

# Gender and caste-based wage discrimination in India: some recent evidence

Tushar Agrawal

Published online: 29 November 2013  
© Institut für Arbeitsmarkt- und Berufsforschung 2013

**Abstract** This paper examines gender and social groups wage discrimination in India using a nationally representative survey. We examine the wage gaps between different sub-groups of population separately in the rural and urban sectors using the Blinder-Oaxaca decomposition method. To account for possibility of the sample selection bias, the Heckman correction model is employed. We find a large wage differential between gender groups and between different social groups. The decomposition analysis reveals that the wage differential between males and females can largely be attributed to discrimination in the labor market. Nevertheless, in case of social groups this gap is mostly driven by differences in endowments.

**Keywords** Gender · Social groups · Wage differential · Blinder-Oaxaca decomposition · India

**JEL Classification** I20 · I21 · J31

## Geschlecht und Kaste-ansässige Lohndiskriminierung in Indien: Einige Neue Beweise

**Zusammenfassung** Dieser Artikel untersucht geschlechts- und sozialgruppierungsspezifische Lohndiskriminierung in Indien basierend auf einer nationalrepräsentativen Umfrage. Unter Anwendung der Blinder-Oaxaca-Zerlegung werden Lohnunterschiede zwischen verschiedenen Bevölkerungsgruppen untersucht, aufgeschlüsselt nach städtischen und

ländlichen Gebieten. Darüber hinaus wird die Heckman-Korrektur angewandt. In diesem Artikel zeigt sich, dass große Lohnunterschiede sowohl zwischen Geschlechtern als auch zwischen verschiedenen Sozialgruppierungen auftreten. Während die Lohnunterschiede zwischen Männern und Frauen größtenteils auf Arbeitsmarktdiskriminierung zurückzuführen sind, entstehen diese Unterschiede bei den untersuchten Sozialgruppierungen durch gruppenspezifische Merkmale, die variierende Preise auf dem Arbeitsmarkt erzielen.

## 1 Introduction

Gender and social discrimination are key issues in the Indian labor market. Despite various efforts by the government such as reservation policies in educational institutions and in government and public sector employment, educational and occupational outcomes of females and disadvantaged social groups such as Scheduled Castes (SCs) and Scheduled Tribes (STs) are unsatisfactory.<sup>1</sup> These groups are associated with low paid, dead-end and insecure employment. Poverty rates are higher and educational attainment are lower among the SCs and the STs (Agrawal 2013a; Das and Dutta 2007; Deshpande 2001; Gaiha et al. 2007). These disadvantaged groups face difficulties in finding regular employment since they face higher transaction costs associated with entry into the labor market (Ito 2009). The share of non-SCs/STs in regular employment is substantially higher than that for SCs/STs (Singh et al. 2013).

Caste is believed to be a defining characteristic in the Indian society. It can be characterized as a hierarchical, hegemonic ranking of social groups found predominantly on the

**Electronic supplementary material** The online version of this article (doi:10.1007/s12651-013-0152-z) contains supplementary material, which is available to authorized users.

T. Agrawal (✉)  
National Council of Applied Economic Research (NCAER),  
Parisila Bhawan, 11 Indraprastha Estate, New Delhi 110002, India  
e-mail: ag.tushar@gmail.com

<sup>1</sup>The Indian constitution provides for affirmative action policies for scheduled castes and scheduled tribes.

Indian subcontinent (Kannabiran 2012: 190). “Caste is undoubtedly an all-India phenomenon in the sense that there are everywhere hereditary, endogamous groups which form a hierarchy, and that each of these groups has a traditional association with one or two occupations” (Srinivas 1966: 3). There are mainly two models of the caste system: ‘varna’ and ‘jati’ (Srinivas 1984). *Varna* is the Vedic classification of the four ranked occupational orders: Brahmin (priest), Kshatriya (warrior), Vaishya (trader and agriculturist), and Shudra (menial). *Jati* is a purely local system of ranked, hereditary, and endogamous groups, each associated with one or more traditional occupations. While the number of *varnas* is invariant, the number of interdependent *jatīs* can vary from a minimum of 10 to about 30 even in a small region. The term *jati* can be applied to a group of castes, a caste, a subcaste, or a subdivision of a subcaste (Beteille 1967). Tribes have generally remained outside the *varna* system. These groups have some similarities with *jatīs* in the sense that almost all tribals are endogamous. These tribe communities have a belief of hierarchy based on descent, ecology, ranking and status (Deshpande 2011: 20–21).

Caste is considered as the most pervasive parameter differentiating the Indian society. The caste system is the primary source of social stratification (Deshpande 2011). Since historic times, caste community has been considered as a vital socio-economic institution (Beteille 2007). Though caste is determined on the basis of birth, it results in determining the distribution of human and physical capital, occupation, power and social status (Dhesi 1996). It is believed to be responsible for major inequalities in access to broad areas such as education, health, jobs and technology (Das and Dutta 2007). The presence of discrimination may result in distorting households’ decisions related to labor allocation and human capital investment (Ito 2009). India has remarkably low levels of occupational and spatial mobility. Caste-based labor market networks have locked entire groups of individuals into narrow occupational categories for generations. Historical occupation patterns induced by the caste-based networks carry on shaping occupational choice and therefore schooling choice for the boys (Munshi and Rosenzweig 2006, 2009). Both SCs and STs have low educational attainment (years of schooling) as well as high inequality in educational attainment (Agrawal 2013a). For instance, the average attainment in 2004–05 for STs, SCs and the remaining population was 2.76, 3.39 and 5.29 years, respectively. SCs/STs habitations are, on average, considerably smaller than of those from the rest. Schools situated in SCs/STs habitations have very few teachers than those in upper caste habitations, and they are more likely to be characterized by ‘multi-grade teaching’ (Kochar 2007).

This paper analyzes wage discrimination between gender and social (caste) groups in India using a set of simple models of wage determination. The term ‘caste’ in this

paper refers to four broad caste divisions or social groups. These are SCs, STs, Other Backward Classes (OBCs), and ‘Others’. The last group includes those not covered in the first three. We also examine the gender differences in Non-Scheduled Castes/Tribes (NSC), and Scheduled Castes and Tribes (SC/ST). Most studies examine wage discrimination either between gender groups or between social groups in the rural or urban sectors. In this paper, we examine wage discrimination separately for the gender as well as for social groups in both the rural and urban sectors. The labor market conditions in India differ very much across the two sectors. For example, in rural India a large workforce is engaged in agricultural and allied activities. A separate examination of the rural and urban sectors may be helpful in identifying some of the important factors relating to gender and castes, and may provide inputs to rural and urban development policies. The analysis in the paper is based on a nationally representative dataset, namely, India Human Development Survey (IHDS) 2005.

The paper is organized as follows. The next section provides a brief overview of literature on wage discrimination in India. Section three gives a general idea of economic models of discrimination. Section four presents methodology and outlines the decomposition method. Section five gives a brief note on the database. Section six discusses the results on decomposition. Section seven provides plausible explanations for our results based on the existing literature. The final section concludes and offers some policy implications.

## 2 Literature review

There have been a number of empirical studies on wage discrimination for both industrialized and developing economies. In India too, there are various studies on labor market and some of them attempt to examine labor market discrimination. Most studies on wage discrimination in the Indian labor market are based on data from different rounds of surveys conducted by the National Sample Survey Organization (NSSO). Divakaran (1996), Deshpande and Deshpande (1997), Kingdon (1997 and 1998), Kingdon and Unni (2001), and Bhaumik and Chakrabarty (2008) examine gender-based discrimination. Banerjee and Knight (1985) and Madheswaran and Attewell (2007) examine caste-based discrimination for urban labor market; Gaiha et al. (2007) for rural India; Ito (2009) for the rural sector of North India; and Kijima (2006) and Das and Dutta (2007) for all India. Deshpande (2001) examines regional variations in inter-caste disparity using a caste deprivation index. Thorat and Attewell (2007) examine job discrimination in private sector enterprises using a field experiment. Deshpande (2011), and Thorat and Newman (2010) provide a comprehensive study on caste system and discrimination in India.

A study by Kingdon (1998) suggests that there exists pro-male bias in investment in education by parents. Females face poor economic incentives to investment in education than males since they reap lower labor market returns to education than males. The study also finds a substantial bias in the estimates of the rates of return to education if family background is ignored. In another study, Kingdon and Unni (2001) find that women experience high levels of wage discrimination in the Indian urban labor market. However, they find that education contributes little to this discrimination. A study by Bhaumik and Chakrabarty (2008) using two rounds of the NSSO's employment-unemployment survey finds that the gender wage gap narrowed considerably between years 1987 and 1999. The narrowing of the earnings gap was attributed largely to a rapid increase in the returns to the labor market experience of women.

As regards to caste-based discrimination, Madheswaran and Attewell (2007) find that SC/ST workers are discriminated against both in the public and private sectors. However, the discrimination effect is more pronounced in the private sector. The study uses the standard Blinder-Oaxaca decomposition method and data from the NSSO's employment and unemployment surveys. Kijima (2006) offers a comparative analysis of disparity in living standards (measured using monthly per capita expenditure) among SC, ST and non-ST/SC households. The study uses the Neumark's wage decomposition method and the data from four rounds of the NSSO's survey (1983, 1987, 1993 and 1999). The findings from the study show that about half of the welfare disparities between the SCs and the non-SCs/STs and two-thirds of the disparities between the STs and the non-SCs/STs are due to differences in characteristics. The characteristic disparities between the STs and the non-SCs/STs are largely due to education and location differences. Further, there was little change in the contribution of the structural (returns) component between the SCs/STs and the non-SCs/STs during the period 1983 to 1999.

A study by Das and Dutta (2007) suggests that caste is still an influential factor in deciding how individuals are remunerated in the wage labor market. The study is based on data from 61st round (2004–05) of the NSSO's survey. They find that the extent of the wage gap is significant (about 0.37 log points) in case of regular workers, and one-third of the gap is attributable to unequal treatment of scheduled caste workers relative to general caste workers. Further, the wage gap among casual workers is very low and mainly accounted for by differences in characteristics. Thorat and Attewell (2007) examine the prevalence of discrimination in the job application process of private sector enterprises using a field experiment. They find that companies discriminate by caste and religion while hiring job applicants with equal qualifications.

In a recent study, Motiram and Singh (2012) examine intergenerational occupational mobility for different caste

groups using transition matrices and the data from the IHDS 2005. Their findings suggest considerable rigidity in class positions, particularly for the lower classes. The study finds that mobility is higher in urban areas as compared to rural areas since there is considerable occupational diversity in urban areas. The persistence in high status occupations (viz., professionals, officials and related) is lower for the SCs/STs than for non-SCs/STs. On the contrary, persistence in occupations that are at the bottom of the occupational hierarchy (viz., such as agricultural laborers and elementary occupations) is much higher for the SCs/STs than for non-SCs/STs. They also document considerable downward mobility of the SCs/STs and show that this is higher for SCs/STs than that for non-SCs/STs.<sup>2</sup>

### 3 Economic models of discrimination

Economic models of discrimination can broadly be classified into two classes: (i) competitive models in which agents act individually, and (ii) collective models in which groups act collectively against each other (Altonji and Blank 1999: 3168). In general, economic analysis has focused on competitive models which underline two broad types of discrimination. The first is "taste-based discrimination" and the second, "statistical discrimination".

Becker (1957) in his pioneering work introduced the first economic model of discrimination which explains discrimination by a "taste for discrimination". What it implies is that discriminators are willing to forgo an extra pay in order to have the group they preferred compared to an equally productive unfavorable group. Becker identified three distinct forms of discrimination in a competitive framework: employer, employee (or coworker) and customer. Some employers hold a taste for discrimination, they think that women (or minority group) are less productive.<sup>3</sup> The discriminating employer in this case is unwilling to hire them unless women or minority workers themselves 'compensate' employers by accepting a lower wage (a wage below the wage paid to men) for identical productivity or by being more productive at a given wage. Employee discrimination exists when employers discriminate in a situation when other employees do not prefer to work with colleague of the opposite sex or minority workers. Consumer discrimination occurs when consumers are not willing to purchase goods and service (for example, stores) from members of a certain group and instead prefer to pay a higher price. In this case employers discriminate not because of their own tastes

<sup>2</sup>In the intergenerational context, downward mobility means children moving to a lower socio-economic position compared to their parents.

<sup>3</sup>It is worth mentionable that Becker's discussion generally referred to race (blacks and whites) discrimination.

but because of their customers' tastes. One of the important results emerging from Becker's work is that in the long run discrimination by employers should decline since it is not economically efficient. His theory suggests that competition will increase the relative demand for minority workers only in the economies where employers have a taste for discrimination.

The statistical discrimination models have been advanced by Arrow (1972), Phelps (1972), and Aigner and Cain (1977). Contrary to the 'taste-based' discrimination, these models assume no prejudice by employers. Rather women (and minority group) earn lower wages because these groups on average have lower productivity. Statistical models are based on the assumption that firms have limited information about the skills of applicants. Employers make decision in the absence of full information, since it is costly and difficult to acquire full information on job applicants. Thus, the firms use observable characteristics (such as gender or race) as a signal of productivity provided that these characteristics are correlated with productivity. In other words, if employers believe that women (or minority group) are on average less productive than men, they use gender or race as a screening device to which an individual belongs while hiring or pay decisions. In turn, they pay minority group a lower wage rate than majority group with identical observable skills. The literature on statistical discrimination has two main strands (Altonji and Blank 1999: 3180–3181). The first strand of literature looks how prior beliefs about the productivity of group members can influence hiring and pay decisions. This issue is addressed by Arrow (1973). The second strand of literature analyzes the consequences of group differences in the precision of the information that employers have about individual productivity. This issue is addressed by Aigner and Cain (1977).

## 4 Methodology

### 4.1 General

The unit of analysis for examining economic discrimination could be: (a) the household or family where the disparity between two subgroups is usually measured by income; or (b) the individual worker where the appropriate unit for examining disparities is wage rates or earnings (Cain 1986). The focus of this paper is on the second type of analysis since we are interested in examining labor market discrimination. There are two approaches that are mostly applied for examining wage discrimination between population sub-groups (gender or caste/race) (Darity and Mason 1998). The first involves estimation of a standard 'Mincer' type earnings/wage function (Mincer 1974) in a single equation form presuming a similar wage structure for both the

groups (male and female). Thus the coefficients of all the explanatory variables remain the same for both males and females and difference between the two groups is captured through a gender dummy. A negative and statistically significant coefficient on the gender dummy (if male is the reference category) can be interpreted as an evidence of discrimination against women. The second approach involves decomposition of the wage function by estimating two wage regressions for the two groups. The latter allows explanatory variables to vary across the groups and decomposes the wage gap into two components: a portion explained by the average group differences in productivity characteristics (endowments) and an unexplained portion which is attributable to labor market discrimination. We apply the conventional Blinder-Oaxaca decomposition method to decompose the average wage gap.

### 4.2 Wage decomposition

The Blinder-Oaxaca decomposition method (Blinder 1973; Oaxaca 1973) can be explained as follows. Assume that there are two groups  $A$  (advantaged) and  $B$  (disadvantaged). The wage equation for each group can be written as:

$$W_A = \beta_A X_A + \varepsilon_A \quad (1)$$

$$W_B = \beta_B X_B + \varepsilon_B \quad (2)$$

where  $W$  is the natural logarithm of wages (hourly),  $X$  is a vector of explanatory variables (productivity characteristics), and  $\varepsilon$  is a random error term. The vector of explanatory variables controls for human capital, demographic and various other characteristics of individuals.

The OLS estimation of the above equations can be expressed as:

$$\bar{W}_A = \hat{\beta}_A \bar{X}_A \quad (3)$$

$$\bar{W}_B = \hat{\beta}_B \bar{X}_B \quad (4)$$

The bars denote mean values of the variables and the hats denote OLS estimates. The difference in average wage rates between the two groups can be expressed as:

$$\bar{W}_A - \bar{W}_B = \hat{\beta}_A \bar{X}_A - \hat{\beta}_B \bar{X}_B \quad (5)$$

If the disadvantaged group is given the same wage structure as the advantaged group (in the absence of wage discrimination) then their average wage would be given as:

$$\bar{W}_{B^*} = \hat{\beta}_A \bar{X}_B \quad (6)$$

By (5) and (6), the overall average wage gap between the advantaged and disadvantaged groups can be expressed as:

$$\bar{W}_A - \bar{W}_B = \hat{\beta}_A (\bar{X}_A - \bar{X}_B) + (\hat{\beta}_A - \hat{\beta}_B) \bar{X}_B \quad (7)$$

The first term on the right-hand side of the above equation is generally referred to as the 'endowment' effect or

‘explained’ component which is the part of the wage differential due to average differences in productivity characteristics of individuals of group *A* and group *B*. The second term is referred to as the ‘treatment’ or ‘unexplained’ (residual) component. The unexplained component shows difference in the returns to similar characteristics (the contribution of differences in the coefficients) between groups *A* and *B*. This component is often interpreted as a measure of wage discrimination. However, the second term also captures the effects of differences in unobserved variables (omitted explanatory variables).

The Blinder-Oaxaca decomposition could suffer from ‘index number problem’, i.e., choice of the non-discriminatory wage structure. The above decomposition is carried out from group *A*’s (say male) perspective, i.e., we assume that without discrimination the male wage structure applied for males and females. To account for the index number problem, decomposition methods based on other reference wage structures (female wages or using weighted combination of male and female wages) are used.<sup>4</sup> Therefore, to understand sensitivity of results, we use the wage structure of both the groups. The decomposition equation using the reference wage structure of the second group can be written as:

$$\bar{W}_A - \bar{W}_B = \hat{\beta}_B(\bar{X}_A - \bar{X}_B) + (\hat{\beta}_A - \hat{\beta}_B)\bar{X}_A \tag{8}$$

It may be noted that the above alternative decomposition can produce quite different results than that from (7).

### 4.3 Heckman sample selection model

The estimation of the wage equation may suffer from problem of sample selection bias in the presence of non-randomly selected samples because of which the ordinary least squares (OLS) estimates of the wage equation may be

<sup>4</sup>There are various other decomposition methods which are similar to the Blinder-Oaxaca method but use different alternative wage structure. To understand some of them, consider  $\beta^*$  as a non-discriminatory reference wage structure given by:  $\beta^* = \Omega\beta_A + (I - \Omega)\beta_B$  where  $\Omega$  is a weighting matrix and  $I$ , an identity matrix. As discussed, Oaxaca (1973) proposes the use of either the current male wage structure ( $\Omega = I$ ) or the current female wage structure ( $\Omega = 0$ ). Reimers (1983) proposes the weighting matrix  $\Omega = 0.5I$ . Cotton (1988) chooses the weighting matrix  $\Omega = 0.5I_wI$ , where  $I_w$  is the fraction of the sample made up by the majority group, since the non-discriminatory structure should be similar to the structure that holds for the larger group. However, neither of these solutions is completely satisfactory as each chooses the weight in an arbitrary manner (Oaxaca and Ransom 1994). Neumark (1988) uses a least squares criterion to estimate  $\beta^*$  from the pooled sample of males and females, i.e.,  $\beta^* = (X'X)^{-1}(X'Y)$  where  $X'X = X'_AX_A + X'_BX_B$ . Oaxaca and Ransom (1994) propose adoption of the weighting matrix  $\Omega = (X'X)^{-1}(X'_AX_A)$  where both  $X'X$  and  $X'_AX_A$  are positive definite matrices.  $X$ ,  $X_A$  and  $X_B$  are the observed matrices for the pooled, group *A* and group *B* sample, respectively. However, it is beyond the scope of this paper to use all the alternative wage structure.

biased and inconsistent. A typical example is the women component in the labor supply. To account for the sample selection bias we employ the Heckman correction method (Heckman 1979). The method involves estimation in two stages: the first step is to estimate a selection equation which determines the probability of labor force participation. After the estimation of the selection equation we get the inverse Mills ratio (selectivity term). The inverse Mills ratio ( $\lambda$ ) is defined as the ratio of the probability density function to the cumulative distribution function of a distribution ( $\hat{\lambda}_i = \frac{\phi(z_i\hat{\varphi})}{\Phi(z_i\hat{\varphi})}$ , where  $z$  is a vector of control variables used in the selection equation, and  $i = 1, \dots, n$  indicates the number of individuals). The second step involves estimating a wage equation by OLS using the inverse Mills ratio as an additional regressor. The selectivity term added to the wage equation measures the sample selection effect due to the lack of observations on the earnings of non-participants (Dolton and Makepeace 1986).<sup>5</sup>

To address the sample selection problem appropriately, it is important to find some identifying variables (exclusion restrictions) that affect the selection equation but can be excluded from the wage equation. In literature, non-labor income (Duraismy 2002; Asadullah 2006), household size (Dutta 2006), number of dependents children in a household (Dutta 2006; Kingdon and Theopold 2008), number of elderly aged 65 and above (Dutta 2006; Kingdon and Theopold 2008), and land ownership (Asadullah 2006; Kingdon and Theopold 2008) have been used as potential exclusion restrictions. Households with land ownership and non-labor income are less likely to attach with wage employment.

After considering the sample selection, the selectivity adjusted wage decomposition equation can be written as follows:

$$\begin{aligned} &(\bar{W}_A - \bar{W}_B) + (\hat{\theta}_B\bar{\lambda}_B - \hat{\theta}_A\bar{\lambda}_A) \\ &= \hat{\beta}_A(\bar{X}_A - \bar{X}_B) + (\hat{\beta}_A - \hat{\beta}_B)\bar{X}_B \end{aligned} \tag{9}$$

where  $\theta$  denotes the covariance between the errors in the selection and wage equations (in each group *A* and *B*).

<sup>5</sup>In the wage regression analysis, adjustment of sample-selection bias using the Heckman two step procedure has become standard practice. However, this approach has been criticized and not always advisable (see Puhani 2000 for discussion). Puhani (2000: 65) concludes that “the cases where the need to correct for selectivity bias are largest are those with a high correlation between the error terms of the selection and the outcome equation, and those with a high degree of censoring. Unfortunately, though, as the Monte Carlo analyses show, in exactly those cases Heckman’s estimator is particularly inefficient and sub-sample OLS may therefore be more robust”. In this paper, we therefore report results from both the OLS and the Heckman estimations.

## 5 Data and variables

### 5.1 Data

This paper draws the data from the India Human Development Survey (IHDS) 2005. The survey was conducted during 2004–05. This dataset is produced by the National Council of Applied Economic Research (NCAER), New Delhi, and the University of Maryland. The IHDS is a nationally representative survey of 41,554 households in 1503 villages and 971 urban neighborhoods across India. These households include 215,754 individuals. The IHDS was conducted in 33 states and Union Territories of India except the Andaman and Nicobar Islands and the Lakshadweep. These states cover 384 districts, 1503 villages and 971 urban blocks located in 276 towns and cities. Villages and urban blocks form the primary sampling unit (PSU) from which the households have been selected. Urban and rural PSUs are selected using a different design (Desai et al. 2010).

The survey has a rich amount of information on household and individual levels characteristics. The household characteristics include information on household residence (rural or urban sector, and state of residence), household size, membership of a social group and religion. The individual characteristics include information on age, educational attainment, gender, occupation, marital status and relation to household head. The survey also has information on wages and salaries, household income, components of household income: farm income, income from interests, dividend or capital gains; property, pension, income from other sources etc.

### 5.2 Variables

#### 5.2.1 Dependent variables

The dependent variable selected for the wage equation is the natural logarithm of hourly wage.<sup>6</sup> For the sake of robustness, the wage distribution is trimmed by 0.1 percent at the top and bottom tails of the distribution.

If an individual works for more than or equal to 240 hours in a year, he/she is considered as part of the workforce.<sup>7</sup> The dependent variable of probit (selection) equation takes a value 1 if an individual is part of the workforce, and 0 otherwise.

<sup>6</sup>In the IHDS survey, for each member of the household the following questions were asked. For how many days did individual do work last year? How many hours did individual work in a usual day? How much was individual paid in cash for that work? The latter information could be expressed on a daily, monthly, or annual basis and accordingly it is also converted on a monthly basis.

<sup>7</sup>This criterion is based on work participation measure used in the IHDS data.

#### 5.2.2 Independent variables

*Human-capital variables* An individual is classified under one of the following six educational levels: illiterate (also includes below primary), primary, middle, secondary, higher secondary, and graduate. It is assumed that an individual spends 0, 5, 3, 2, 2 and 3 additional years, respectively in these educational levels. Experience is captured through potential experience which is a proxy for the actual experience. The potential experience is defined as age minus numbers of year of schooling minus five.<sup>8</sup>

*Demographic variables* Demographic controls are sex, place of residence (rural or urban sector), membership of a social group, marital status, and state control. All these are used as dummy variables. Each household belongs to one of the following social groups: Scheduled Castes (SCs), Scheduled Tribes (STs), Other Backward Classes (OBCs) and 'Others'. As discussed, SCs and STs are two historically disadvantaged groups in India and while examining gender differences in social groups we regroup them into two groups: NSC (OBCs and Others), and SC/ST (SCs and STs). State dummies are also included to capture the state-specific variations.

*Family background* Family background is captured through household head's education.<sup>9</sup> Education of household head is classified into the following five categories: illiterate (and below primary), primary, middle, secondary (secondary and higher secondary), and graduate.

*Exclusion restrictions* The identifying variables included in the selection equation for correcting sample selection are household size and the number of children (aged less than 15) in a household. Both the variables are continuous variables. In addition to these, we use household land possession for the rural sector, and non-labor income of an individual or household for the urban sector as identifying variables. Both these variables are being used as a dummy variable.<sup>10</sup> The non-labor income includes income from renting property and/or income from interest, dividends, or capital gains. We do not use the variable household land possession for the urban sector because as suggested by Dutta (2006) it is not a reasonable restriction in the urban context. We

<sup>8</sup>We assume that an individual starts schooling at the age of five and starts working immediately after that (Agrawal 2012; Duraisamy 2002).

<sup>9</sup>In case the household head is an observation, his/her father's education is used.

<sup>10</sup>We group both land possession and non-labor income into four categories. We do not use them as a continuous variable since in the data we observe that quite a high proportion of households do not possess agriculture land or do not have non-labor income.

expect negative signs on the household size and non-labor income since individuals living in larger households and/or with non-labor income are less likely to enter in wage employment whereas a positive sign on the number of children in a household since individuals in a household with more number of dependents (children) are more likely to seek wage work (Agrawal 2012).

The study is restricted for the working age population aged 15 to 65. A separate analysis for the rural and urban sectors is carried out since labor market conditions in India differ across the two sectors. Appendix 1 (online) provides description of all the variables used in the analysis.

## 6 Results

### 6.1 Descriptive statistics

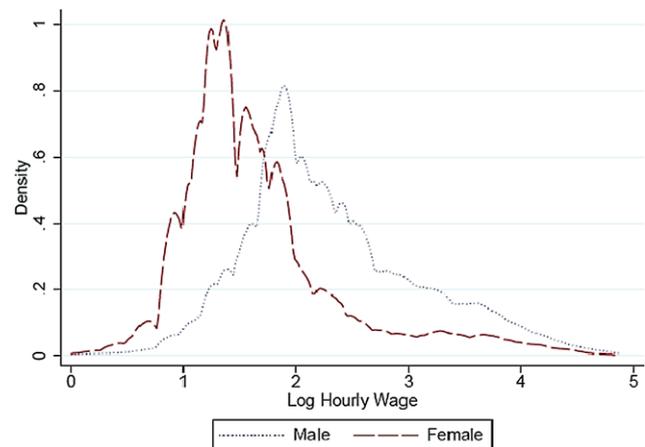
Table 1 gives the mean and the standard deviation of the variables used in the Heckman estimation. The table shows that the people residing in the rural sector earn about one-half the wages on an hourly basis than those in the urban sector. Educational differences in the two sectors are also large. Individuals in the rural sector have low educational attainment and a substantial proportion does not even have formal schooling. Whereas in the urban sector individuals have high educational levels; particularly, the proportion of the individuals with graduation is quite high in the urban sector than that in the rural sector. The table also shows large differences in educational attainment between NSC and SC/ST. For instance, more than 50 percent population of SC/ST is illiterate or has below primary education in the rural sector compared to 36 percent in the urban sector.

Figure 1 presents kernel density estimates of the distributions of log hourly wages for males and females in the age group 15–65 years. The wage distribution for female population is skewed towards the lower tail. Among the social groups, except for the ‘Others’, the wage distributions for the remaining groups are skewed towards the lower tail (Fig. 2). While the distribution for ‘Others’ group nearly approximates the normal distribution, the distribution for SCs has a long peak, and the distribution of OBCs follows that of the SCs but with less peakedness. The distribution of STs is bimodal with two long peaks.

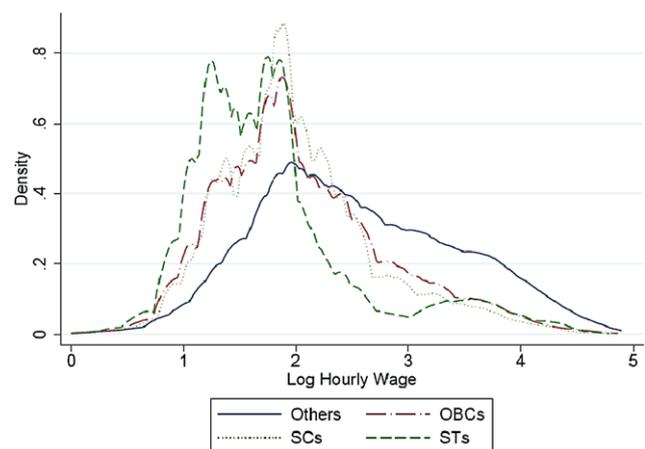
### 6.2 Male-female wage discrimination

Table 2 shows decomposition of the augmented earnings function for the gender groups using the Blinder–Oaxaca method.<sup>11</sup> The table reports selectivity-corrected estimates

<sup>11</sup>In the results section, we only report the decomposition results. The detailed regression results on which the decomposition results are



**Fig. 1** Wage distribution by gender groups



**Fig. 2** Wage distribution by social groups

separately for the rural and urban sectors.<sup>12</sup> The block 1 of the table shows the mean log hourly wages for the two groups. The mean log hourly wage for men is 2.09 and 2.75 in rural and urban populations, respectively. The corresponding values for women are 1.41 and 2.21. Thus, there is a wage gap of 0.68 (log points) in the rural population and 0.54 in the urban population between males and females groups.

We notice that the raw gender wage gap is more pronounced in the rural sector than that in the urban. In the rural sector, women are mainly employed in agriculture and related occupations. Though the wages in these occupations are low for both men and women in general, but women are more likely to work with much lower wages than men. On the other side, in the urban sector women are employed in other occupations such as professional and

based are provided in online Appendix 2 for the rural sector and in online Appendix 3 for the urban sector.

<sup>12</sup>The OLS estimates are provided in online Appendix 4.

**Table 1** Means and standard deviations of variables used in Heckman estimation

Variable	Rural			Urban		
	All	NSC	SC/ST	All	NSC	SC/ST
Log Hourly Wage	1.907 (0.678)	1.981 (0.712)	1.813 (0.621)	2.653 (0.854)	2.695 (0.859)	2.520 (0.825)
Work participation	0.542	0.484	0.641	0.362	0.346	0.421
<i>Educational level:</i>						
Illiterate & below primary	0.416	0.358	0.515	0.199	0.177	0.284
Primary	0.140	0.141	0.139	0.108	0.103	0.127
Middle	0.157	0.165	0.142	0.161	0.158	0.173
Secondary	0.175	0.200	0.133	0.243	0.248	0.220
Higher secondary	0.069	0.082	0.048	0.132	0.138	0.107
Graduate	0.043	0.054	0.023	0.158	0.176	0.089
Experience	23.457 (16.243)	22.758 (16.375)	24.648 (15.945)	20.784 (15.362)	20.697 (15.355)	21.120 (15.386)
Experience squared	814.045 (924.687)	786.054 (924.335)	861.754 (923.354)	667.966 (806.630)	664.148 (802.113)	682.743 (823.759)
Sex (Dummy, Ref. -Male)	0.484	0.497	0.461	0.532	0.538	0.509
Marital Status (Dummy, Ref.-Unmarried)	0.715	0.697	0.744	0.685	0.686	0.681
<i>Social group:</i>						
Others	0.240	–	–	0.419	–	–
OBC	0.390	–	–	0.376	–	–
SC	0.252	–	–	0.170	–	–
ST	0.118	–	–	0.035	–	–
<i>Household head's education:</i>						
Illiterate & below primary	0.550	0.483	0.665	0.271	0.239	0.395
Primary	0.180	0.195	0.154	0.154	0.152	0.162
Middle	0.109	0.124	0.085	0.154	0.157	0.143
Secondary and higher secondary	0.135	0.165	0.083	0.286	0.300	0.230
Graduate	0.026	0.033	0.014	0.135	0.152	0.070
Household size	6.216 (3.129)	6.360 (3.282)	5.971 (2.832)	5.597 (2.577)	5.587 (2.650)	5.636 (2.270)
No. of children	1.876 (1.787)	1.868 (1.830)	1.891 (1.711)	1.379 (1.463)	1.355 (1.471)	1.472 (1.427)
<i>Agriculture land (in acre):</i>						
No land	0.484	0.434	0.568	–	–	–
>0 & ≤4 acre	0.399	0.420	0.363	–	–	–
>4 & ≤15 acre	0.099	0.121	0.061	–	–	–
>15 acre	0.018	0.025	0.007	–	–	–
<i>Non-labor income (in rupees):</i>						
No non-labor income	–	–	–	0.918	0.911	0.945
>0 & ≤2500 rupees	–	–	–	0.018	0.018	0.014
>2500 & ≤8000 rupees	–	–	–	0.019	0.020	0.013
>8000 rupees	–	–	–	0.046	0.051	0.027

Notes: The sample consists of individuals aged 15–65 in the IHDS (2005) data. Standard deviations in parentheses and is not reported for dummy variables. Please see online Appendix 1 for description of variables used in the analysis. 'NSC' denotes non-scheduled castes group (other backward classes and others)

**Table 2** Decomposition of gender wage differential: Blinder-Oaxaca decomposition with sample selection

Components	Rural			Urban		
	ALL	NSC	SC/ST	ALL	NSC	SC/ST
Prediction (Male)	2.091 (0.004)	2.166 (0.006)	1.986 (0.006)	2.748 (0.007)	2.783 (0.008)	2.631 (0.015)
Prediction (Female)	1.411 (0.020)	1.406 (0.030)	1.412 (0.028)	2.207 (0.138)	2.324 (0.189)	2.083 (0.169)
Raw differential	0.680 (0.021)	0.760 (0.030)	0.574 (0.028)	0.541 (0.139)	0.459 (0.190)	0.547 (0.170)
<i>Using males wage structure:</i>						
Explained differential	0.226 (0.006)	0.244 (0.008)	0.179 (0.008)	0.100 (0.013)	0.080 (0.015)	0.112 (0.024)
Unexplained differential	0.454 (0.021)	0.516 (0.031)	0.394 (0.029)	0.441 (0.138)	0.379 (0.189)	0.436 (0.169)
Proportion of explained differential	0.332	0.321	0.312	0.185	0.174	0.204
Proportion of unexplained differential	0.668	0.679	0.688	0.815	0.826	0.796
<i>Using females wage structure:</i>						
Explained differential	0.266 (0.014)	0.289 (0.020)	0.196 (0.019)	0.082 (0.021)	0.068 (0.023)	0.064 (0.050)
Unexplained differential	0.414 (0.017)	0.471 (0.025)	0.372 (0.023)	0.459 (0.133)	0.391 (0.186)	0.483 (0.152)
Proportion of explained differential	0.391	0.380	0.342	0.151	0.148	0.118
Proportion of unexplained differential	0.609	0.620	0.658	0.849	0.852	0.882
Observations	31720	17842	13878	15063	11461	3602

Notes: 'NSC' denotes non-scheduled castes group (other backward classes and others), and 'SC/ST' denotes scheduled castes group (scheduled castes and scheduled tribes). Standard errors are in parentheses

clerical and the wages between men and women remain comparable. Further, one can expect more bargaining power of women in the urban sector than that in the rural sector.

In the block 2, the raw wage differential is divided into two components: the differences in endowments and the differences in coefficients, using the wage structure of males as the reference wage structure. We find that the discrimination component (coefficients) is larger than the endowment component in both the rural and urban sectors. This suggests that the wage difference between males and females is largely because of discrimination in the labor market. The differences in endowments account for 33 and 19 percent wage gap in the rural and urban populations, respectively whereas discrimination explains 67 and 81 percent of the total wage differential. The block 3 of the table shows the decomposition results using the females wage structure as the reference wage structure. The proportion of wage differential in the rural sector due to endowments is 39 percent and the remaining 61 percent can be attributed to discrimination whereas in the urban sector these figures are 15 and 85 percent, respectively.

An intriguing observation is that the proportion of discrimination component is higher in the urban sector than that in the rural sector. However, this does not necessarily

indicate that women in the urban sector suffer more discrimination compared to those in the rural. In the rural sector, human capital differences between men and women are much larger as compared to the same in the urban sector. For instance, we find that the proportions of graduate men and women in the rural sector are 4.7 and 1.5 percent respectively whereas in the urban sector the same proportions are 19.4 and 19.3 percent.<sup>13</sup> It has also been documented that the distribution of education is more unequal in rural areas of the country as compared to that in urban areas (Agrawal 2013b). However, the entire unexplained component cannot be considered as discriminatory. This is because the data does not adequately capture some human capital differences between men and women. Many women remain out of the labor force because of various reasons such as household and child bearing activities. Even if they are employed, they typically work fewer hours per week than men (Kingdon 1998).

It is worthwhile to mention that findings from a meta-analysis on the gender wage gap by Weichselbaumer and Winter-Ebmer (2005) suggest that the raw gender wage differentials worldwide have fallen from 65 percent to 30 per-

<sup>13</sup>Beside this, a large proportion of women in the rural sector (in our sample of individuals aged between 15–65 years) either does not have schooling or has education up to below primary level.

cent between the 1960s to the 1990s. Most of this decrease is attributed to better labor market endowments of females. The above figures from the meta-analysis however suggest that the incidence of the raw gender wage gap is much higher in India. A study by Blau and Lawrence (2003) using micro-data from the International Social Survey Programme for 22 countries over the 1985–94 period finds that countries with more compressed male wage structures (a narrower male earnings distribution) and lower female net supply are associated with a lower gender pay gap.

### 6.3 Gender wage discrimination across social groups

Next we examine the gender wage differential in NSC and SC/ST groups separately. The decomposition of the gender wage differential for both NSC and SC/ST is presented in the Table 2. Interestingly, we note that the gender wage differential in NSC group is higher than the same in SC/ST group in the rural sector but is smaller in the urban sector. Across the two sectors, the wage differential in each group (NSC or SC/ST) is higher in the rural sector than in the urban sector. A large part of the wage differential is explained by the discrimination component for both NSC and SC/ST in both the sectors. These findings remain consistent irrespective of the reference being the male wage structure or the female wage structure.

### 6.4 Wage discrimination between NSC and SC/ST

In Table 3, we present the decomposition of the wage differential between NSC and SC/ST.<sup>14</sup> We notice a wage gap of 0.21 in the rural sector and 0.31 in the urban. The wage gap is more pronounced in the urban sector than in the rural sector. Using the wage structure of NSC, we find that 62 percent of the total wage differential in the rural sector could be attributed to group differences in the characteristics and the remaining 38 percent could be attributed to discrimination. The same figures in the urban sector are 69 and 31, respectively. Therefore, discrimination between social groups is higher in the rural sector as compared to the same in urban sector. In a similar type of study, Madheswaran and Attewell (2007) using the NSSO's data of 1999–2000 find that endowment explains 79 percent and discrimination explains 21 percent of the lower wages of scheduled castes (SCs and STs) as compared to non-scheduled castes in the urban labor market. However, it is also important to note that the difference in endowments itself may be a result of past discrimination which cannot be measured directly.

<sup>14</sup>The OLS estimates are provided in online Appendix 5.

**Table 3** Decomposition of social groups wage differential: Blinder-Oaxaca decomposition with sample selection

Components	Rural	Urban
Prediction (NSC)	1.981 (0.005)	2.695 (0.008)
Prediction (SC/ST)	1.776 (0.012)	2.389 (0.041)
Raw differential	0.205 (0.013)	0.306 (0.042)
<i>Using NSC wage structure:</i>		
Explained differential	0.128 (0.006)	0.210 (0.012)
Unexplained differential	0.077 (0.012)	0.096 (0.041)
Proportion of explained differential	0.622	0.686
Proportion of unexplained differential	0.378	0.314
<i>Using SC/ST wage structure:</i>		
Explained differential	0.138 (0.008)	0.228 (0.017)
Unexplained differential	0.067 (0.013)	0.078 (0.043)
Proportion of explained differential	0.672	0.746
Proportion of unexplained differential	0.328	0.254
Observations	31709	15081

Notes: 'NSC' denotes non-scheduled castes group (other backward classes and others), and 'SC/ST' denotes scheduled castes group (scheduled castes and scheduled tribes). Standard errors are in parentheses

## 7 Discussion

Some interesting facts emerge from the preceding section. Why does discrimination explain a large part of the wage differential for gender groups and not for social groups? Why is it that the disadvantaged groups are worse-off in the rural sector than in the urban sector? These questions could be explained mainly by the large human capital differences among these groups and across the two sectors. It may be noted that gender and caste are two different entities. A large share of the SC/ST population lives in villages which often lack the educational infrastructure whereas the residents of urban areas are relatively well off as the infrastructure is easily available and accessible. The quality of education across the two sectors may also account for part of the unexplained differential in the two sectors. Further, occupational distribution of social groups is quite different across the two sectors. For instance, the share of regular wage workers is high in the urban sector whereas casual workers dominate the rural sector, and the share of each social group also varies in these employment categories.

The large difference in educational attainment between caste groups is an important cause of inequality in employment status (Ito 2009). Some of these groups have high incidence of illiteracy. ST population stays in villages where well-paid jobs are less available and, as a consequence, they are likely to get less returns to formal schooling which deters their investment in formal schooling (Kijima 2006). Though the occupational structure has undergone a profound change over time, caste divisions have remained more or less the same (Deshpande 2000). A major section of the SC/ST population is concentrated in the primary occupations such as farm, fish, hunt and log (Deshpande 2001). These two groups are disproportionately clustered in the lower ladder of occupations: casual labor, agricultural labor and unemployed whereas the upper castes dominate the more prestigious and better-paying occupations (Deshpande 2011).

Spatial location and schools could also explain caste discrimination in labor market. Borjas (1995) in the context of the US shows that residential segregation and the external effect of ethnicity are linked partly because ethnic capital summarizes the socioeconomic background of the neighborhood where the children are raised. Neighborhoods isolate people of similar backgrounds and promote a set of cultural attitudes, social contacts, and economic opportunities that affect individuals throughout their lives. Social networks show an underlying segregation (Deshpande 2000). In India, upper and lower caste intermarriage is rare and the practice of 'untouchability' is still continued mainly in rural areas.

The large endowment difference, observed in the case of social groups, suggests that pre-labor market discriminatory practices with respect to education, health, and nutrition are more crucial in explaining wage differentials than labor market discrimination (Madheswaran and Attewell 2007). However, it may be noted that the whole part of discrimination component cannot be attributed to current discrimination. It has been argued that unequal labor market outcomes have their roots in discrimination in the past that has caused more harm to deprived backgrounds of the disadvantaged workers. Pre-labor market discrimination affects earnings indirectly by means of lower out of school investments, poor quality of education, field of study, accessibility to higher education, poorer nutrition and health status, and lower social capital. These may result in lower endowments and persistent wage differentials over time (Altonji and Blank 1999; Das and Dutta 2007). Further, the discrimination in access to schooling and to wage employment cannot be controlled for and explained through this analysis.

## 8 Conclusions and policy implications

This paper examines wage discrimination between gender groups and between social groups in India using a na-

tionally representative household survey. We use the traditional Blinder-Oaxaca decomposition method corrected for the sample selection bias. We find a strong picture of the wage differential between males and females, and between social groups. The wage gap is more pronounced in the rural sector than in the urban sector. The decomposition analysis shows that the wage gap between males and females is largely because of discrimination in the labor market. This result remains valid when we separately examine gender differences in NSC and SC/ST groups. However, the wage differential between NSC and SC/ST is due to the differences in endowments.

A large part of the wage differential between males and females, which is due to discrimination, indicates that human capital of females could not be realized as completely as that of males. Policies should be planned for more accessible employment opportunities for females. Equal employment opportunities and anti-discriminations legislations should be followed strictly. Policies should also be aimed to encourage women's education particularly in rural areas. Kingdon (1998) advocates that policies to discourage gender-discrimination in the labor market will increase the rates of return to females' education and enhance their incentives to invest in education. Educating females has some other benefits like reduced infant mortality and lower fertility. A large endowments difference between social groups indicates that there is a need to promote educational opportunities for socially backward sections of the population. The government should also ensure that the disadvantaged groups of the society get full participation in schooling as they have high incidence of illiteracy. Therefore, policies should be aimed to reduce inequalities in access to education.

**Acknowledgements** I thank Sripad Motiram, S. Chandrasekhar, Ankush Agrawal and Edgar Baake for comments and suggestions. I am grateful to two anonymous referees of this journal for providing insightful comments on an earlier draft of this paper. Institutional support provided by Indira Gandhi Institute of Development Research, Mumbai is gratefully acknowledged. The usual disclaimer applies.

## References

- Agrawal, T.: Returns to education in India: some recent evidence. *J. Quant. Econ.* **10**(2), 131–151 (2012)
- Agrawal, T.: Educational attainment in educationally backward states of India: some implications for the right to education act. *Int. J. Educ. Econ. Dev.* **4**(1), 89–99 (2013a)
- Agrawal, T.: Educational inequality in rural and urban India. *Int. J. Educ. Dev.* (2013b) doi:[10.1016/j.ijedudev.2013.05.002](https://doi.org/10.1016/j.ijedudev.2013.05.002)
- Aigner, D.J., Cain, G.G.: Statistical theories of discrimination in labor markets. *Ind. Labor Relat. Rev.* **30**(2), 175–187 (1977)
- Altonji, J.G., Blank, R.M.: Race and gender in the labor market. In: Ashenfelter, O., Card, D. (eds.) *Handbook of Labor Economics*, pp. 3143–3259. Elsevier, Amsterdam (1999)

- Arrow, K.J.: Some mathematical models of race discrimination in the labor market. In: Pascal, A.H. (ed.) *Racial Discrimination in Economic Life*, pp. 187–204. Heath Press, Lexington (1972)
- Asadullah, M.N.: Returns to education in Bangladesh. *Educ. Econ.* **14**(4), 453–468 (2006)
- Banerjee, B., Knight, J.B.: Caste discrimination in the Indian urban labour market. *J. Dev. Econ.* **17**(3), 277–307 (1985)
- Becker, G.S.: *The Economics of Discrimination*. University of Chicago Press, Chicago (1957)
- Beteille, A.: Race and descent as social categories in India. *Daedalus* **96**(2), 444–463 (1967)
- Beteille, A.: Classes and communities. *Econ. Polit. Wkly* **42**(11), 945–952 (2007)
- Bhaumik, S.K., Chakrabarty, M.: Does move to market have an impact on earnings gap across gender? Some evidence from India. *Appl. Econ. Lett.* **15**(8), 601–605 (2008)
- Blau, F.D., Lawrence, M.K.: Understanding international differences in the gender pay gap. *J. Labor Econ.* **21**(1), 106–144 (2003)
- Blinder, A.S.: Wage discrimination: reduced form and structural estimates. *J. Hum. Resour.* **8**(4), 436–455 (1973)
- Borjas, G.J.: Ethnicity, neighborhoods, and human-capital externalities. *Am. Econ. Rev.* **85**(3), 365–390 (1995)
- Cain, G.G.: The economic analysis of labor market discrimination: a survey. In: Ashenfelter, O., Layard, R. (eds.) *Handbook of Labor Economics*, pp. 693–785. Elsevier, Amsterdam (1986)
- Cotton, J.: On the decomposition of wage differentials. *Rev. Econ. Stat.* **70**(2), 236–243 (1988)
- Darity, W.A., Mason, P.L.: Evidence on discrimination in employment: codes of color, codes of gender. *J. Econ. Perspect.* **12**(2), 63–90 (1998)
- Das, M.B., Dutta, P.V.: Caste matter for wages in the Indian labor market: caste pay gaps in India. Human Development Unit, The World Bank (2007)
- Desai, S., Dubey, A., Joshi, B.L., Sen, M., Shariff, A., Vanneman, R.: *Human Development in India: Challenges for a Society in Transition*. Oxford University Press, New Delhi (2010)
- Deshpande, A.: Recasting economic inequality. *Rev. Soc. Econ.* **LVIII**(3), 381–399 (2000)
- Deshpande, A.: Caste at birth? Redefining disparity in India. *Rev. Dev. Econ.* **5**(1), 130–144 (2001)
- Deshpande, A.: *The Grammar of Caste: Economic Discrimination in Contemporary India*. Oxford University Press, New Delhi (2011)
- Deshpande, S., Deshpande, L.K.: Gender-based discrimination in the urban labour market in India. *Indian J. Labour Econ.* **40**(3), 545–562 (1997)
- Dhesi, A.S.: Unequal opportunities in education and labour market in India. *Int. Policy Rev.* **6**(1), 61–69 (1996)
- Divakaran, S.: Gender based wage and job discrimination in urban India. *Indian J. Labour Econ.* **39**(2), 235–257 (1996)
- Dolton, P.J., Makepeace, G.H.: Sample selection and male-female earnings differentials in the graduate labour market. *Oxf. Econ. Pap.* **38**(2), 317–341 (1986)
- Duraisamy, P.: Changes in returns to education in India, 1983–1994: by gender, age-cohort and location. *Econ. Educ. Rev.* **21**(6), 609–622 (2002)
- Dutta, P.V.: Returns to education: new evidence for India, 1983–1999. *Educ. Econ.* **14**(4), 431–451 (2006)
- Gaiha, R., Thapa, G., Imai, K., Kulkarni, V.S.: Disparity, deprivation and discrimination in rural India. *BWPI working paper 13*, Brooks World Poverty Institute (2007)
- Heckman, J.J.: Sample selection bias as a specification error. *Econometrica* **47**(1), 153–161 (1979)
- Ito, T.: Caste discrimination and transaction costs in the labor market: evidence from rural North India. *J. Dev. Econ.* **88**(2), 292–300 (2009)
- Kannabiran, K.: A cartography of resistance: the national federation of Dalit women. In: Ray, R. (ed.) *Handbook of Gender*, pp. 185–206. Oxford University Press, New Delhi (2012)
- Kijima, Y.: Caste and tribe inequality: evidence from India, 1983–1999. *Econ. Dev. Cult. Change* **54**(2), 369–404 (2006)
- Kingdon, G.G.: Labour force participation, returns to education, and sex discrimination in India. *Indian J. Labour Econ.* **40**(3), 507–526 (1997)
- Kingdon, G.G.: Does the labour market explain lower female schooling in India? *J. Dev. Stud.* **35**(1), 39–65 (1998)
- Kingdon, G.G., Unni, J.: Education and women's labour market outcomes in India. *Educ. Econ.* **9**(2), 173–195 (2001)
- Kingdon, G.G., Theopold, N.: Do returns to education matter to schooling participation? Evidence from India. *Educ. Econ.* **16**(4), 329–350 (2008)
- Kochar, A.: Do schooling policies contribute to schooling inequality? School location policies in rural India. *SIEPR policy brief*. Stanford Institute for Economic Policy Research, Stanford University (2007)
- Madheswaran, S., Attewell, P.: Caste discrimination in the Indian urban labour market: evidence from the national sample survey. *Econ. Polit. Wkly* **42**(41), 4146–4153 (2007)
- Mincer, J.: *Schooling, Experience and Earnings*. National Bureau of Economic Research, New York (1974)
- Motiram, S., Singh, A.: How close does the apple fall to the tree? Some evidence from India on intergenerational occupational mobility. *Econ. Polit. Wkly* **XLVII**(40), 56–65 (2012)
- Munshi, K., Rosenzweig, M.: Traditional institutions meet the modern world: caste, gender, and schooling choice in a globalizing economy. *Am. Econ. Rev.* **96**(4), 1225–1252 (2006)
- Munshi, K., Rosenzweig, M.: Why is mobility in India so low? social insurance, inequality, and growth. NBER working paper No. 14850, National Bureau of Economic Research, Cambridge, MA (2009)
- Neumark, D.: Employers' discriminatory behavior and the estimation of wage discrimination. *J. Hum. Resour.* **23**(3), 279–295 (1988)
- Oaxaca, R.: Male-female wage differentials in urban labour markets. *Int. Econ. Rev.* **14**(3), 693–709 (1973)
- Oaxaca, R.L., Ransom, M.R.: On discrimination and the decomposition of wage differentials. *J. Econom.* **61**(1), 5–21 (1994)
- Phelps, E.S.: The statistical theory of racism and sexism. *Am. Econ. Rev.* **62**(4), 659–661 (1972)
- Puhani, P.: The Heckman correction for sample selection and its critique. *J. Econ. Surv.* **14**(1), 53–68 (2000)
- Reimers, C.W.: Labor market discrimination against Hispanic and black men. *Rev. Econ. Stat.* **65**(4), 570–579 (1983)
- Singh, A., Das, U., Agrawal, T.: How inclusive has regular employment been in India? A dynamic view. *Eur. J. Dev. Res.* **25**(3), 486–494 (2013)
- Srinivas, M.N.: *Social Change in Modern India*. Orient BlackSwan, New Delhi (1966)
- Srinivas, M.N.: Some reflections on the nature of caste hierarchy. *Contrib. Indian Sociol.* **18**(2), 151–167 (1984)
- Thorat, S., Attewell, P.: The legacy of social exclusion: a correspondence study of job discrimination in India. *Econ. Polit. Wkly* **42**(41), 4141–4145 (2007)
- Thorat, S., Newman, K.S.: *Blocked by caste: economic discrimination and social exclusion in Modern India*. Oxford University Press, New Delhi (2010)
- Weichselbaumer, D., Winter-Ebmer, R.: A meta-analysis of the international gender wage gap. *J. Econ. Surv.* **19**(3), 479–511 (2005)

**Tushar Agrawal** is an Associate Fellow at the National Council of Applied Economic Research (NCAER), New Delhi. His research is in the area of education and labor issues in India.