Indian Journal of Applied Economics and Business
Volume 7; Number 2; 2025 : pp. 385-408
ISSN: 2582-4325
https://DOI:10.47509/IJAEB.2025.v07i02.06

ARF INDIA
Academic Open Access Publishing
www.arfjournals.com

SEX DISCRIMINATION AND WAGE GAP IN THE INDIAN LABOUR MARKET: HECKMAN ESTIMATION AND OAXACA DECOMPOSITION OF EARNINGS DIFFERENTIALS

T. Lakshmanasamy

Formerly ICSSR Senior Fellow and Professor, Department of Econometrics, University of Madras, Chennai. E-mail: tlsamy@yahoo.co.in

Article History
Received : 04 November 2025; Revised : 10 December 2025; Accepted : 15 December 2025; Published : 29 December 2025

Abstract: The existence of an earnings differential or wage gap is one of the many indicators of inequality and a manifestation of discrimination against women. The observed wage gap may be due to factors like social and cultural factors, differential preferences and tastes, low attachment of women in the labour market, occupational segregation, and, importantly, discrimination. This paper empirically estimates the extent of discrimination in the Indian labour market using the 2011-12 Indian Human Development Survey. The labour force participation decision is estimated by the probit method, and separate wage equations are estimated for men and women by OLS. Applying the Oaxaca-Blinder decomposition method, the observed gender wage gap is decomposed into a part that is justified or explained by the differentials in the labour market characteristics of men and women, and the other part that is the unjustified or unexplained differentials that are mainly attributable to discrimination, among other plausible reasons. The decomposition results show that there exists a 26% wage gap against females, attributable to sex discrimination in the Indian labour market. Despite globalisation and labour market reforms, there has been hardly any decline in the gender wage gap in Indian labour since 1981, when 28% wage discrimination against women was observed.

To cite this paper:
T. Lakshmanasamy (2025). Sex Discrimination and Wage Gap in the Indian Labour Market: Heckman Estimation and Oaxaca Decomposition of Earnings Differentials. *Indian Journal of Applied Economics and Business*. 7(2), 385-408. *https://DOI:10.47509/IJAEB.2025.v07i02.06*

Keywords: Gender wage gap, discrimination, labour force participation decision, probit estimation, wage estimation, Oaxaca-Blinder decomposition, discrimination

INTRODUCTION

The wage or earnings or pay gap by gender is normally defined as the relative difference in the average gross earnings of women and men in the economy. It is one of the many indicators of gender inequality in a country when examining labour market participation in terms of gender. According to the International Labour Organisation, progress in reducing the gender pay gap is very slow in every developed and developing country. In developed countries, especially in the United States, a sizable pay gap is very persistent (ILO, 2018). In developing countries, specifically in Latin American countries, the size of the gap as well as the trends towards its growth or reduction vary.

In India, as per the 2011 Census of India, women population constitutes almost half of the country's population, accounting for 496 million in absolute numbers. More than 90% of these women in rural areas and about 69% of women in urban areas have no technical skills. According to the NSSO, in the year 2011-12, the share of regular and salaried wage earners in total employment in India was 12.7% for women workers compared to 19.8% for their male counterparts. The 2011 Census shows that the work participation rate of females is 25.5% and that of males is 53.2%. Rural women show a higher work participation rate of 30% compared with 53% of urban male work participation rate.

As per the NSSO, between 2004-05 and 2009-10, the increase in overall employment was 1.1 million, but between 2009-10 and 2011-12, employment grew from 459 million to 472.9 million only. Between 1999-2000 and 2011-2012, male employment grew by 1.9% per annum while female employment increased by only 0.3% per annum (Verick, 2018). The labour force participation rate for women aged 15 and above in India declined from 37.3% in 2004-05 to 27.2% in 2011-12. From 2004-05 to 2009-10, the number of women workers dropped by 21.3 million, of which 19.5 million are in rural areas, while the number of men workers increased by 22.4 million. Urban areas accounted for 57.2% of the employment growth. As per the latest NSSO 68th round 2011-12, in rural areas, 55% of males and 25% of females, and in urban areas, 56% of males and 16% of females participate in the Indian labour market.

Though women are increasingly participating in the labour force, with varying levels of education and skills, they are mostly concentrated in the private sector as informal and casual workers, and the wages and salaries earned by women are generally much lower than those of their male counterparts. Even then, the annual growth in employment of women in the private sector is only 0.52%, just 0.04% in the government sector, and 0.27% in the entire formal employment (Mohanty et al., 2014). The female work participation rate is lower than the male work participation rate in all states in India, and the scenario is much worse in urban areas.

According to the Self-Employed Women's Association, the wages of women workers were below the minimum wages in some cases and at the same time, the average monthly income of women workers was ₹1815 compared to ₹3842 for male workers (SEWA, 2000). The NSSO 68th round data show that for 2011-12, the actual wage received by women is ₹308 per day compared to ₹407 per day by men. The daily wage/salary earnings of a regular rural wage/salaried worker in the active age of 15-59 years have been ₹298.96, and those of an urban worker have been ₹449.65. By gender, rural males earn ₹322.28, rural females ₹201.56, urban males ₹469.87, and urban females earn ₹366.15 per day, showing a significant rural-urban as well as male-female wage gap in 2011-12 in the Indian labour market. Thus, the 2011-12 NSSO data reveal that women in India earn approximately 75% less than the wages of men, 62% less in rural areas and 78% less in urban areas.

Tables 1 to 4 present the work participation rate in India, the relative wage gap between males and females in both rural and urban areas and the relative wage gap with respect to the educational levels in India, respectively. From Table 1, it is observed that while the work participation rate of males has increased by 1.6%, the female work participation rate has slightly declined by

Table 1: Work Participation Rate and Gender Gap in India, 2001 and 2011

Gender	*2001 (per cent)*	*Gender gap*	*2011 (per cent)*	*Gender gap*	*Percentage change*
Persons	39.1	26.1	39.8	27.8	+0.7
Male	51.7		53.3		+1.6
Female	25.6		25.5		-0.1

Source: Census of India, 2011.

Note: The 2011 census of India is the latest available census data. The decennial 2021 census is postponed due to COVID-19 and has yet to be conducted.

0.1% between 2001 and 2011. However, the gender gap in work participation rate between males and females has increased by 0.9% only, suggesting almost stagnant employment.

From Table 2, it is observed that the female work participation rate is lower than the male work participation rate in all states in India, and the scenario is much worse in urban areas. The women's work participation rate in Bihar is pathetic, whereas Himachal Pradesh seems to be a much more women-friendly labour market.

Table 2: Work Participation Rate in India, 2011-2012

State	*Rural*			*Urban*		
	Male	*Female*	*Total*	*Male*	*Female*	*Total*
Andhra Pradesh	61.2	44.8	52.8	57.6	18.0	38.0
Assam	56.4	12.9	35.9	57.3	9.7	34.8
Bihar	48.7	5.8	28.4	44.1	5.4	26.7
Gujarat	60.2	27.9	44.8	60.7	13.5	38.7
Haryana	53.2	16.4	36.5	53.5	10.2	33.1
Himachal Pradesh	54.7	52.9	53.8	61.2	23.6	43.3
Jammu & Kashmir	55.9	26.3	41.5	56.3	14.5	36.2
Karnataka	62.0	28.9	45.4	59.4	17.1	38.8
Kerala	58.3	25.8	41.0	56.7	22.2	38.6
Madhya Pradesh	56.4	23.9	40.7	53.3	11.9	33.4
Maharashtra	58.2	38.9	49.0	56.0	17.2	37.4
Orissa	60.6	25.1	42.7	60.3	15.8	39.5
Punjab	57.9	23.7	41.4	58.6	14.1	37.9
Rajasthan	50.0	34.9	42.7	50.7	14.4	33.6
Tamil Nadu	60.7	38.6	49.5	59.9	21.1	40.3
Uttar Pradesh	49.6	17.8	34.1	53.3	10.6	33.1
West Bengal	60.2	19.4	40.0	63.0	18.6	41.9
All India	55.3	25.3	40.6	56.3	15.5	36.7

Source: NSSO 68th Round (July 2011–June 2012)

Table 3: Average Wage/Salary Earnings per day by Regular Employees, 2011-12 (in ₹)

State	*Rural*				*Urban*			
	Male	*Female*	*Relative wage gap, 2011-12*	*Relative wage gap, 2009-10*	*Male*	*Female*	*Relative wage gap, 2011-12*	*Relative wage gap, 2009-10*
Andhra Pradesh	251.28	225.01	10.45	52.6	427.82	244.30	42.90	27.25
Assam	343.97	179.71	47.75	61.7	615.23	561.63	8.71	22.43

State	*Rural*				*Urban*			
	Male	*Female*	*Relative wage gap, 2011-12*	*Relative wage gap, 2009-10*	*Male*	*Female*	*Relative wage gap, 2011-12*	*Relative wage gap, 2009-10*
Bihar	450.49	188.42	58.17	7.60	417.10	369.02	11.52	4.84
Gujarat	268.69	173.13	35.57	5.05	326.34	271.86	16.69	27.79
Haryana	396.44	357.38	3.04	32.46	810.93	635.59	21.62	44.22
Himachal Pradesh	434.72	250.69	42.33	37.58	426.03	306.55	28.04	10.62
Jammu & Kashmir	453.56	222.37	50.97	2.2	497.61	484.71	2.59	15.22
Karnataka	237.53	151.85	36.07	78.15	518.58	391.97	24.41	24.32
Kerala	386.44	240.45	33.12	26.6	519.84	412.47	20.65	29.30
Madhya Pradesh	270.94	108.56	59.93	10.32	459.66	320.58	30.26	28.86
Maharashtra	369.14	306.76	16.90	43.99	516.55	370.30	28.31	29.16
Orissa	245.30	223.23	9.00	48.36	457.66	286.42	37.42	10.83
Punjab	302.79	157.61	47.95	48.12	352.58	399.38	-13.27	33.55
Rajasthan	328.61	177.86	45.88	56.8	417.14	412.89	1.02	8.57
Tamil Nadu	292.55	199.44	31.81	37.05	420.76	297.63	29.26	15.11
Uttar Pradesh	296.51	171.27	42.24	37.0	496.53	378.00	23.87	13.26
West Bengal	297.35	119.76	59.72	46.05	454.61	323.56	28.83	29.73
All India	322.28	201.55	37.46	37.05	469.87	366.15	22.07	18.13

Source: NSSO 66th Round (July 2009–June 2010) and 68th Round (July 2011–June 2012)

Table 3 presents the average wage received by regular wage and salaried employees in India during 2011-12. The wages for males and females differ substantially in rural and urban areas. Overall, females earn lower wages compared to men in both rural and urban areas. The relative wage gap has been calculated to show the comparable difference in wages between the sexes. This shows that there is a prominent wage gap between male and female workers. In every state, there is a significant level of wage difference between male and female workers. Madhya Pradesh has the highest level of wage gap between males and females in rural areas, whereas Andhra Pradesh has the highest difference in gender wages in urban areas. Also, the wage gap is relatively low in urban areas compared to rural areas. Surprisingly, the relative wage gap is negative 13.27% in urban Punjab in 2011-12, dropping from a positive 33.55% gender earnings gap in 2009-10, while there is 48% relative wage gap in rural Punjab. At the all-India level, during 2011-12, women in rural India have been earning more than one-third less and in urban India one-fifth less

per day relative to the respective per-day earnings of men. While the relative wage gap is almost the same between men and women in rural areas, the urban gender wage gap has increased between 2009-10 and 2011-12.

Table 4 shows the relative wage gap with respect to the educational levels. The gender gap remains very high among workers without formal education and with lower educational standards, who are mostly working in informal activities. While the gender gap declines with educational attainments in rural areas, except at higher education levels, there has been a rise in the wage gap for all educational levels in urban areas, except in the case of secondary and higher secondary levels. Moreover, the relative wage gap has widened between 2009-10 and 2011-12 in almost all cases, indicating rising income inequality and differential returns to higher education.

Table 4: Relative Gender Wage Gap by Education Level in India (%)

Education level	*Rural*					*Urban*				
	1993-94	*1999-2000*	*2004-05*	*2009-10*	*2011-12*	*1993-94*	*1999-2000*	*2004-05*	*2009-10*	*2011-12*
Not literate	42.5	43.4	50.7	51.8	48.78	73.0	40.9	50.7	40.9	40.55
Literate upto middle	47.9	45.2	51.6	49.8	48.50	76.1	38.7	41.9	37.8	44.02
Secondary and higher secondary	20.3	14.9	36.6	43.3	43.66	13.3	13.3	17.6	19.0	14.38
Graduate and above	26.1	27.6	36.0	29.0	31.33	29.5	29.5	26.6	21.3	24.03

Source: NSSO: Employment and Unemployment Situation in India (various rounds)

The gender wage gap mainly occurs when both males and females are in the same industry or in the same occupation, where women are being paid less compared to men. The observed wage gap is generally explained on the basis of occupational differences, educational qualification, industry, religion and community, marital status, etc. Men and women, on average, work in different industries and occupations, and this differential occupation pattern may explain some portion of the wage gap. The hours of work also differ between males and females, which may also account for some portion of the gender wage gap in earnings. Generally, women work fewer hours relative to men. The gender wage gap can go beyond education level, occupation, experience, etc. According to Blau and Khan (2017), the wage gap can also be due to some of the factors that can not be observed but are evident.

One such unobserved factor that prompts the observed gender wage gap is discrimination. Discrimination takes place when one group (males) are paid a higher wage rate than females for the same level of work in the same occupation, with similar education and skills. According to the neoclassical approach to labour market discrimination, wage differences arise among equally productive workers not based on the objective criteria of productive characteristics, but on the basis of unproductive characteristics like colour, sex or caste. Such non-market group-specific characteristics are more valued in the market, and the values attached to these unproductive characteristics are determined by the preferences of the employer. For Becker (1971), who pioneered the work on (racial) discrimination, discrimination is practised because employers have a 'preference' or 'taste' for discrimination. The existence of this 'taste for discrimination' explains the process whereby people who, in some sense, are equal do not receive the same treatment in economic transactions.

The issue of gender discrimination in the labour market in the form of the gender wage gap and low labour force participation for women is a well-researched area in labour economics. Gender wage discrimination takes place between the male and female groups based on various factors. Women workers get discriminated against more in terms of wages if they also belong to socially and economically backward castes and religious minorities. Lakshmanasamy and Madheswarn (1995) observed the existence of caste discrimination even in the Indian scientific labour market to the extent of 74% earnings disadvantage for the SC personnel compared to the non-SC personnel. Using the 1981 Degree Holders and Technical Personnel survey and the Blinder-Oaxaca decomposition method, Lakshmanasamy and Madheswaran (1996) showed that there exists gender discrimination in the Indian labour market to the extent of 28%. Subsequent studies also demonstrated the continued existence of gender differences in wages and the persistence of sex discrimination in the labour market of India. Thus, the gender wage gap continues to persist in the Indian labour market, despite improvements in women's empowerment, labour market regulations and globalisation facilitating much mobility. The wage gap is partly due to some observable factors and some discriminatory factors. Therefore, it is important and also interesting to see how far discrimination against females operating through the gender wage gap has changed/declined since 1981 in the Indian labour market.

Thus, the main objective of this paper is to examine the extent of gender wage discrimination in the Indian labour market. The paper analyses the gender wage gap, using the India Human Development Survey (IHDS) 2011-2012 data, applying the Heckman two-step estimation (Heckit) procedure of wage estimation, and the Oaxaca-Blinder decomposition method to quantitatively measure the extent of wage discrimination. In the empirical estimation of male and female wage equations, two econometric techniques are followed: the first step is the ML probit estimation of selection (labour force participation) equation, and the second step is the OLS estimation of the wage equation using the participants only, including the moral hazard (the inverse mills ratio) estimated from the first step as an additional regressor. Then, the wage gap between the sexes is decomposed as one part due to the observed labour market characteristic differences and the other part due to unexplained differences between the sexes, normally attributed to discrimination against females.

LITERATURE REVIEW

There are only a handful of econometric studies that analyse the gender wage gap in India using nationwide data. The empirical analysis of the extent of discrimination, especially the wage gap between the sexes, relies heavily on the Oaxaca-Blinder decomposition method. Navarro (2015) analysed the gender wage gap in the Dominican Republic using the 2013 Labour Force National Survey. The key determinants of wages were age, education, type of occupation, marital status and number of children. The empirical results show that there exists a gender wage gap in the Dominican Republic, and there is discrimination against women. Higher education narrows the degree of the gender wage gap; the hourly wage for college-educated women is 8% less than the hourly wage for college-educated men, compared with 24-32% less in other levels of education.

Sengupta and Das (2014) analysed gender wage discrimination across social and religious groups in India using the 50th and 66th rounds of the NSSO data. Heckman's two-step estimation model was used to estimate the gender wage gap. It was observed that the probability of women's participation rate in wage employment was lower than that of men's participation. However, the difference has been reduced over time. Hindu women workers were in a better position than Muslim women workers, partly due to religious customs

and partly due to religious discrimination. The study also shows that female workers were subject to stark discrimination in terms of participation in the regular/salaried job market as well as wages. Further, discrimination was more severe for women from minority religious and tribal groups.

Khanna (2012) examined the gender wage gap using wage data for regular wage workers from the NSSO employment-unemployment data for 2009-2010 and the Oaxaca-Blinder decomposition and the quantile regression methods. The focal point of the analysis was to look at log wage gaps at different points of the distribution, not just at the mean. The main finding is that of a sticky floor effect, that is, the phenomenon of declining gender log wage gaps across the quantiles. The Oaxaca-Blinder decomposition shows the contribution of discrimination to the gender log wage gap of at least over 80%. The quantile regression technique shows that the factors that contribute to the gender wage gap, however, are not the same at the bottom and the top of the distribution. Further, being married actually has a wage-reducing effect for high-wage-earning women. The decreasing coefficients still mean that the wage dispersion among women falls as the proportion of married women increases, as low-wage earners see an increase, while the high-wage earners see a reduction in wages.

Sarkhel (2012) estimated the wage gap in the Indian labour market and gender bias, using the 2004-05 IHDS data. The study explains the payment gap by integrating the labour market performance with that of family decision-making practice. The Mincer wage functions estimated incorporate patriarchy as one of the explanatory variables. The Oaxaca-Blinder method was used for decomposing inequality in male-female hourly wage earnings into the contributing factors, such as observed and unobserved. The results show that women from patriarchal families are earning less than men due to the severe participation constraint.

Deshpande et al. (2018) decomposed gender wage gaps not only at the mean but also along the entire wage distribution by applying the quantile regression method, using the 55th (1999-2000) and 66th (2009-10) rounds of NSSO data. The estimated gender wage gap was 55% in 1999-2000 and 49% in 2009-10. It is shown that though the proportion of educated women among women is higher than that among men, the gender wage gap is almost 38%. Further, they observed that between the two periods, the gender wage gap at the median has declined from 76% to 53%.

In short, most studies indicate that there exists a significant gender wage gap between male and female groups in the labour market. This is so despite marked improvements in both education and labour force participation of women relative to men. Given that the gender wage gap is common in the labour market, the factors that have been used to analyse the gender wage gap are mainly age, education, community, religion, marital status and occupation. To analyse the gender wage gap and the contribution of discrimination to the male-female wage gap, the most commonly used method is the Oaxaca-Blinder decomposition technique, which identifies the discriminatory part of the gender wage gap.

DATA AND METHODOLOGY

To align with the latest available census data (2011 Census of India), the data from the Indian Human Development Survey 2011-12 (IHDS-II) was used to analyse the gender wage gap and the decomposition of the male-female earnings differentials to measure the extent of discrimination in the Indian labour market. The IHDS survey, by the NCAER and the University of Maryland, is a nationally representative survey consisting of 41554 households spread across 1503 villages and 971 urban neighbourhoods. The sample is spread across 33 States and UTs of India, except Andaman and Nicobar Islands and Lakshadweep. The sample size of the data is 204568, out of which 102062 are males and 102506 are females. A detailed discussion of the methodology and survey can be found in the IHDS technical report (Desai et al., 2010). The data is a wealth of information on individual age, income source, employment, wages, occupation, education level, educational expenditure, earnings, number of household members, residence, house ownership, etc. The data contain two categories of income source, the salary employed and self-employed, and wages are reported only for the salaried employees. There are 131,365 wage-employed individuals and 73,150 self-employed individuals in the data. In the empirical analysis, the income source, a dichotomous variable, was used as the dependent variable in the selection equation estimated by probit regression, and the log of wage was used as the dependent variable in the wage equations for males and females.

Following Mincer (1974), the wage equation was estimated separately for men and women. A statistical issue that arose during the estimation of the

wage equations was that in the data, there were households that did not have wage income. Some households within the sample and some members within a household received no wage as they selected themselves not to work and hence reported no wage income. When empirical estimation is carried out, ignoring the households with no wage income, then the samples become non-random or incidentally truncated, and the problem of sample selection bias arises.

In order to overcome this selectivity bias in wage estimation, Heckman (1976, 1979) proposed a two-step estimation (Heckit) procedure. The first step is to estimate the selection equation, the labour market participation decision model, by the maximum likelihood estimation of a probit equation. Then, in the second step, the wage equations are estimated by the ordinary least squares method, incorporating the inverse Mills ratio obtained from the first step estimation as an explanatory variable. This way, the self-selection bias was taken care of. Then, the gender wage gap between the two groups was decomposed by the Oaxaca-Blinder decomposition method, as explained and unexplained parts, and the unexplained part was treated as discrimination against women.

Heckman Two-Step Estimation Method

The wage equation and selection equation included in the model are:

$$\text{Wage equation: } \ln W_i = X_i\beta + \varepsilon_i \quad (1)$$

$$\text{Ln(wage)} = \beta_0 + \beta_1 \text{ Age} + \beta_2 \text{ Age Square} + \beta_3 \text{ Education} + \beta_4 \text{ Religion} + \beta_5 \text{ Community} + \beta_6 \text{ Own house} + \beta_7 \text{ Married} + \beta_8 \text{ Residence} + u_i \quad (2)$$

$$\text{Selection equation: } Y_i^* = \gamma Z_i + u_i, \quad (3)$$

where $Y_i = 1$ if $Y_i^* > 0$

$Y_i = 0$ otherwise

The selection equation is a latent model, where the Z vector is not observed, and the latent dependent variable represents the individual evaluation of work choice.

$$\text{Employed} = \beta_0 + \beta_1 \text{ No. Persons} + \beta_2 \text{ Education} + \beta_3 \text{ Religion} + \beta_4 \text{ Community} + \beta_5 \text{ Own house} + \beta_6 \text{ Married} + \beta_7 \text{ Residence} + \varepsilon_i \quad (4)$$

where Employment status = 1, if salary employed

= 0 if self-employed

In the second step of the model, the estimated inverse Mills ratio (λ) from the selection equation was included in the estimation of the wage equation as an additional regressor to correct for the selectivity bias. The significance level of the estimated inverse Mills ratio implies whether there is sample selection bias or not. The inverse Mills ratio is the ratio of the probability density function to the cumulative density function of a distribution. If Y is a random variable distributed normally with mean μ and variance σ^2, then it is possible to show that,

$$E(Y \mid X > \alpha) = \mu + \sigma[\{\phi(\alpha - \mu)/\sigma\} / \{1 - \Phi(\alpha - \mu)/\sigma)\}] \quad (5)$$

where α is a constant, ϕ denotes the standard normal density function, and Φ denotes the standard normal cumulative distribution function. The selectivity bias-corrected estimating wage equation is then given by,

Wage equation: $\ln W_i = X_i\beta + \sigma\lambda + \varepsilon_i$ (6)

where $\lambda = \frac{-\phi(\hat{\beta}' X)}{\Phi(\hat{\beta}' X)}$ if Y = 1 and $\lambda = \frac{\phi(\hat{\beta}' X)}{\{1 - F\Phi(\hat{\beta}' X)\}}$ if Y = 0 (7)

Oaxaca-Blinder Decomposition Method

The Oaxaca-Blinder decomposition method was used to analyse the wage gap between male and female groups. To estimate the gender wage gap between the two groups, the wage functions of men and women were estimated separately. The estimating wage functions are specified as,

$$\ln W_i^m = X_i^m \beta^m + \varepsilon_i^m \quad (8)$$

$$\ln W_i^f = X_i^f \beta^f + \varepsilon_i^f \quad (9)$$

where the subscript m denotes male and f denotes female. The average log wage gap is denoted as,

$$\overline{\ln w_m} - \overline{\ln w_f} = \overline{X_m}\beta_m - \overline{X_f}\beta_f \quad (10)$$

The Oaxaca-Blinder decomposition of the wage gap is,

$$\overline{\ln w_m} - \overline{\ln w_f} = (\overline{X_m} - \overline{X_f})\beta_m + (\beta_m - \beta_f)\overline{X}_f \quad (11)$$

The wage gap was decomposed by imposing the returns (β) for characteristics of one group on the other group's characteristics. That is, $\overline{X_m}\beta_m - \overline{X_f}\beta_f$

explains the observed wage difference for the differential characteristics of the male and female groups. The unexplained wage difference, $\overline{X_m}\beta_m - \overline{X_f}\beta_m$ obtained by imposing the male returns on the female characteristics is usually attributed to discrimination. This difference shows the unjustified part of the observed wage gap due to some observed factors, primarily discrimination, as well as some personal preferences or similar factors.

Probit Model Estimation

The probit model is a popular specification for an ordinal or a binary response model. The probit model is a type of regression where the dependent variable is a qualitative and dichotomous variable which can only take two values, either the presence or the absence of an attribute, for example, working or not working. The purpose of the model is to estimate the probability that an observation with particular characteristics will fall into a specific alternative, i.e. one of the categories. Moreover, if the estimated probability is greater than ½ is treated as classifying an observation into a predicted category, the probit model is a type of binary classification model. The probit model, which employs a probit link function, is most often estimated using the standard maximum likelihood procedure. The behaviour of the dichotomous dependent variable needs a suitable cumulative distribution function, and the probit regression emerges from a normal cumulative distribution function.

The normal cumulative distribution function (CDF) is the function where a variable Y follows the normal distribution with mean μ and variance σ^2. Then the probability density function (PDF) for Y is given by,

$$\phi(Y) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(Y-\mu)^2/2\sigma^2} \tag{12}$$

and its cumulative distribution function is,

$$\Phi(Y) = \int_{-\alpha}^{Y_0} \frac{1}{\sqrt{2\sigma^2\pi}} e^{-(Y-\mu)^2/2\sigma^2} \tag{13}$$

where Y_0 is some specified value of Y. Consider the following regression model,

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{14}$$

where X is a vector of explanatory variables and Y=1 if the individual works in salaried employment and 0 if the individual does not work in salaried

employment. Assume that the decision of the i^{th} person to go for salaried employment or not depends upon the unobservable utility index I_i (the latent variable), which is determined by a set of explanatory variables X_i in such a way that the larger the value of the index I_1, the greater the probability of going for salaried employment. The index I_i is expressed as,

$$I_i^* = \beta_1 + \beta_2 X_i \tag{15}$$

Let Y=1 if a salary employed and Y= 0 if not wage employed. It is reasonable to assume that there is a critical or threshold level of the index, I_i^*, such that if I_i exceeds I_i^*, the person is wage employed, otherwise not wage employed. The threshold I_i^*, like I_i, is not observable, but if we assume that it is normally distributed with the same mean and variance, it is possible to not only estimate the parameters of the index given in the equation $I_i = \beta_1 + \beta_2 X_i$ but also to get some information about the unobservable index itself.

Given the assumption of normality, the probability that I_i^* is less than or equal to I_i can be computed from the standardised normal CDF as,

$$P_i = P(Y=1|X) = P(I_i^* \leq I_i) = P(Z_i \leq \beta_1 + \beta_2 X_i) = \Phi(\beta_1 + \beta_2 X_i) \tag{16}$$

where P(Y=1|X) means the probability that an event occurs given the values of X, and Z_i is the standard normal variable, i.e. $Z \sim N(0, \sigma^2)$. Φ is the standard normal CDF, which can be written explicitly in the present context as,

$$\Phi(I_i) = \frac{1}{2\pi}\int_{-\alpha}^{I_i} e^{-z^2/2} dz \tag{17}$$

$$\Phi(I_i) = \frac{1}{2\pi}\int_{-\alpha}^{\beta_1+\beta_2 X_i} e^{-z^2/2} dz \tag{18}$$

where P represents the probability that an event will occur, is measured by the area of the standard normal curve from – α to I_i. To obtain information on I_i, the utility index on β_1 and β_2, take the inverse of the above equation to obtain

$$I_i = \Phi^{-1}(I_i) = \Phi^{-1}(P_i) = (\beta_1 + \beta_2 X_i) \tag{19}$$

where Φ^{-1} is the inverse of the normal CDF. The estimating probit model would be,

$$Y_i = 1 \text{ if } I_i^* = X_i\beta + \varepsilon_i > 0 \tag{20}$$
$$= 0 \text{ otherwise}$$

EMPIRICAL ANALYSIS

The focus of the paper is to estimate the extent of the gender wage gap and discrimination between males and females in India. Table 5 presents the earnings of the male and female workers in each sector (self-employed and wage-employed), and there is a difference of almost 22413.3 and 40415.4 for self-employed and wage-employed, respectively.

Table 5: Gender Wage Gap (₹/annum)

Employment	Male earnings	Female earnings	Earnings gap
Wage-employed	76085.9	35670.5	40415.4
Self-employed	39687.5	17274.2	22413.3

Table 6 presents the mean wage earnings between males and females for different characteristics. On average earnings, by every characteristic, males have higher earnings compared to females. Therefore, it is evident that there is a wage gap between males and females in every respect; the wage gap is almost 20-30 per cent between male and female groups. This shows that the gender wage gap is prominent in every aspect of the labour market.

Table 6: Average Earnings of Males and Females for Characteristics (₹/annum)

Characteristics	*Description*	*Wage earnings*		*Log of wage earnings*	
		Male	*Female*	*Male*	*Female*
Education	Illiterate and primary	41663.5	15718.4	9.09	9.13
	Secondary and above	113006.7	90918.3	11.03	11.0
Religion	Hindu	62838.3	25694.3	10.40	9.31
	Non-Hindu	66933.1	39849.9	10.58	9.70
Community	General	91575.1	55660.3	10.80	9.91
	OBC, SC or ST	54355.5	22083.3	10.32	9.25
Marital status	Married	68973.7	25845.2	10.51	9.29
	Unmarried	46722.0	32261.0	10.21	9.54
Residence	Owning house	60606.0	25803.5	10.02	9.31
	Not owning a house	97895.0	53264.1	10.99	10.02
Region	Urban	103440.6	65485.4	11.11	11.11
	Rural	42244.6	15950.0	10.07	9.07

Table 7 presents the descriptive statistics of the variables that were used in the empirical analysis. The overall mean earnings are ₹52608 per year. The average wage gap between males and females based on their average earnings

is 36003. The logarithmic difference in earnings between the genders is 1.08. This wage gap between male and female workers shows that there are chances of discrimination in the market by gender. About 64% is the proportion of individuals going from salary-employed work to self-employed work. The overall mean value of the education levels indicates that 0.27 is the proportion of males who have more than primary education, compared to 0.19 the proportion of females with secondary education. Similarly, 0.80 and 0.70 are the proportions of Hindu and backward households, 0.92 is the proportion of households owning a house, and 0.48 is the proportion of married households. Similarly, 0.33 is the proportion of urban households.

Table 7: Descriptive Statistics of Variables

Variable	*Description*	*Mean (standard deviation)*		
		All	*Males*	*Females*
Wage	Wage earnings (₹/annum)	52608.2	63605	27602
Ln(Wage)	Ln (Wages) (₹/annum)	10.111 (1.313)	10.44 (1.19)	9.363 (1.252)
Employment	Wage-employed = 1 Self-employed = 0	0.642 (0.479)	0.642 (0.479)	0.641 (0.479)
Age	Age (yrs)	29.82 (20.36)	29.344 (20.33)	30.29 (20.38)
Age square	Age squared (yrs)	1303.9 (1525.2)	1274.5 (1509.4)	1333.3 (1540.2)
Education	Secondary and above = 1 Illiterate and primary = 0	0.2291 (0.420)	0.2692 (0.4435)	0.1893 (0.39)
Religion	Hindu = 1, others = 0	0.8010 (0.399)	0.8008 (0.3993)	0.8012 (0.3990)
Community	Community: OBC, SC or ST = 1, General = 0	0.7076 (0.454)	0.7067 (0.4552)	0.7086 (0.4543)
Own house	Owning house = 1, Not owning house = 0	0.9280 (0.2584)	0.9280 (0.2584)	0.9279 (0.2585)
Marriage	Married =1, Unmarried = 0	0.4816 (0.499)	0.4692 (0.4990)	0.4940 (0.4999)
Residence	Urban =1, Rural = 0	0.3394 (0.473)	0.3417 (0.4743)	0.3372 (0.4727)
N	Sample size	204568	102062	102506

HECKMAN TWO-STEP ESTIMATES

Table 8 presents the probit estimates of the decision to work for wages, along with marginal effects. The estimated coefficients of probit estimation have no direct interpretation, but the marginal effects have a direct interpretation as the effect of an independent variable on the dependent variable. The marginal effects show that the number of persons in a household has a negative impact on wage employment decisions. The lower the family size, the lower the chance for participation in wage employment (salaried work). One extra individual in the household reduces the probability of a male being in salaried work by 0.62 percentage points and 0.46 percentage points for a female. Similarly, the probability of being in salaried work by a male or a female with secondary and higher education increases by 0.08 for both. This shows that one extra level of education increases the number of males and females working for a salary in the labour force by 0.28 percentage points.

The probability of an individual who is Hindu going for salaried work is negative, and higher for males than for females. The marginal effect is almost equally negative at 0.63 and 0.65 percentage points for males and females, respectively, whereas the probability increases for individuals from other religions to participate in the labour market. Also, the individuals who are from SC, ST and OBC communities have a lower probability of salaried work, and the marginal effect of these communities on labour market participation is lower by 12% and 16% for males and females, respectively. Similarly, the probability of wage work of individuals owning a house is negative, and it reduces the probability of salaried work by 0.23 percentage points. The probability for an individual who is married to work for wages is also negative, and the marginal effect is lower for women than for men. The urban residence significantly increases the probability of labour force participation, with almost a 33 percentage point increase for both males and females. Overall, from the probit estimates of Table 7, both males and females have almost equal probability of labour market participation and being wage earners or salaried employees.

Given the more or less equal labour market participation, the next question to be answered is why there is such a noticeable gender wage gap. To understand this, the wage equations for males and females were estimated separately, which is the second step of the Heckman estimation. Table 9 presents

Table 8: Probit Estimates of Selection Equation with Marginal Effects

Variable	*Employment status (dependent variable)*			
	Male		*Female*	
	Coefficient	*Marginal effect*	*Coefficient*	*Marginal effect*
No. persons in the household	-0.174* (11.46)	-0.0062	0.0128* (8.59)	-0.0046
Secondary and above education	0.0800* (7.80)	0.0285	0.0794* (6.76)	0.0283
Hindu	-0.1816* (16.21)	-0.0635	-0.1870* (16.75)	-0.0656
OBC, SC or ST community	-0.0399* (3.46)	-0.0121	-0.0451* (4.63)	-0.0162
Own house	-0.8367* (6.04)	-0.2375	-0.8196* (6.09)	-0.2353
Married	-0.0521* (33.64)	0.0187	-0.0139* (28.60)	-0.0047 (-28.60)
Urban	1.0434* (10.47)	0.3340*	1.0277* (18.77)	0.3301
Constant	1.1515* (19.08)	-	1.1129* (18.58)	-
LRchi2	16440.46		16024.35	
Prob > chi2	0.0000		0.0000	
Pseudo R-square	0.1241		0.1202	
Log likelihood	-58022.048		-58627.758	
Oservations	101593		102092	

Note: Absolute t-values in parentheses. * Significant at 1% level.

the OLS estimates of wage equations of males and females with and without selectivity bias correction. The statistical significance of the inverse Mills ratio (λ) shows the presence of sample selection bias in the OLS estimation. The inverse Mills ratio is negatively signed, suggesting that the error terms in the wage equations and the selection equation are negatively correlated, and there is sample selection bias in the estimation. The estimated values have reduced by 1 to 5 percentage points when the selectivity correction term is included, compared to the uncorrected wage estimates by the simple regression for both males and females. The wage earnings increase for both males and females with an increase in age by 8%. The negative sign on the age square shows that wage earnings increase at a decreasing rate.

Relative to illiterate or primarily educated individuals, secondary or higher education increases wage earnings of males by about 6% and female

earnings by 10%. This shows that the returns to education are higher for females compared to those of males. An individual of the Hindu religion has a negative wage earnings of 14% compared to non-Hindus. Similarly, belonging to SC, ST or OBC communities has a negative impact on the wage earnings of both male and female workers, the marginal effect being 0.17 and 0.15 for males and females, respectively. Thus, being forward or in the general community results in a positive increase in the wage earnings of both males and females. A married male will have a positive increase in the wage earned by 0.05, whereas for females, being married will negatively affect the wage earnings by 0.42, implying women working after marriage

Table 9: OLS Estimates of Wage Equations by Gender
Dependent variable: Log (wage)

Variable	*Male*		*Female*	
	Selectivity bias uncorrected	*Selectivity bias corrected*	*Selectivity bias uncorrected*	*Selectivity bias corrected*
Age	0.078* (30.31)	0.077* (30.11)	0.080* (22.0)	0.076* (21.25)
Age square	-0.0008* (30.09)	-0.0008* (29.99)	-0.0008* (19.86)	-0.0008* (19.39)
Secondary and above education	0.585* (47.84)	0.572* (41.55)	1.074* (42.79)	1.010* (36.99)
Hindu	-0.154* (10.98)	-0.149* (10.04)	-0.184* (7.47)	-0.145* (5.68)
OBC, SC or ST community	-0.179* (13.84)	-0.178* (13.80)	-0.157* (6.64)	-0.155* (6.59)
Own house	-0.137* (6.72)	-0.088* (2.84)	-0.223* (6.52)	0.026 (0.48)
Married	0.101* (6.15)	0.055** (2.06)	-0.217* (10.48)	-0.425* (10.29)
Urban	0.828* (69.12)	0.761* (22.63)	0.802* (37.31)	0.501* (8.94)
Constant	8.792* (174.15)	8.915* (115.80)	8.039* (102.1)	8.576* (70.83)
Inverse Mills ratio (λ)	-	-0.106* (2.13)	-	-0.472* (5.83)
Prob>F	0.00	0.00	0.00	0.00
R-square	0.24	0.27	0.25	0.29
Number of observations	36925		16284	

Note: Absolute t-values in parentheses; * Significant at 1% level.

have lower labour force attachment. Owning a house has a negative effect on the wage earnings of males and a positive effect on female work earnings. Also, being in an urban area positively increases the wage earnings for both males and females, by 0.76 and 0.50, respectively, as the opportunity for salaried work is higher in urban areas.

Thus, from the Heckman two-step estimation, it is observed that both males and females have an equal probability of being wage or salaried workers, but the returns for wage determinants in the labour market are not equal between the sexes. The returns for males are somewhat higher than those for females. Therefore, here we used the Oaxaca-Blinder decomposition method to find the gender wage gap between male and female groups. The Oaxaca-Blinder decomposition method analyses how much of the gender wage gap is due to the explained factors and due to the unexplained factors, which is the discrimination part.

DECOMPOSITION OF THE GENDER WAGE GAP

The Blinder-Oaxaca decomposition method decomposes the gender gap in the mean earnings between the two groups into that part that is due to the group differences in the wage-determining characteristics of males and females, and another part that is due to the differential returns to the same wage-determining characteristics. Thus, the total log wage gap of 1.072 between the sexes is decomposed into the explained and the unexplained part, together constituting the gender wage gap in the labour market.

Table 10 explains the decomposition results. The positive results express how much the gender wage gap will be reduced if males and females have equal returns for those attributes, and the negative numbers indicate how much the gender wage gap will fall if females face the same wage functions as males. The unexplained part is generally attributed to gender wage discrimination. Almost 74% of the mean wage gap between males and females is justified by the differences in male and female wage-determining characteristics, 26% of the unexplained part captures gender discrimination, and also the plausible unobserved gender differences in productivity and tastes in the labour market. Thus, it is evident that there is a sizable gender wage gap, and nearly one-quarter of this wage gap can be attributable to the discriminatory rewards against females in the Indian labour market.

Table 10: Decomposition of the Gender Wage Gap

Total log of wage gap	1.072
Log of mean wage gap attributable to:	
Age	-.077
Age square	0.047
Education	0.046
Religion	0.00006
Community	0.023
Own house	-0.0001
Marital status	0.002
Urban residence	0.003
Constant	0.7524
Explained (justified) part (Characteristic difference)	0.7994 (74%)
Unexplained (unjustified) part (Discrimination)	0.2632 (26%)

CONCLUSION

The paper analyses the extent of gender wage discrimination in the Indian labour market against females. Taking the log of wage earnings by sex, using the 2011-12 India Human Development Survey, the observed gender wage gap was decomposed into two parts by applying the Oaxaca-Blinder decomposition method. As labour market participation is not likely to be random due to self-selection in labour force participation, the wage equations have been estimated using the Heckman two-step estimation method. First, the probability of labour market participation was estimated using the total sample by the probit model, obtaining the inverse Mills ratio. In the second step, wage equations for males and females were estimated separately by the OLS method, including the inverse Mills ratio as an explanatory variable to correct for selection bias.

The study reveals that the probability of women's participation in wage employment is lower than male participation. Women have higher levels of education than men, but their wages and the probability of going to work are lower compared to those of men. Women workers with higher levels of education are found to have a higher wage rate than women with lower education levels. Women of the socially backward community have low wage earnings compared to men. Women workers who are Hindu have significantly lower wages than men. Women who are from SC, ST and OBC communities also

hurt their wage earnings. Similarly, married women also earn less, but married males earn more, which can also be because the labour force participation of married women is generally low. Women workers in urban areas have a high probability of going for salaried employment, but receive significantly lower wages than men. There is a high level of wage gap between males and females. Therefore, there is a sizable gender wage gap in the Indian labour market.

The Oaxaca-Blinder decomposition of the gender wage gap revealed that there is a significant 26% wage gap that is due to unexplained factors. This part is generally attributed to discrimination, besides social and personal factors, and differential tastes and preferences. The 74% of the wage gap, which is explained by the variables, is that part of the gender gap that could be justified based on differentials in the wage-determining characteristics of males and females. Thus, the observed gender wage gap is partly the result of the discrimination that takes place in the Indian labour market. Compared to the estimated 28% wage discrimination in 1996, discrimination against females in 2011-12 is estimated to be 26%, suggesting only a marginal decline in the gender wage gap and labour market discrimination. Thus, the persistence of gender differences in labour force participation, wages, and gender wage gap reveals the continuing phenomenon of sex discrimination in India, even in the 21st century and in the age of globalisation and women's empowerment.

References

Becker, G. S. (1971). *The economics of discrimination*. University of Chicago Press.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and sources. *Journal of Economic Literature, 55*(3), 789–865. https://doi.org/10.1257/jel.20160995

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

Desai, S., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., & Vanneman, R. (2009). *India Human Development Survey: Design and data quality* (IHDS Technical Paper No. 1).

Deshpande, A. (2018). Bad karma or discrimination: Male–female wage gaps among salaried workers in India. *World Development, 102*, 331–344. https://doi.org/10.1016/j.worlddev.2017.10.011

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5*(4), 475–492.

Heckman, J. J. (1979). Sample selection as a specification error. *Econometrica, 47*(1), 153–161. https://doi.org/10.2307/1912352

International Labour Organization. (2018). *World employment and social outlook: Trends for women*. ILO.

Javeed, S., & Manuhaar, A. (2013). Women and wage discrimination in India: A critical analysis. *International Journal of Humanities and Social Science Invention, 2*(4), 6–12.

Kapsos, S. (2008). *The gender wage gap in Bangladesh* (ILO Asia-Pacific Working Paper Series). International Labour Organization.

Khanna, S. (2012). *Gender wage discrimination in India: Glass ceiling or sticky floor?* (Working Paper No. 214). Delhi School of Economics.

Lakshmanasamy, T., & Madheswaran, S. (1995). Discrimination by community: Evidence from the Indian scientific and technical labour market. *Indian Journal of Social Science, 8*(1), 59–77.

Lakshmanasamy, T., & Madheswaran, S. (1996). Occupational segregation and earnings discrimination by sex: Evidence from India. *Artha Vijnana, 38*(4), 372–386.

Mincer, J. (1974). *Schooling, experience and earnings*. Columbia University Press.

Mohanty, S. S., Bisht, M., & Mohapatra, P. (2014). Gender wage gap among salaried workers in India: Results of a principal component analysis of state-level statistics. *Economic Affairs, 59*(2), 175–190.

Navarro, A. G. (2015). *Gender wage gap and education: Case in the Dominican Republic* (All Graduate Plan B and Other Reports, Paper 485). Utah State University.

National Sample Survey Office. (2013). *Key indicators of employment and unemployment in India: 2011–12*. Ministry of Statistics and Programme Implementation, Government of India.

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

Ozcan, Y. Z., Ucdogruk, S., & Ozcan, K. M. (2003). Wage differences by gender, wage and self-employment in urban Turkey. *Journal of Economic Cooperation, 24*(1), 1–24.

Sarkhel, S. (2012). *Wage gap in labour market, gender bias and socio-cultural influences: A decomposition analysis for India* (Working paper). St. Xavier's College, Kolkata.

Sengupta, A., & Das, P. (2014). Gender wage discrimination across social and religious groups in India. *Economic and Political Weekly, 49*(21), 71–76.

Self-Employed Women's Association. (2000). *Labouring brick by brick: A study of construction workers*. SEWA.

Taniguchi, K., & Tuwo, A. (2014). *New evidence on the gender wage gap in Indonesia* (Economics Working Paper Series No. 404). Asian Development Bank.

Verick, S. S. (2018). *The puzzles and contradictions of the Indian labour market: What will the future of work look like?* (IZA Discussion Paper No. 11376). Institute of Labor Economics.