Economic Affairs, Vol. 62, No. 4, pp. 609-620, December 2017
DOI: 10.5958/0976-4666.2017.00074.2

Effect of Endowments on Gender Wage Differentials: A Decomposition Analysis for Indian Labour Market

Suraj Sharma

Assistant Professor I, Amity School of Economics, Amity University, Noida, India

Corresponding author: ssharma46@amity.edu

ABSTRACT

The earning function clearly supports the existing evidence of the significant positive coefficients for education and the marginal wage effects are increasing with the level of education for both the genders. There are clear evidence of caste bias for males, location and regional bias for both the genders in earning. Our decomposition results show that endowment component which shows the existence of pre-market discrimination is smaller than the discrimination component. Discrimination explains 66.1 per cent of the lower wages of female individuals when compared to males. Discrimination component is the highest for the production workers (81.3 per cent) followed by professionals (77.6 per cent), agriculture/allied workers (77.4 per cent), clerical workers (65.9 per cent) and is least for sales/services workers (61.4 per cent). Gender wage discrimination is very high for the urban areas (86.3 per cent) than the rural settings (71.3 per cent). Large discrimination differences are a matter of concern for thepolicy makers.

JEL Classifications: I21, J30, J31.

Keywords: Gender, wage, discrimination, decomposition, occupation

Indian labour market is characterised by the low rate of mobility; imperfect markets, the problem of unemployment, casual labour, difference in occupational choices that lead to occupational dissimilarities (Petersen and Snartland, 2004) and most importantly wage differentials. The growing sense of informalisation of work is a serious threat, which adds fuel to the fire. The wage differentials are not only characterised with caste prejudices but it also has long deep roots of gender disparities which are more prevalent for the women of the marginalised sections, who themselves are marginal. So, women are double burdened or discriminated, firstly as they belong to a certain gender and secondly as they belong to a particular caste. The undervaluation of female work, stereotypical consideration of primary responsibility of female in household and the care work limits the representation of females in the labour market (Chen *et al.* 2005).

Gender wage discrimination has always been an objective for the scholars and the measurement[1] of this discrimination was mostly based on human capital theory, which is tied to productivity. According to this theory, the more the human capital a person attains, more the productivity will be and vice versa. In the presence of non-discriminatory environment, the wages of males and females are supposed to be equal to a particular level of human capital (education level). The wage differentials are solely because of the difference in productivity levels and on the other hand discrimination is said to be in existence, as male-female wage differentials cannot be solely explained by the difference in the productivity levels.

Very few studies can be found on caste discrimination in earnings (wage) and occupation (job) in the Indian scenario. The pioneer work in this context is done by Banergee and Knight (1985) who found that the "discrimination" component accounts for the significant part of raw wage differentials and that "wage discrimination" superimposed the "job discrimination". Similar results were also found by Borooah *et al.* (2007) that job discrimination is only a part of the observed wage differentials. On the other hand studies of Madheswaran and Attewell (2007) found discrimination accounted for a large part of the gross earning difference between HCs and SC/STs, with "job discrimination" being more important than "wage discrimination". However, there is also an evidence that the gap is declining over time.

Thorat and Attewell (2007) and Siddique (2011) used a correspondence study to determine the extent of caste-based discrimination in the Indian private sector and found that on an average, low-caste applicants need to send 20 per cent more resumes than high-caste applicants to get the same callback. Differences in callback which favoured high-caste applicants are particularly large when hiring is done by male or by Hindu recruiters. A study (Thorat and Sadana 2009) found that even the ownership of private enterprise continued to be highly skewed along caste lines. Indeed, Chakravarty and Somanathan (2008) studied the placement outcomes at Indian Institute of Ahmedabad, which revealed no caste discrimination may be possible because campus selection was done in an organized manner.

In addition, studies like Tilak (1980), Kingdon and Unni (2001), Goel (2009) and Sengupta and Das (2014) investigated gender-based discrimination in the Indian urban labour market. Tilak (1980) examined the returns to education for males and females separately in India and found that lower levels of gender wage gaps were associated with higher education groups of females. On the other hand, Kingdon and Unni (2001) using NSS data for two Indian states namely Tamil Nadu and Madhya Pradesh found that there exist high gender wage discrimination and education played an insignificant role in combating this discrimination.

Bhaumik and Chakrabarty (2008) using the NSS data from 1987 to 1999 found that gender wage gap reduced significantly asthis period owed to the increasing returns of experience for females in the Indian labour market. Using 66th round (2009-10) round data of NSS, Khanna (2010) found that higher gender wage gap is significantly associated with the lower end of wage distribution, suggesting that among high wage groups this gender gap is low compared to low wage groups and these low wage groups are more vulnerable to wage discrimination.

Sengupta and Das (2014) in their study using 50th and 66th round unit level data of Employment Unemployment situations in India by NSS found that gender wage gap has regional differences as at every educational levelit has been increasing in the rural areas and on the other hand it has been decreasing in the urban areas, moreover this gap remained very high among the illiterate workers. Women have been paid lower wages and this gender discrimination was more profound in the socially backward classes like SC/STs and religious minorities like Muslims and gender discrimination, superimposed caste and religious discrimination,thereby, have madeit too difficult to beat par with others for the women belonging to the lower caste and religious minorities. Deininger et al. (2013) focused on wage discrimination in informal labour markets, an issue largely neglected in the Indian literature despite the fact that informal markets are the main destination for the poorest section of the population. Their results suggest that gender wage discrimination is larger in informal labour markets than in formal labour markets and more pronounced in the agricultural sector.[2]

Most empirical studies on wage discrimination in India have found females earn significantly lower wages than males. Not only females but weaker social groups like SCs and STs, rural workers also earn substantially lower wages. In the light of the above studies reviewed so far, this study tries to focus on gender-based discrimination in earnings, the effect of increasing education and its return, not only by productivity differences but also with the post-labour market discrimination by various econometric tools and techniques described in the methodology of the study. This study improves the literature as it includes various important variables like regional differences, nature of job, etc.

Hypotheses

$H0_A$: There is no gender wage differential in Indian labour market.

$H1_A$: There is significant gender wage differential in Indian labour market.

$H0_B$: There is no gender wage discrimination in different occupational categories for females.

$H1_B$: There is significant gender wage discrimination in different occupational categories for females.

Data and Methodology

The study uses nationally representative unit level secondary data from Indian Human Development Survey (IHDS) collected first in 2004-05 and again in 2011-12. The study is based on the individual data. The study confines itself to individuals aged between 15 and 65 years, either employed or seeking employment (unemployed). The lower bound of the age group ensures that the individual is not a child labourer. The survey has detailed demographic information (e.g., age, gender, marital status, household size, religion, social group, sector,and place of residence) and socioeconomic position (e.g., land ownership, educational attainment, occupation and industry, type of job, wages and earnings) among several other characteristics.[3]

Variables

In the analysis, we use the following variables:

- **Age:** Age of an individual in years.
- **Education:** Education of an individual is grouped in one of the following categories: (i) Illiterate or below primary (0-2 years), (ii) Primary (3-5), (iii) Middle (6-8), (iv) Secondary (9-10), (v) Higher secondary (11-12), and (vi) Graduate (above 12 years).
- **Marital status:** Married or unmarried. The married group includes married, divorced and widowed.
- **Number of children:** Number of children (0 – 14 years of age).
- **Social groups:** Forward caste (FCs), other backward classes (OBCs) Scheduled Castes (SCs).
- **Religion:** Hindus or non-Hindus.
- **Sector of residence:** Rural or urban.
- **Earnings/wages:** The earnings variable is an hourly wage, obtained by dividing the total amount received during a year (or per day or month) by the number of days worked in a year and the number of hours an individual usually works in a day. The wage distribution is trimmed by 0.1 percent at both the ends of the distribution.
- **Region:** To capture the regional variations, we group all the states of the country into four regions: Northern, Eastern, Southern, and Western.[4]
- **Occupational characteristics:** In the dataset, occupations are recorded using the National Classification of Occupations - 1968 (NCO-68) scheme at the two-digit level. We prefer to work with the broadest classification of occupations (at the one-digit level). We have seven occupational categories at the one-digit level: (i) Professional, technical and related workers (codes 0 and 1), (ii) Administrative, executive and managerial workers (2), (iii) Clerical and related workers (3), (iv) Sales workers (4), (v) Service workers (5), (vi) Farmers, fishermen, hunters, loggers and related workers (6), and (vii) Production and related workers, transport equipment operators and labourers (7, 8 and 9). In our analysis occupation codes 0, 1 and 2 are clubbed together to make it a single category as "Professional and Administration". Accordingly, codes 4 and 5 are clubbed as "Sales and Services", 3 and 6 as a separate category "Clerical" and "Agriculture" respectively and lastly codes 7, 8, 9 as "Production".
- **Type of job:** Part time, full time and overtime[5].

The objective of the study is to find out any pieces of evidence of gender discrimination in earning functions, occupation and returns to education. First, the study uses single equation models to predict earnings separately for both males and females from the characteristics of all the individuals. In this case, the results present marginal differences between genders, holding other control variables (characteristics) at their mean value, and thereby yields a biased result because it constraints the value of coefficient of the explanatory variables, such as education, age (experience), gender and location, etc.

The Oaxaca Blinder decomposition

This second approach parts the observed wage gap into two "endowment" and a "coefficient" component called "decomposition techniques". The "endowment" part is such that it shows the component of wage differentials explained by individual 'characteristics' like education, age, and others, and the later part is derived as an unexplained residual and shows the component of wage differentials explained by 'discrimination'. This method was first developed by Blinder (1973) and Oaxaca (1973), which is called The Blinder Oaxaca decomposition method. The details on how this decomposition is done canbe seen in their research papers mentioned earlier, Jann (2008).

The Blinder-Oaxaca decomposition technique can be explained as follows:

Here this study has two groups of males and females with an outcome variable, Y (log of hourly wages) and a set of variables like education, age etc. Now the difference of mean outcome is to be computed:

$$D = E(Y_M) - E(Y_F)$$

Where E (Y) denotes the expected value of outcome variable, is accounted by the group difference in the regressors. The linear model is as follows:

$$Y_i = X_i'\beta_i + \varepsilon_i \qquad \ldots(1)$$

Where, E (ε_i) = 0 and i € (male, female)

Where, X is a vector containing the predictors and a constant β contains the slope parameters and the intercept, and ε_i is the error term. The mean outcome difference can be expressed as the difference in the linear prediction at the group-specific means of the regressors. That is,

$$D = E(Y_M) - E(Y_F) = E(X_M)'\beta_M - E(X_F)'\beta_F \qquad \ldots(2)$$

Because,

$$E(Y_i) = E(X_i'\beta_i + \varepsilon_i) = E(X_i'\beta_i) + E(\varepsilon_i)$$
$$= E(X_i)'\beta_i$$

Where, E (β_i) = β_i and E (ε_i) = 0 by assumption

To identify the contribution of group differences in predictors to the overall outcome difference, (1) can be rearranged, for example, as follows:

$$D = \{E(X_M) - E(X_F)\}'\beta_F + E(X_F)'(\beta_M - \beta_F) + \{E(X_M) - E(X_F)\}'(\beta_M - \beta_F) \qquad \ldots(3)$$

Here what we get is called the 'threefold' decomposition that is the mean wage difference (D) is divided into three components:

$$R = E + C + I$$

The first component, $E = \{E(X_M) - E(X_F)\}'\beta_F$ amounts to the part of the differential that is due to group difference in the regressors (the "endowments effect"). The second component, $C = E(X_F)'(\beta_M - \beta_F)$ measures the contribution of differences in the coefficients (including differences in the intercept, "discrimination" component). And the third component $I = \{E(X_M) - E(X_F)\}'(\beta_M - \beta_F)$ is an interaction term accounting for the fact that differences in endowments and coefficients exist simultaneously between the two groups.

The decomposition shown in (3) is formulated from the viewpoint of the females. That is, the group differences in the regressors are weighted by the coefficients of females (β_F) to determine the endowments effect (E) and similarly for the C component the differences in coefficients are weighted by female predictor levels. Naturally, the differential can also be expressed from the viewpoint of males, yielding the reverse 'threefold' decomposition (eq. 4),

$$D = \{E(X_M) - E(X_F)\}'\beta_M + E(X_M)'(\beta_M - \beta_F) + \{E(X_M) - E(X_F)\}'(\beta_M - \beta_F)$$

Scholars used either of these two equations (equation 3 or 4) based on their assumptions about the existing market wage structure. It can be argued that under discrimination males are paid competitive wages more than the females. Coefficient should be used as the non-discriminatory wage structure. Therefore, the issue in literature is to determine how the wage structure would prevail in the absence of discrimination. This choice poses the well-known 'index number' problem given that we would use either the male or the female wage structure as the non-discriminatory benchmark.

To overcome this problem and to extend the wage discrimination component further, Cotton (1988), Neumark (1988) and Oaxaca and Ransom (1994) have proposed an alternative decomposition prominent in the discrimination literature results from the concept that there is a nondiscriminatory coefficient vector that should be used to determine the contribution of the differences in the predictors. Let β^* be such a nondiscriminatory coefficient vector. The outcome difference can then be written as;

$$D = \left\{E\left(X_M\right) - E\left(X_F\right)\right\}' \beta^* + E\left(X_M\right)' \left(\beta_M - \beta^*\right) + E\left(X_F\right)' \left(\beta^* - \beta_F\right) \quad \ldots(5)$$

Where the first term on RHS of equation (5) is the part of the wage differential that is explained by group differences in the regressors (skill difference), the second term is the overpayment to males due to favouritism and the third term is underpayment to females due to discrimination. The equation (6) is operationalised under the assumption of non-discriminatory wage structure by assigning proportions of males (P_M) and females (P_F) weight to the wage structure and β^* is defined as;

$$\beta^* = P_M \beta_F + P_F \beta_F \quad \ldots(6)$$

ECONOMETRIC ANALYSIS AND RESULTS

The gender wage gap

The table on the descriptive statistics for the study is given in Appendix. Fig. 1 shows probability density (kernel density[6] plot) as to capture the existing gender wage gaps persisted in earnings.

Fig. 1: Kernel density estimates of log of hourly wages for males and females

Source: *Computed by the author from IHDS II unit level data.*

The distance between the densities of males and females' wage distribution (densities) at any point represents the extent of the raw wage gap. As evident from the figure that male wage density is placed or skewed rightward with respect to female wage density, indicating towards significant gender wages gaps.

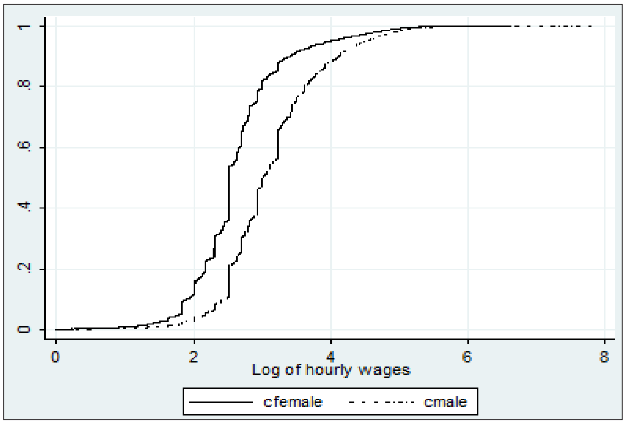

Fig. 2: Cumulative density estimates of log of hourly wages for males and females

Source: *Computed by the author from IHDS II unit level data.*

The gender gap can be better viewed in Fig. 2, which shows the cumulative density function (CDF) of male and female (log) hourly wages. The horizontal distance between the two functions is the gender wage gap and the plot indicates towards raw gender wage gap.

OLS regression for earnings and returns to education

The linear regression modeling (single equation technique) uses the logarithm of hourly wage rate as our regress and age, level of education, caste, location, region, marital status and number of children as our regressors to estimate the earning function of males and females separately (Table 1). The results clearly support the existing evidence of significant positive coefficients for education and the marginal wage effects are increasing with the level of education for both genders.

Table 1: OLS result of earning function (hourly wages)

Variables	Males	Females
Age	0.0312654***	0.0214162***
Age^2	-0.0002782***	-0.0001921***
Below primary	0.0725216***	0.0552119***
Primary	0.1392288***	0.1182708***
Upper primary	0.2070691***	0.1590479***
Secondary	0.3543888***	0.2818276***
Higher secondary	0.464217***	0.6313025***

Graduate and above	0.9249455***	1.214806***
OBC	-0.0934517***	-0.0421573***
SC	-0.0400013***	0.010678
Location (Urban)	0.2431335***	0.1697621***
Marital status (Married)	0.0671954***	-0.0141168***
Number of children	-0.0216873***	-0.0141168***
North (Dummies)	0.1037539***	0.0111143
South (Dummies)	0.213496***	0.0251017
West (Dummies)	-0.1254931***	-0.1201325***
Constant	2.039798***	1.918377***
R-square	0.2715	0.2699
F-value	663.22***	203.57***
N	35907	15800

Source: *Computed by the author from IHDS II unit level data (Robust results) Note: N stands for sample size and *** p < 0.01, ** p<0.05, * p<0.10*

It is evident from Table 1 that return to education as a whole varies between genders. Lower levels of education were favourable to males upto secondary level of education, but after that it favours females from the higher secondary and above, which shows that for better wage outcomes females have to get educated more than that of a male counterpart and with lower levels of education a female is going to get less than that of a male with the same level of education. The results also point out that if a female is educated to a certain level as mentioned above can perform better than those of a man as a female will receive more wages than a man. The coefficient of experience (age) also favoured the male and clears that experience bias is also there for females.

There are clear evidence of caste bias in earnings as OBCs and SCs are vulnerable categories. Males from the OBC and SC social groups are getting less than the general social category and even females from upper caste are vulnerable of wage differentials. Caste bias in earnings is more prevalent for males than females and SC females are getting somewhat more than the generals but the part is not statistically significant so we cannot come to a conclusion here. There is clear location bias for both genders in favour of the urban settlement and most importantly the bias is more prevalent for females as evident from the low coefficient. Regional disparities in earnings are clear, more prone to males and western states of India face more wage discrimination than the others. Number of children is negatively related to the earnings of an individual and being married is definitely going to hinder the earnings of a female individual as the coefficient of marital status is negative for females.

As the discussion goes more on the relationship between educational attainment and hourly wages for male and female separately, we draw Fig. 3 which shows the log of hourly wages for male and female on the vertical axis and level of education on the horizontal axis. The figure indicates that the gender wage gap shows a pervasive increase as the level of education increases up to secondary and further it decreases for higher level of education. The rural-urban wage differentials for both males and females are substantial, which is biased for rural areas and females.

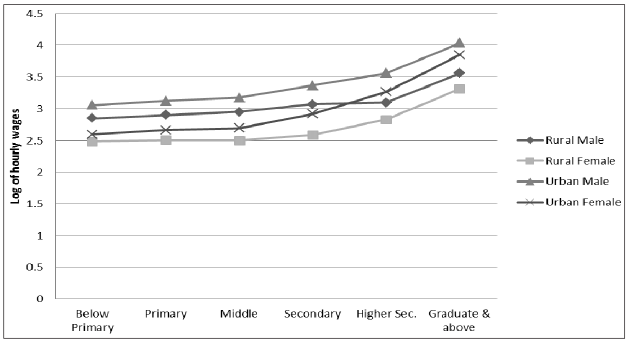

Fig. 3: Log of hourly wages by level of education

Source: *Computed by the author from IHDS II unit level data.*

Fig. 3 also shows a sudden rise in the hourly wages after the secondary level of education and it is prevalent and very clear for urban females upto secondary level of education. The rural males were getting more than urban females of the same education level but after the secondary level of education the scenario changed and urban females got more wages than the rural males. This perfectly goes with the finding by Agrawal (2012) who had stated about the returns to education increases after middle schooling in both rural and urban areas in India.

The kernel density estimates the log hourly wage for males and females at low and high education level[7] separately for both the rural and urban sectors. The plots indicate that the female distributions are more skewed or placed towards the left than the male distributions for both sectors but the four plots show clearly different wage distributions. As the figures show that even for the highly educated urban

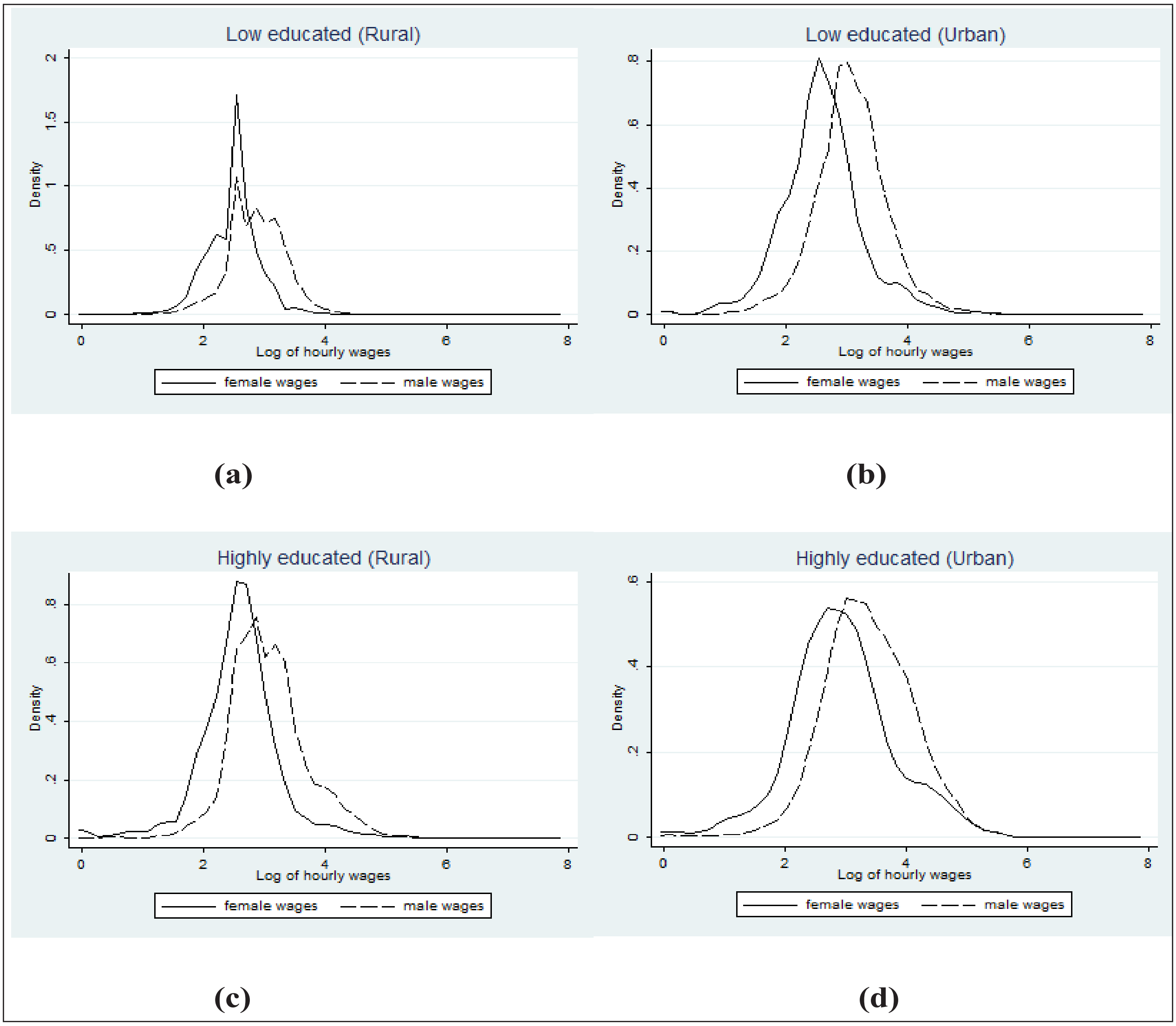

Fig. 4(a) to 4(d): Kernel density estimates of log of hourly wages by education level

Source: *Computed from IHDS II unit level data.*

sector the wage gap between males and females is prevalent which is an alarming situation and may lead to the feeling of inferiority among females who are equally capable of males in terms of human capital have been paid less than that of males.

Decomposition analysis

It can be seen that the coefficient (discrimination) component is much larger than the endowment component (characteristics) which proves that an evidence is there for post labour market discrimination other than individual characteristics like educational attainment, experience, health, etc. Here discrimination explains 78.5 per cent of the lower wages of female individuals when compared to males in 2011-12 and the interaction of endowment and discrimination favours females (negative coefficient). Endowment reasons only 22.9 per cent of the lower wages for the females when compared to males in 2011-12. It can be seen from Table 2 that over 2004-05 to 2011-12 the endowment part has somewhat increased from 19.7 to 22.9 showing that individual characteristics have played more role in deciding individuals wages in 2011-12 than in 2004-05 but the discrimination component is very large and it showed very little progress over the years.

It is noticeable here that the small endowment components show that individual characteristic like education and other endowment differences or pre-market discrimination or pure productivity differences, which explains a small part of wage differentials among the genders. However, discrimination component is very large and is significant which results in lower wages of female individuals with the same productivity levels as male individuals.

Table 2: Blinder-Oaxaca decomposition results using female mean wage

Components of decomposition (I)	Female vs Male (2005) (II)	Female vs Male (2012) (III)
Difference	0.6552878***	0.500216***
Due to endowment (E)	0.1294273***	0.1144804***
Due to coefficient (C)	0.530779***	0.3928575***
Due to interaction (I)	-0.0049185***	-0.0071219
Endowment as per cent of total (E/R)	19.7	22.9
Discrimination as per cent of total (C/R)	81.0	78.5

Source: *Computed by the author from IHDS I and II unit level data;* ***Note:*** *Raw Differential (E+C+I) = (R), N stands for sample size and *** $p < 0.01$, ** $p<0.05$, * $p<0.10$*

Table 3 shows different decomposition results using different approaches. It can be seen that according to Blinder-Oaxaca decomposition wage differentials due to discrimination component is 77.1 per cent using female means as weight and 78.5 per cent using male means as weight, which somewhat underestimates the true value of the skill differentials. The Oaxaca-Ransom (pooled) decomposition estimates the discrimination component at 76.8 per cent and the Cotton/Neumark decomposition yields discrimination component at 66.1 per cent and has been further decomposed, which indicates the magnitude of more favouritism to males (45.9 per cent) than a disadvantage to females (20.2 per cent). The component of the unexplained wage gap springs from the component (female disadvantage) measures the female disadvantage due to labour market discrimination which is the equivalent to the ratio between the wage females should receive if the non-discriminatory wage structure were enforced and the wage they actually receive. This is essentially an indication of male favouritism in the labour market. As this Cotton/Neumark decomposition has minimum standard errors, the estimates are perhaps most reliable among others.

It can be seen in Table 4 that earnings differential due to caste discrimination were 7.1 per cent for SCs in comparison to the general category and

Table 3: Various decomposition results

Components	Blinder-Oaxaca using female means as weight	Blinder-Oaxaca using male means as weight	Oaxaca-Ransom (pooled)	Cotton/Neumark
Explained (skill differences)	22.9*** (0.004563)	21.5*** (0.005363)	23.2*** (.004155)	33.9*** (.004303)
Unexplained (discrimination)	77.1*** (0.006474)	78.5*** (0.006981)-	76.8*** (.006408)	66.1*** (.005585)
Overpayment to males	—	—	—	45.9*** (.004021)
Underpayment to females	—	—	—	20.2*** (.001861)

Source: *Author's calculation and corresponding proportions are in the parentheses;* ***Note:*** **** $p < 0.01$, ** $p<0.05$, * $p<0.10$ and standard errors are in the parentheses.*

Table 4: Cotton/Neumark decomposition results for caste (SC vs General) wage gap

Components	Cotton/Neumark
Explained (skill differences)	92.9*** (0.0072512)
Unexplained (discrimination)	7.1*** (0.0083312)
Overpayment to Generals	4.5*** (0.0053422)
Underpayment to SC	2.6*** (0.0030631)

Source: *Author's calculation and corresponding proportions are in the parentheses;* ***Note:*** **** $p < 0.01$, ** $p<0.05$, * $p<0.10$ and standard errors are in the parentheses.*

leads to more magnitude towards the favouritism to generals than the discrimination to SCs. However, it is evident from comparing the Table 3 with Table 4 that gender discrimination surpasses caste discrimination. The study of Das (2012) using national-level data supports these findings and our results are consistent with the same.

Table 5: Cotton/Neumark decomposition results for gender wage gap

Components	Urban	Rural
Differences	0.421539***	0.4590436***
Explained (skill differences)	13.7*** (.0097243)	28.7*** (.0037862)
Unexplained (discrimination)	86.3*** (.0137798)	71.3*** (.0057293)
Overpayment to males	66.7*** (.0108685)	46.9*** (.0039713)
Underpayment to females	19.6*** (.0034002)	24.4*** (.0021679)

Source: *Author's calculation and corresponding proportions are in the parentheses;* ***Note:*** **** p < 0.01, ** p<0.05, * p<0.10 and standard errors are in the parentheses.*

The same decomposition was done separately for rural and urban settlements of India to capture the rural-urban divide in earnings of both genders and we found that the overall difference in hourly wages are more in rural areas when compared to urban settlements. On the other hand, discrimination in earnings was more prevalent in urban areas which has diversified occupation structures than the rural areas and more the discrimination in urban areas when compared to the rural settlements. In the same way, more favouritism was there for a male in both the areas but the magnitude of favouritism was much more in urban settlements, which indicates the need of more intense studies on wage differentials for urban occupations.

Before going to our separate decomposition results for different occupations we should have a look at the concentration of males and females for different occupation categorized as NCO 1968. It can be seen from table 6 that approximately half of female workforce is engaged in agricultural activities which are predominantly rural, followed by production activities (31.6 per cent), sales & services (8.9 per cent), Professionals (8.8 per cent) and Clerical (3.9 per cent) activities. On the other hand, males are predominantly in the production activities (53.6 per cent) and their proportion in agricultural activities (22.9 per cent) is very low when compared to females. One of the reasons for this is migration from rural settlements to urban counterparts of males for their livelihood.

Table 6: Gender and Occupation

Component	Professionals	Clerical	Sales & Services	Agriculture	Production	Total
Female	1,383 (8.78)	610 (3.87)	1,398 (8.87)	7,384 (46.86)	4,983 (31.62)	15,758 (100.0)
Male	2,134 (6.03)	2,643 (7.47)	3,527 (9.97)	8,098 (22.89)	18,969 (53.63)	35,371 (100.0)
Total	3,517 (6.88)	3,253 (6.36)	4,925 (9.63)	15,482 (30.28)	23,952 (46.85)	51,129 (100.0)

Source: *Author's calculation and corresponding proportions are in the parentheses.*

Table 7 provides detailed Cotton/Neumark decomposition analysis of wage differentials by gender and as the study includes occupational variables in the model we get separate decomposition results for each occupational category. Here the discrimination component is much higher for production workers (81.3 per cent) followed by professionals (77.6 per cent), agriculture/allied workers (77.4 per cent), clerical workers (65.9 per cent) and is least for sales/services workers (61.4 per cent). It is important to notice here that discrimination component is highest for 'production' workers which include predominantly the construction workers. The existence of discrimination in production activities is mostly due to the informal nature of these activities and the magnitude of favouritism to males is also the highest for production activities (64.4 per cent). Large discrimination components in professionals and clerical works is a matter of concern that even in these type of activities which are considered to be highly paid and respectable jobs, discrimination is accounting for more than other occupations and not surprisingly, there is a large discrimination component in agriculture occupation category too, which predominantly is the labourer activities and a disadvantage to females as underpayments is the highest (36.9 per cent) for agriculture activities.

Conclusion and Policy Implications

To analyse gender-based discrimination in earnings and the effect of increasing education and its return, the study estimated first the linear regression

Table 7: Cotton/Neumark decomposition results for gender wage gap

Components	Professionals	Clerical	Sales & Services	Agriculture	Production
Differences	0.4927736***	0.3808636***	0.4948602***	0.3955295***	0.4644228***
Explained (skill differences)	22.4*** (0.0196406)	34.1*** (0.023128)	38.6*** (0.0156111)	22.6*** (0.0043351)	18.7*** (0.0047283)
Unexplained (discrimination)	77.6*** (0.0301964)	65.9*** (0.0340366)	61.4*** (0.0221142)	77.4*** (0.0064064)	81.3*** (0.0086488)
Overpayment to males	47.1*** (0.0191438)	53.5*** (0.0282321)	44.0*** (0.0163928)	40.5*** (0.0037332)	64.4*** (0.0071203)
Underpayment to females	30.5*** (0.012612)	12.4*** (0.00675)	17.4*** (0.0067333)	36.9*** (0.0034094)	16.9*** (0.0020999)

Source: *Author's calculation and corresponding proportions are in the parentheses;* ***Note:*** **** p < 0.01, ** p<0.05, * p<0.10 and standard errors are in the parentheses.*

analysis followed by decomposition regression techniques. The earning function clearly supports the existing evidence of significant positive coefficients for education and the marginal wage effects are increasing with the level of education for both males and females. But lower levels of education were favourable to males upto secondary level of education and after that, it favours females from higher secondary and above. There are clear evidence of caste bias for males, location and regional bias for both in earnings.

Our decomposition results show that endowment component which shows the existence of pre-market discrimination is smaller than the discrimination component. Discrimination explains 66.1 per cent of the lower wages of female individuals when compared to males. Discrimination component is the highest for production workers (81.3 per cent) followed by professionals (77.6 per cent), agriculture/allied workers (77.4 per cent), clerical workers (65.9 per cent) and is the least for sales/services workers (61.4 per cent). Gender wage discrimination is very high for urban areas (86.3 per cent) than the rural settings (71.3 per cent).

Large discrimination differences are a matter of concern for policy makers. Constitutional safeguards like equal pay for same work have not contributed to an extent to this regard. As far as endowment differences is concerned, improving the level of education for marginalised, job opportunities in government (public sector) as well as in private sectors and safeguards from discriminatory practices like norms on maternal/pregnancy holidays and many invisible barriers that prevent women from getting higher rank jobs, should be provided to mitigate the huge human capital gap between females and males. Caste and regional parity in earnings are of utmost importance. Our occupational decomposition results indicate that a huge labour market in India is out of the ambit of affirmative action policies and here the government should initiate to reserve the right to equal earnings and opportunity for marginalised sections of the society like women, SC/STs, and rural people.

ABBREVIATIONS

IHDS: Indian Human Development Survey; FC: Forward Castes; OBC: Other Backward Classes; OLS: Ordinary least squares; SC: Scheduled Castes; ST: Scheduled Tribes; NSSO: National Sample Survey Office; NCO: National Classification of Occupations

ACKNOWLEDGEMENTS

The views expressed in this paper are those of the author and do not reflect the views of Amity School of Economics, Amity University, NOIDA, the institution that the author is affiliated to. I am grateful to the participants of the National Seminar on "Gender Equality and Empowerment of Women and Girls" of National Institute of Rural Development & Panchayati Raj (NIRD&PR) held at NIRD&PR, Hyderabad, India and to the related technical session's chairperson Dr. C Ganesh (Dept. of Sociology, Osmania University) for many insightful comments, which greatly improved the paper.

AVAILABILITY OF DATA AND MATERIALS

This paper is prepared by using unit level secondary data from Indian Human Development Survey (IHDS) collected in 2004-05 and 2011-12 at the individual as well as household level. The IHDS data set and questionnaires are publicly available online.

END NOTES

1. The appropriate methods for measurement of discrimination in wages have been described in the methodology section of the study.
2. The same estimates for the non-agriculture sector were insignificant.
3. See Desai *et al.* (2010) for the survey sampling and more information about the survey.
4. The 33 states (and Union Territories) are grouped as follows. The northern region includes nine states: Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu and Kashmir, Punjab, Rajasthan, Uttar Pradesh and Uttarakhand. The eastern region consists of 12 states: Arunachal Pradesh, Assam, Bihar, Jharkhand, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Sikkim, Tripura and West Bengal. The southern region includes five states: Andhra Pradesh, Karnataka, Kerala, Puducherry and Tamil Nadu, and the western region covers seven states: Chhattisgarh, Dadra and Nagar Haveli, Daman and Diu, Goa, Gujarat, Madhya Pradesh and Maharashtra.
5. Individuals working 10 hours or more in a day are considered working overtime, from seven to nine hours as full-time and less than seven hours as part time workers.
6. Kernel density estimation is a special type of Probability distribution function (PDF) and a useful non-parametric technique for visualising the underlying distribution of a continuous random variable.
7. Low educated samples have education below the secondary level (up to middle level) while the highly educated samples comprises those who have secondary or beyond secondary level of education.

REFERENCES

Agrawal, T. 2012. Returns to education in India: Some recent evidence. *Journal of Quantitative Economics*, **10**(2): 131-151.

Banerjee, B. and Knight, J.B. 1985. Caste discrimination in the Indian urban labour market. *Journal of Development Economics*, **17**(3): 277–307.

Bhaumik, S.K. and Chakrabarty, M. 2008. Does move to market have an impact on earnings gap across gender? Some evidence from India. *Applied Economics Letters*, **15**(8): 601-605.

Blinder, A. 1973. Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, **8**(4): 436–455.

Borooah, V.K., Dubey, A. and Iyer, S. 2007. The effectiveness of jobs reservation: Caste, r*eligion and economic status in India. Development Change*, **38**(3): 423–445.

Chakravarty, S. and Somanathan, E. 2008. Discrimination in an elite labour market? Job Placements at IIM-Ahmedabad. *Economic and Political Weekly*, **43**(44): 45-50.

Chen, M., Vanek, J., Lund, F., Heintz, J., Jhabvala, R. and Bonner, C. 2005. *Women work and poverty*. New York, USA: United Nations Development Fund for Women (UNIFEM). Retrieved from http://www.un-ngls.org/orf/women-2005.pdf

Cotton, C.J. 1988. On the decomposition of wage differentials. *Review of Economics and Statistics*, **70**(2): 236-243.

Das, P. 2012. Wage inequality in India: Decomposition by sector, gender and activity status. *Economic and Political Weekly*, **47**(50): 58-64.

Deininger, K., Jin, S. and Nagarajan, H. 2013. Wage discrimination in India's informal labour markets: Exploring the impact of caste and gender. *Review of Development Economics*, **17**(1), 130-147.

Desai, S., Dubey, A., Joshi, B.L., Sen, M., Sharif, A. and Vanneman, R. 2010. Human development in India: Challenges for a society in transition, New Delhi: Oxford University Press. Retrieved from http://www.sonaldedesai.org/human-development-in-india.pdf

Goel, M. 2009. *Trends in wage inequality in India* (Working Paper Series). Columbus, OH: The Ohio State University.

Jann, B. 2008. The Blinder–O*axaca decomposition for linear* regression models. *The Stata Journal*, **8**(4), 453–479.

Khanna, S. 2012. *Gender wage discrimination in India: Glass ceiling or sticky floor? (CDE Working* Paper No. 214) New Delhi: Centre *for Development Economics, Department of Economics, Delhi School of* Economics. Retrieved from http://www.cdedse.org/pdf/work214.pdf

Kingdon, G.G. and Unni, J. 2001. Education and women's labour market outcomes in India. *Education Economics*, **9**(2): 173–195.

Madheswaran, S. and Attewell, P. 2007. Caste discrimination in the Indian urban labour market: Evidences fr*om the national sample survey*. *Economic and Political Weekly*, **42**(41): 4146-4153.

Oaxaca, R. 1973. Male–Female wage differentials in urban labour markets. *International Economic Review*, **14**(3): 693–709.

Oaxaca, R. and Ransom, M. 1994. On discrimination and the decomposition *of wage differentials. Journal of Econometrics*, **61**(1): 5-21.

Petersen, T. and Snartland, V. 2004. *Firms, wages, and incen*tives: Incentive systems an*d their impacts on wages, productivity, and risks. Research in Social Stratification and Mobility*, **21**: 255–288.

Sengupta, A. and Das, P. 2014. Gender *wage discrimination across* social and religious *groups in India: Estimates with unit level data*. *Economic and Political Weekly*, **49**(21): 71-76.

Siddique, Z. 2011. Evidence on caste based discrimination. *Labour Economics*, **18**(S1): 146–159.

Thorat, S. and Attewell, P. 2007. *The legacy of social exclusion: A* correspondence study of job discrimination in India. *Economic and Political Weekly*, **42**(41): 4141-4145.

Thorat, S. and Sadana, N. 2009. Caste and owne*rship of private enterprises*. *Economic and Political Weekly*, **44**(23): 13-16.

Tilak, J.B.G. 1980. Education and labour market discrimination. *Indian Journal of Industrial Relations*, **16**(1): 95-114.

Appendix I: Descriptive statistics of main variables used in the analysis by gender (2011-12)

Variables	Description	Males	Females
Hrwage	Hourly wages	31.35 (37.9803)	19.28 (26.44054)
Lnwage	Log of hourly wages	3.14 (0.7217193)	2.64 (.6906291)
Age	Age in years	35.44 (14.19057)	35.51 (14.12212)
Illit	If respondent is illiterate =1; 0 otherwise	0.07 (0.2631619)	0.07 (.2483692)
Prim	Completed primary school = 1; 0 otherwise	0.07 (0.2605177)	0.08 (.2687042)
Middle	Completed middle school = 1; 0 otherwise	0.29 (0.454261)	0.23 (.4181897)
Secondary	Completed secondary school = 1; 0 otherwise	0.17 (0.3750397)	0.12 (.3266353)
H. Sec	Completed hr. secondary school = 1; 0 otherwise	0.13 (0.3305016)	0.09 (.2857073)
Grad	Completed graduation and above = 1; 0 otherwise	0.10 (0.3012483)	0.07 (.2474278)
Obc	If respondent is OBC = 1; 0 otherwise	0.41 (0.491068)	0.41 (.4912901)
Sc	If respondent is SC = 1; 0 otherwise	0.21 (0.4065825)	0.21 (.4056145)
Married	If respondent is married = 1; 0 otherwise	0.68 (0.4680514)	0.80 (.4001095)
Urban	If respondent is working in urban = 1; 0 otherwise	0.36 (0.4810532)	0.36 (.4785917)
Nchild	Number of children	1.40 (1.526073)	1.52 (1.552823)
Overtime	If worker is doing overtime = 1; 0 otherwise	0.11 (0.3159735)	0.03 (.1736841)
North	If worker is from North India = 1; 0 otherwise	0.34 (0.4735546)	0.34 (.4744797)
South	If worker is from Southern India = 1; 0 otherwise	0.21 (0.4100844)	0.22 (.4140666)
West	If worker is from Western India = 1; 0 otherwise	0.24 (0.4262166)	0.23 (.4219491)
Hindu	If worker is Hindu = 1; 0 otherwise	0.81 (0.3960204)	0.81 (.3956732)
Professional/Admn	If worker' is in professional/administrative = 1; 0 otherwise	0.63 (7.383079)	0.32 (4.717508)
Clerical	If worker is in clerical job = 1; 0 otherwise	0.63 (7.383092)	0.26 (4.715176)
Sales & services	If worker is in sales & services job = 1; 0 otherwise	0.65 (7.38263)	0.31 (4.717409)
Agriculture	If worker is in agricultural job = 1; 0 otherwise	0.78 (7.378946)	0.69 (4.717254)
Production	If worker is in production = 1; 0 otherwise	1.08 (7.361437)	0.54 (4.720947)

Source: *Computed by the author from IHDS II unit level data;* ***Note:*** *standard errors are in the parentheses.*