, that is, almost as if rural areas are morphing into urban areas. It would be interesting to study this concept in a broader framework if the data throws up a hint of such a development.

Dilip Mookherjee said that the findings in the paper were not very surprising, and that the authors should, perhaps, have asked a different set of questions. It is well known that the returns to education are much higher in urban than in rural areas for obvious reasons. Skilled labor is more complementary to capital, and there is more capital in urban sector jobs than in rural sector jobs. Irrespective of the factors contributing to endogenous growth, agglomeration and technical change is happening much more in urban areas. Hence, the results derived by the authors in the paper are to be expected. These results are a natural part of the development process. Instead, the question that needed greater focus in the paper is: Why is there so little structural transformation in India?

Devesh Kapur said that the authors should look more carefully at the Indian definition of "urban" since it is not the same as the standard UN definition. For instance, Census towns do not show up in the Indian categorization of urban areas. Hence, a very significant fraction of the so-called rural population could exhibit very urban-like characteristics, if it were not measured by the very peculiar Indian definition of what is urban and if India were to adopt the international definition of urban. Many parts of what we call peri-urban areas or areas that are actually legally rural have completely urban characteristics in many more ways, such as in terms of occupation.

He asked the authors to think much more about this definition since it could shape many of their results.

Sudipto Mundle suggested an explanation for the conundrum that consumption had not converged between urban and rural India. Convergence can be a very positive story or not so positive depending on the growth context. In the case of a situation of dynamic employment growth, convergence leads to a positive story or a classic Lewis-type model. However, in a situation where employment is not growing or, at least, quality jobs are not growing, which has been occurring in India in recent times, the convergence may actually point to simply a survival strategy.

Sonalde Desai noted that occupation codes had changed between 2004–05 and 2009–10, so that the big jump in the number of white-collar workers may have something to do with the way in which the codes had changed. She also noted that the occupation trends being examined by the authors, who were probably combining both male and female workers, did not take into account the recent significant transformation in women's work. There has been a substantial expansion of government-created white-collar jobs for women in rural areas, such as those for *Anganwadi* and health workers, and simultaneously, there has been a big decline in self-employment and agricultural work. The authors should consider producing separate employment figures for men and women.

Rinku Murgai asked the authors to speculate on why consumption was diverging when wages were converging. She also noted that using the NSS data, it is possible to create a series separating the small towns from the big cities. There is a big difference between the million-plus cities and rural areas but much less difference with the smaller towns: The small towns and villages are beginning to resemble each other.

The session video, the paper, and all presentations for this IPF session are hyperlinked on the IPF program available on the NCAER website by scanning this QR code or going to the following URL:
https://www.ncaer.org/IPF2019/Agenda/Agenda_IPF_2019.pdf