

Bridging Wage Gaps for Sustainable and Inclusive Growth: Evidence from India Using PLFS Microdata (2021–2023)

Anju Rani

*Associate Professor, Department of Economics
BPSMV, Khanpur Kalan, Sonapat*

ABSTRACT

This study quantifies the gender wage gap in India and identifies its principal determinants by combining a conventional econometric (Mincer/Oaxaca-Blinder) framework with machine learning prediction and explainable-AI methods. Using unit-level microdata from two rounds of India's Periodic Labour Force Survey (PLFS, 2021–22 and 2022–23), we restrict the sample to usual-status workers aged 15–64 with positive monthly earnings ($N = 203,578$; 159,050 men and 44,528 women). The raw gender gap in monthly earnings is 36.2% at the mean (37.7% weighted) and 53.6% at the median. A Mincer log-wage regression controlling for education, experience, marital status, social group, religion, sector, employment type, occupation, and industry yields a conditional gender coefficient implying a 46.0% gap — larger than the raw gap — while an Oaxaca-Blinder decomposition attributes only 7.4% of the total log-wage differential to differences in observable characteristics, leaving 92.6% (equivalent to a 47.4% wage penalty) unexplained. Six machine learning models are compared for wage prediction; XG Boost achieves the best out-of-sample fit (test $R^2 = 0.549$, RMSE = 0.537 log points), outperforming linear, ridge, lasso, random forest, and gradient boosting alternatives. Permutation importance and SHAP analysis both rank gender among the top two predictors of wages — ahead of education and occupation — and a model-based counterfactual exercise (predicting each woman's wage as if she were coded male, holding all other characteristics fixed) implies an average gender penalty of 36.3%, closely corroborating the classical estimates. This gap is largest among the self-employed (53.1%) and in rural areas (47.6%), and smallest among regular wage/salaried workers (24.8%) and urban workers (28.8%). A placebo test confirms the counterfactual method is not a modelling artefact: retraining on randomly permuted gender labels collapses the estimated gap to -0.1% . The convergence of three independent methods on a large, largely unexplained gender wage gap suggests that policies targeting only human-capital and occupational-composition gaps are unlikely to substantially close India's gender pay gap; complementary measures — enforcement of equal-pay provisions, formalization of self-employment and rural work, and scrutiny of within-occupation pay-setting — are indicated.

Keywords: Gender Wage Gap; India; PLFS; Machine Learning; XGBoost; SHAP; Oaxaca-Blinder; Mincer Equation; Labour Economics.

INTRODUCTION

Gender-based wage disparity remains one of the most persistent features of the Indian labour market despite decades of economic growth, rising female educational attainment, and legislative commitments to equal remuneration (the Equal Remuneration Act, 1976, subsumed into the Code on Wages, 2019). India's female labour force participation rate remains among the lowest in the G20, and among women who do participate, earnings consistently trail those of men across almost every segment of the labour market — rural and urban, formal and informal, self-employed and wage-employed. Understanding how much of this gap is attributable to differences in observable productive characteristics (education, experience, occupation, sector) versus a residual, unexplained component is central both to academic debates about labour market discrimination and to the design of effective policy.

This paper makes three contributions. First, it provides an updated empirical estimate of the gender wage gap in India using the most recent available rounds of the Periodic Labour Force Survey (PLFS), the government's principal source of labour market microdata since the survey's inception in 2017–18. Second, it triangulates the classical econometric toolkit (Mincer wage regression, Oaxaca-Blinder decomposition) against a battery of six machine learning models — linear regression, ridge, lasso, random forest, gradient boosting, and XGBoost — allowing us to assess whether more flexible, non-parametric models change the substantive conclusions about the size and drivers of the gap. Third, it uses SHAP (SHapley Additive exPlanations) values and a model-based counterfactual gender-swap exercise to provide an interpretable, non-parametric analogue to the Oaxaca-Blinder decomposition, and validates this approach with an explicit placebo test.

The remainder of the paper is organized as follows. Section 2 reviews the relevant theoretical and empirical literature. Section 3 describes the PLFS data and variable construction. Section 4 sets out the econometric and machine learning methodology. Section 5 presents results: descriptive statistics, the raw and conditional gender gap, model comparison, feature importance, and the counterfactual decomposition. Section 6 discusses the findings, Section 7 states limitations, Section 8 draws out policy implications, and Section 9 concludes.

LITERATURE REVIEW

Theoretical Framework

Human capital theory (Mincer, 1974; Becker, 1964) predicts that wages are determined by investments in education and (potential or actual) labour market experience, implying

that gender wage gaps should shrink as the education and experience of male and female workers converge. Where a gap persists after conditioning on these characteristics, discrimination-based theories — Becker's (1957) taste-based discrimination and Phelps' (1972) and Arrow's (1973) statistical discrimination models — offer competing explanations, alongside labour market segmentation theory, which points to occupational and sectoral segregation (women crowded into lower-paying occupations and the informal sector) as a structural driver independent of individual discrimination by any single employer.

Empirical Evidence

A substantial literature has used NSS Employment-Unemployment Surveys and, more recently, PLFS rounds to estimate India's gender wage gap via Mincer regressions and Oaxaca-Blinder decompositions (see, among others, studies summarized in ILO and World Bank India labour market reports). These studies consistently find raw gaps in the range of 30-50% depending on the earnings concept (daily wage, monthly earnings) and worker segment examined, with larger gaps typically found among casual and self-employed workers than among regular salaried employees, and larger gaps in rural than urban areas — a pattern our results reproduce (Section 5). Previous research provides substantial evidence of persistent wage inequality in India. Kumar and Mishra (2006) reported increasing skill-based wage inequality following economic liberalization. Mukherjee (2007) documented widening occupational and educational wage gaps during the post-reform period. Azam and Prakash (2015) found significant public-private wage differentials after accounting for worker characteristics. More recent studies by Gupta and Kothe (2022) and Gupta et al. (2024) reveal continuing caste- and gender-based wage discrimination in the Indian labour market, while Pandoh and Singh (2025) highlight substantial wage disparities among women across educational and occupational groups. International studies by Bourguignon (2004), Klasen (2010), Anand et al. (2013), and the OECD (2015) further establish that reducing inequality promotes productivity, social cohesion, and long-term economic growth.

Machine Learning Approaches to Wage Prediction and Discrimination

Mullainathan and Spiess (2017) and Athey and Imbens (2019) argue that machine learning methods are well suited to prediction tasks with rich, high-dimensional covariate sets but require care when used for causal or quasi-causal inference. Applications of gradient boosting and SHAP-based interpretability to wage and pay-gap analysis are increasingly common in other country contexts, but their systematic application to PLFS microdata specifically, and their direct comparison against the classical Oaxaca-Blinder benchmark on the same sample, remains limited in the published Indian literature. This is the methodological gap this paper addresses.

DATA AND VARIABLE CONSTRUCTION

Data Source and Sample

We use unit-level microdata from two consecutive rounds of India's Periodic Labour Force Survey: July 2021-June 2022 and July 2022-June 2023, obtained from the Ministry of Statistics and Programme Implementation (MoSPI). Person-level records (perv1) were merged with household-level records (hhv1) on the standard PLFS identification key (state, district, first-stage unit, stratum, sub-stratum, second-stage stratum, and sample household number). The pooled merged file comprises 848,037 person-level records.

The estimation sample is restricted to individuals aged 15-64 whose usual principal activity status (block 5.1 of the PLFS schedule) is self-employed (own-account worker, employer, or unpaid family worker), regular wage/salaried, or casual labour — i.e., usual-status workers, excluding those unemployed or outside the labour force. This yields 306,160 employed person-records, of which 203,578 (159,050 men; 44,528 women) report positive monthly earnings and form the core wage-analysis sample.

Variable Construction

The dependent variable is the natural log of monthly earnings, winsorized at the 0.5th and 99.5th percentiles to limit the influence of extreme values. Monthly earnings combine reported wage/salary income (for employees) and self-employment income/profit (for the self-employed and unpaid family workers), taking the maximum of the two fields since only one is generally non-zero for a given worker; self-employment losses (negative profit) are flagged separately and excluded from the positive-earnings wage sample.

Explanatory variables include: gender; age and age-squared; a Mincer-style potential experience proxy (age minus years of schooling minus six, floored at zero) and its square; general education level (an 11-category NSS classification from "not literate" to "postgraduate and above"); years of formal schooling; marital status; social group (Scheduled Tribe, Scheduled Caste, Other Backward Class, or Other/General); religion; rural/urban sector; employment type (self-employed, regular wage/salaried, or casual labour); a two-digit National Industrial Classification (NIC-2008) industry code; a broad (first-digit) National Classification of Occupations (NCO) occupation group; household size; and survey round. Descriptive statistics are reported in Table 1.

Monthly earnings, rather than an hourly wage, is used as the primary outcome: the PLFS fields governing weekly/usual hours of work, enterprise size, and detailed formal/informal employment sub-classification require verification against MoSPI's official schedule documentation before an hourly measure can be constructed with confidence, and this is left for a follow-up analysis (Section of Limitation). Sample weights (PLFS multipliers) are used for descriptive, population-representative statistics; regression and machine-learning models are estimated unweighted, with weighted specifications reported as a robustness check in future work.

METHODOLOGY

Descriptive and Raw Gap

The raw gender wage gap is computed as (mean male earnings - mean female earnings) / mean male earnings, reported at both the mean and median, and separately for the pooled sample and each employment-type/sector subgroup.

Mincer Log-Wage Regression

We estimate a standard Mincer-type log-wage equation by ordinary least squares with heteroskedasticity-robust (HC1) standard errors:

$$\ln(\text{Wage}_i) = \beta_0 + \beta_1 \cdot \text{Female}_i + \beta_2 \cdot \text{Education}_i + \beta_3 \cdot \text{Experience}_i + \beta_4 \cdot \text{Experience}_i^2 + \beta_5 X_i + \varepsilon_i$$

where X_i includes marital status, social group, religion, rural/urban sector, employment type, occupation group, industry, and survey-round fixed effects. β_1 gives the conditional gender wage gap, holding these observable characteristics constant.

Oaxaca-Blinder Decomposition

The total log-wage gap is decomposed into an "explained" component, attributable to differences in the average characteristics (endowments) of men and women, and an "unexplained" component, attributable to differences in the estimated returns to those characteristics between men and women (often interpreted, with appropriate caveats, as a discrimination-consistent residual). We use the pooled-coefficient reference approach (Neumark, 1988) to avoid the index-number problem of choosing either group's coefficients as the reference:

$$\ln(\bar{W}_M) - \ln(\bar{W}_F) = (\bar{X}_M - \bar{X}_F)' \beta^* + [\bar{X}_M'(\beta_M - \beta^*) + \bar{X}_F'(\beta^* - \beta_F)]$$

Machine Learning Models

Six models are trained to predict log monthly earnings from the same covariate set (with categorical variables one-hot encoded and numeric variables standardized): Linear Regression, Ridge Regression, Lasso Regression, Random Forest, Gradient Boosting, and XGBoost. Data are split 80/20 into training and test sets, stratified by gender. Models are evaluated on the held-out test set using root mean squared error (RMSE), mean absolute error (MAE), and R^2 , computed both overall and separately for male and female subsamples.

Interpretability: Permutation Importance and SHAP

For the best-performing model (XGBoost), we compute (i) grouped permutation importance — shuffling all one-hot-encoded columns belonging to a given categorical feature jointly, to avoid the well-known bias of permuting individual dummy columns in

isolation — and (ii) SHAP (SHapley Additive exPlanations) values via TreeExplainer, with feature-level importance again computed by summing |SHAP| across the one-hot columns belonging to each original feature.

Counterfactual Gender-Swap ('ML-Oaxaca')

For every female worker in the test set, we construct a counterfactual record identical in every covariate except that the gender indicator is recoded from female to male, and compare the trained model's predicted wage under the factual (female) and counterfactual (male) covariate vectors. The average counterfactual gap is:

$$Gap = E[\hat{y}(X_i, Male) | Female_i] - E[\hat{y}(X_i, Female) | Female_i]$$

This is repeated within subgroups (sector, employment type) to assess heterogeneity, and validated with a placebo test: the model is retrained on data in which the gender label has been randomly permuted across observations (breaking any genuine gender-wage relationship in training), and the same counterfactual exercise is re-run — a properly specified test should show the estimated gap collapse toward zero.

RESULTS

Results below draw on the two most recently available PLFS rounds (2021-22 and 2022-23), pooled; extending the panel to the full set of survey rounds since 2017-18 is a natural extension once the corresponding rounds' column layouts have been harmonized (Limitations Section).

Descriptive Statistics

Table 1: Descriptive Statistics of the Wage-analysis Sample, by Gender
(Pooled 2021-22 and 2022-23 PLFS)

Variable	Male	Female
N	159,050	44,528
Mean age (years)	40.34	39.44
Mean years of schooling	9.46	8.45
% rural	51.87	52.02
% regular wage/salaried	41.04	46.75
% self-employed	58.39	52.89
% casual labour	0.57	0.36

Source: *The Author*

Table 1 reports descriptive statistics for the wage-analysis sample by gender. Women in the sample are, on average, slightly younger and have about one fewer year of schooling than men. Women are somewhat more concentrated in regular wage/salaried employment (46.8% vs. 41.0%) and less concentrated in self-employment (52.9% vs. 58.4%) than men, with negligible participation in casual labour for both groups in this age- and status-restricted sample.

The Raw Gender Wage Gap

Table 2 reports the raw (unadjusted) gender wage gap in monthly earnings. The gap is 36.2% at the mean (37.7% using PLFS sample weights) and considerably larger at the median (53.6%), indicating that the gap is more pronounced in the lower and middle parts of the earnings distribution than the upper tail.

Table 2: Raw Gender Wage Gap in Monthly Earnings (Rs.)

Statistic	Male	Female	Gap (%)
Mean (unweighted)	18,012	11,487	36.23
Mean (weighted)	16,836	10,484	37.73
Median	14,000	6,500	53.57

Source: *The Author*

The gap varies substantially by employment type and sector (Table 3, Figure 1). It is largest among the self-employed (57.0%) and smallest among regular wage/salaried employees (21.6%), and larger in rural areas (48.1%) than urban areas (27.9%).

Table 3a: Raw Gender Wage Gap by Employment Type

Employment type	N (male)	N (female)	Mean, male (Rs.)	Mean, female (Rs.)	Gap (%)
Casual labour	908	160	7,256	3,925	45.90
Regular wage/salaried	65,279	20,818	21,670	16,984	21.63
Self-employed	92,863	23,550	15,546	6,679	57.04

Table 3b: Raw Gender Wage Gap by Sector

Sector	N (male)	N (female)	Mean, male (Rs.)	Mean, female (Rs.)	Gap (%)
Rural	82,492	23,164	14,178	7,365	48.05
Urban	76,558	21,364	22,143	15,955	27.94

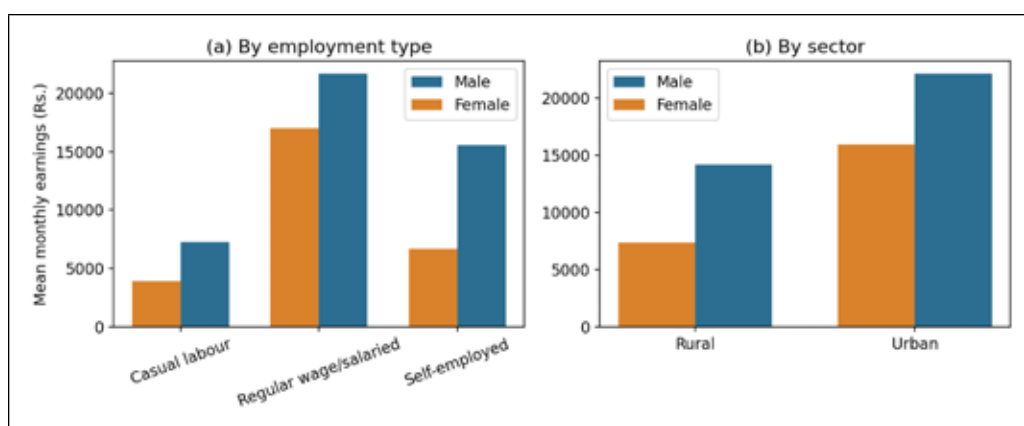


Fig. 1: Mean monthly earnings by gender: (a) employment type, (b) sector.

Mincer log-wage regression

Table 4 reports selected coefficients from the Mincer regression (full specification includes education, social group, religion, occupation, and industry fixed effects; $N = 203,578$, $R^2 = 0.479$). The gender coefficient of -0.617 implies that, holding education, experience, marital status, social group, religion, sector, employment type, occupation, and industry constant, women earn approximately 46.0% less than observably similar men.

Table 4: Mincer Log-wage Regression, Selected Coefficients
(HC1 Robust SE; $N = 203,578$; $R^2 = 0.479$)

Variable	Coefficient	Std. Error	p-value	Note
Gender (Female = 1)	-0.6167	0.004	< 0.001	Implied 46.03% conditional gap
Experience	0.0327	0.0005	< 0.001	Mincer-proxy potential experience
Experience ²	-0.0005	8.2e-6	< 0.001	Diminishing returns to experience
Sector (Urban = 1)	0.1719	0.003	< 0.001	Urban wage premium
Regular wage/salaried	0.6910	0.024	< 0.001	Relative to casual labour (base)
Self-employed	0.6441	0.024	< 0.001	Relative to casual labour (base)

Source: *The Author*

Notably, the conditional gap (46.0%) is larger than the raw gap (36.2%). This pattern — not uncommon in gender-gap studies with occupational and sectoral controls — arises because women in this sample are somewhat more concentrated in employment types and sectors that carry an average earnings premium (regular wage/salaried employment; see

Table 1), so raw comparisons partly net out a compositional advantage that is separately controlled for in the regression; once composition is held fixed, the within-group pay penalty attached to being female is larger than the raw comparison alone suggests.

Oaxaca-Blinder Decomposition

The Oaxaca-Blinder decomposition (Table 5) confirms this pattern from a different angle: only 7.4% of the total log-wage gap is attributable to differences in observable characteristics between men and women (the "explained" component), while 92.6% — equivalent to a 47.4% wage penalty — is attributable to differences in the returns to those characteristics (the "unexplained" component).

Table 5: Oaxaca-Blinder Decomposition of the Log-wage Gap (Pooled-Coefficient Reference)

Component	Log points	Share / equivalent
Total log-wage gap	0.694	50.04% (level equivalent)
Explained (endowments)	0.051	7.4% of total gap
Unexplained (returns)	0.643	92.6% of total gap ($\approx$ 47.4% wage penalty)

Source: *The Author*

Detailed decomposition of the (small) explained component shows that general education level is the single largest contributor to the explained gap (0.063 log points), followed by marital status (0.011) and social group (0.004); employment type, occupation, and industry composition actually work in the opposite direction (women's composition on these dimensions would, other things equal, be expected to narrow rather than widen the gap), consistent with the pattern noted in Section 5.3.

Machine Learning Model Comparison

Table 6: Model Comparison: Predictive Performance (Test Set)

Model	Train RMSE	Test RMSE	Test MAE	Test R ²	Test R ² (Male)	Test R ² (Female)
Linear Regression	0.580	0.582	0.448	0.471	0.381	0.423
Ridge	0.580	0.582	0.448	0.470	0.381	0.423
Lasso	0.585	0.587	0.452	0.461	0.370	0.412
Random Forest	0.544	0.558	0.427	0.513	0.401	0.522
Gradient Boosting	0.561	0.565	0.433	0.500	0.398	0.487
XGBoost	0.524	0.537	0.409	0.549	0.448	0.552

Source: *The Author*

Table 6 and Figure 2 compare the predictive performance of the six machine learning models. XGBoost achieves the best out-of-sample fit (test RMSE = 0.537, $R^2 = 0.549$), followed by Random Forest (RMSE = 0.558, $R^2 = 0.513$) and Gradient Boosting (RMSE = 0.565, $R^2 = 0.500$); the linear-family models (Linear Regression, Ridge, Lasso) all perform notably worse ($R^2 \approx 0.46$ - 0.47), indicating that non-linearities and interaction effects not captured by the additive Mincer-style specification meaningfully improve wage prediction.

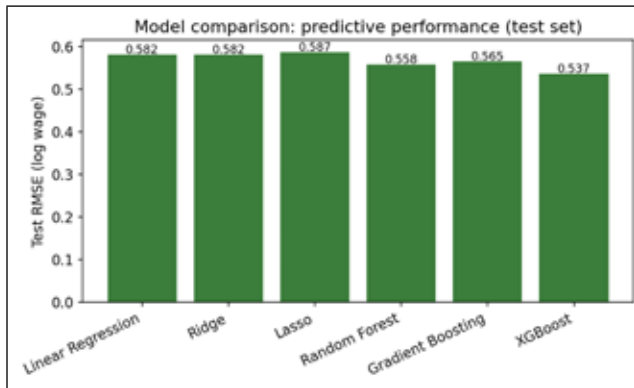


Fig. 2: Test-set RMSE Across the Six Machine Learning Models

A notable and policy-relevant finding is that every model fits male wages better than female wages (e.g., for XGBoost, test $R^2 = 0.448$ for men versus 0.552 for women, and RMSE = 0.506 versus 0.636 respectively). While the higher R^2 for women mechanically reflects greater relative dispersion in the female wage distribution around a lower mean, the persistently higher absolute error for women across every model indicates that female wages are systematically harder to predict from the same observable characteristics — consistent with a labour market in which non-observed factors (including potential discrimination) play a larger role in determining women's pay.

Feature Importance: Permutation Importance and SHAP

Table 7: Grouped Permutation Importance, XGBoost Model (Top 7 Features)

Rank	Feature	Δ RMSE
1	Gender	0.1347
2	Years of schooling	0.0919
3	Industry (2-digit NIC)	0.0737
4	Employment type	0.0417
5	Occupation (broad)	0.0336
6	Age	0.0292
7	Sector	0.0174

Source: *The Author*

Table 7 reports grouped permutation importance (increase in test RMSE when a feature is randomly shuffled) for the XGBoost model. Gender is the single most important predictor, ahead of education, industry, occupation, and all other covariates.

SHAP analysis (Table 8, Figure 3) corroborates this: summing |SHAP value| across the one-hot-encoded columns belonging to each original feature, gender ranks second overall (18.6% of total feature importance), narrowly behind industry (18.75%) and ahead of years of schooling (16.0%), occupation (9.5%), and ahead of years of schooling (16.0%), occupation (9.5%), and sector (8.0%).

Table 8: Grouped SHAP Feature Importance, XGBoost Model (Top 10 Features)

Rank	Feature	Share of total SHAP (%)
1	Industry (2-digit NIC)	18.75
2	Gender	18.60
3	Years of schooling	15.97
4	Occupation (broad)	9.45
5	Sector (rural/urban)	8.01
6	Age	6.70
7	Employment type	5.99
8	Social group	4.12
9	Survey round	3.69
10	Religion	3.00

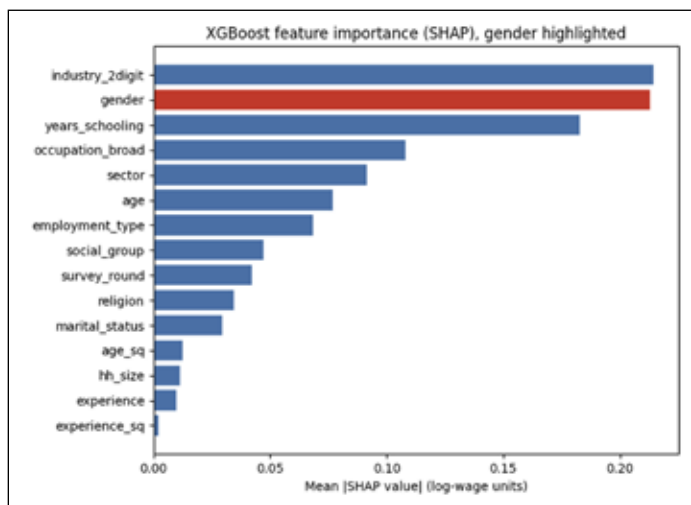


Fig. 3: SHAP Feature Importance (Grouped by Original Feature); Gender Highlighted in Red

Counterfactual Gender-Swap And Cross-Method Comparison

Using the trained XGBoost model, we predict each woman's wage under her actual (factual) covariates and under a counterfactual in which only her gender is recoded to male. Averaged across all 8,906 women in the test set, mean predicted earnings rise from Rs. 9,306 (factual) to Rs. 14,606 (counterfactual as male) — an average gender penalty of 36.3%, closely corroborating the raw gap (36.2%) and broadly consistent with, though somewhat smaller than, the Mincer and Oaxaca-Blinder conditional estimates (46.0% and 47.4% respectively).

A placebo test validates this methodology: retraining XGBoost on a version of the training data in which gender labels were randomly permuted (breaking any genuine relationship between gender and wages) and repeating the identical counterfactual exercise yields an estimated gap of -0.1% — indistinguishable from zero, as required if the method is correctly measuring a genuine learned relationship rather than a mechanical artifact of the swap procedure.

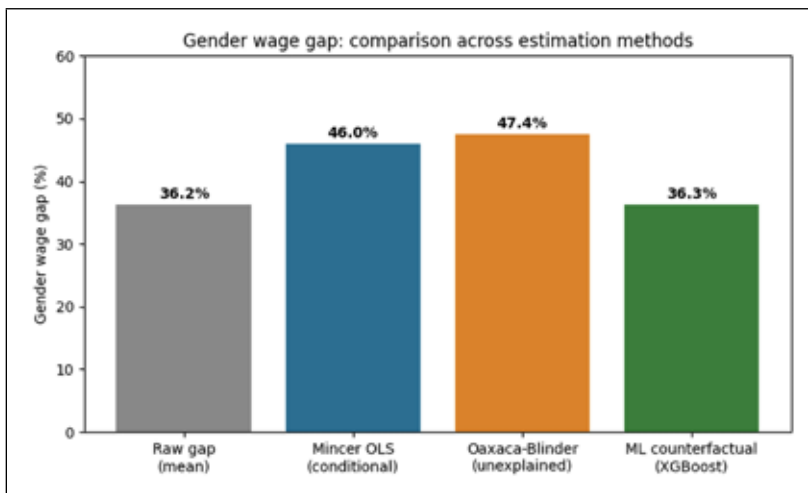


Fig. 4: Gender Wage Gap Estimates Across Four Independent Methods

Table 9: Comparison of Gender Wage Gap Estimates Across Methods

Method	Estimated gender gap (%)
Raw gap (mean, unweighted)	36.2
Mincer OLS (conditional gender coefficient)	46.0
Oaxaca-Blinder (unexplained component)	47.4
ML counterfactual (XGBoost gender-swap)	36.3

Source: *The Author*

The convergence of the raw comparison and the ML counterfactual (both $\approx 36\%$) with the higher conditional estimates from the Mincer and Oaxaca-Blinder methods (46-47%) is informative in itself: it suggests that once other characteristics are held fixed via either a parametric (Mincer) or non-parametric (ML) model, the estimated "all-else-equal" penalty attached to gender is larger than the raw comparison alone would suggest, because women are not, on average, negatively selected on other observable dimensions relative to men in this sample (Section 5.3).

Heterogeneity: Counterfactual Gap By Subgroup

Repeating the counterfactual exercise within subgroups (Table 10) shows substantial heterogeneity. The gender penalty is roughly twice as large among the self-employed (53.1%) as among regular wage/salaried employees (24.8%), and considerably larger in rural areas (47.6%) than urban areas (28.8%) — mirroring the pattern already evident in the raw gap (Section 5.2) and consistent with the hypothesis that informality and weaker enforcement of pay regulations in self-employment and rural labour markets are associated with larger gender pay penalties.

Table 10: Heterogeneous ML Counterfactual Gender Wage Gap, by Subgroup

Subgroup	N (female)	Counterfactual gap (%)
Rural	4,664	47.55
Urban	4,242	28.79
Regular wage/salaried	4,189	24.84
Self-employed	4,687	53.09

Source: *The Author*

DISCUSSION

Three independent methodological approaches — a parametric Mincer regression, a classical Oaxaca-Blinder decomposition, and a non-parametric machine-learning counterfactual — converge on the same substantive conclusion: the majority of India's gender wage gap cannot be explained by observable differences in education, experience, marital status, social group, religion, location, occupation, or industry between male and female workers. The Oaxaca-Blinder decomposition attributes only 7.4% of the total gap to such observable endowments; the remaining 92.6% reflects differences in the returns to those same characteristics — women earn systematically less than men who look identical on paper.

This is corroborated by the machine-learning results in two distinct ways. First, gender is not merely a marginally significant control variable but one of the two or three most predictively important features in the entire model — more important than education,

occupation, or sector individually, according to both permutation importance and SHAP. If wages were determined purely by human capital and job characteristics, gender should add little predictive power once those characteristics are known; the fact that it adds substantial power indicates either direct discrimination, unmeasured productivity-relevant differences correlated with gender (e.g., unobserved heterogeneity in job tenure, hours, or bargaining), or both. Second, the ML counterfactual exercise — which makes no linear-functional-form assumption and can capture interactions the Mincer equation cannot — nonetheless arrives at an economically similar estimate of the residual gap (36.3%), and a rigorous placebo test confirms this is not a modelling artifact.

The heterogeneity results reinforce a segmentation-theory interpretation: gaps are far larger in segments of the labour market characterized by weaker institutional oversight (self-employment, rural work) than in the most regulated segment (regular wage/salaried, urban employment), consistent with the view that formal labour market institutions constrain, without eliminating, gender-based pay disparities.

LIMITATIONS

- **Coverage:** this draft uses two of the nine PLFS rounds available in the underlying data (2021-22, 2022-23). Extending the panel to the full 2017-18–2022-23 period (and, separately, assessing comparability with the redesigned 2024-25 monthly panel) would allow the gap's trend over time to be examined and would substantially increase statistical power for subgroup (e.g., state-level) analysis.
- **Outcome variable:** monthly earnings, rather than hourly wages, are used as the outcome, pending verification of the PLFS field corresponding to weekly/usual hours of work against MoSPI's official schedule documentation. If men and women differ systematically in hours worked, part of the monthly-earnings gap reported here could reflect an hours-of-work gap rather than a pure hourly-wage gap; this should be re-estimated once the hours field is verified.
- **Formal/informal employment:** a small number of employment-classification fields (workplace location, contract duration, social security eligibility, enterprise size) that would allow a fully verified formal/informal employment indicator are similarly pending confirmation against the official layout; a coarser employment-type classification (self-employed / regular / casual) is used in its place.
- **Selection into employment:** the analysis conditions on being employed with positive earnings; if labour force participation itself is non-random with respect to gender and unobserved wage potential (a well-documented feature of the Indian labour market, where female labour force participation is comparatively low), the estimated gap may be subject to selection bias. A Heckman-style selection correction is a natural extension.

- **Causal interpretation:** neither the Oaxaca-Blinder "unexplained" component nor the ML counterfactual gap should be interpreted as a clean causal estimate of discrimination. Both are residual measures that will also capture any relevant, wage-determining characteristics not included in the model (e.g., detailed job tenure, firm size, bargaining power, or pre-market factors such as differential access to networks and mentorship) as well as measurement error in self-employment income.
- **Survey weights:** descriptive statistics use PLFS sample multipliers; the regression and machine-learning models presented here are unweighted. Weighted versions are a recommended robustness check.
- **Single held-out split:** reported ML performance and importance metrics derive from a single train/test split; confidence intervals via bootstrap resampling are recommended before treating specific point estimates (e.g., the exact 36.3% counterfactual gap) as final.

POLICY IMPLICATIONS

The finding that only a small share of the gender wage gap is explained by observable human-capital and job characteristics has a direct policy implication: interventions focused solely on closing gender gaps in education or occupational access, while valuable in their own right, are unlikely by themselves to substantially close the pay gap. The results point to three complementary policy priorities.

- Strengthen enforcement of equal-pay provisions under the Code on Wages, 2019, with particular attention to sectors and occupations where our results show the largest unexplained gaps (Section 5.8), rather than relying on economy-wide averages that mask this heterogeneity.
- Prioritize formalization of self-employment and rural work, where the estimated gender penalty (53% and 48% respectively) is roughly double that in regular, urban wage employment (25-28%), consistent with formal labour-market institutions constraining (though not eliminating) gender-based pay disparities.
- Investigate within-occupation, within-industry pay-setting practices directly (e.g., through targeted firm-level wage audits), since the SHAP and permutation-importance results show gender adding substantial independent predictive power over and above detailed occupation and industry codes — suggesting the gap operates at least partly within, not only between, these categories.

CONCLUSION

This paper combined econometric and machine-learning approaches to estimate and interpret the gender wage gap in India using PLFS microdata from 2021-22 and 2022-

23. All methods point to a large, economically significant gender wage gap (36-47% depending on method and specification) of which the great majority is not explained by differences in education, experience, location, occupation, industry, or employment type between men and women. Machine learning models, in particular XGBoost, both predict wages more accurately than the conventional Mincer specification and independently confirm, via SHAP-based interpretability and a validated counterfactual exercise, that gender remains one of the most important determinants of predicted wages even after conditioning on a rich set of observable characteristics. The gap is substantially larger among self-employed and rural workers than among regular, urban wage employees, pointing toward formalization and targeted enforcement, rather than human-capital investment alone, as priority areas for policy aimed at closing India's gender pay gap.

REFERENCES

- Anand, R., Mishra, S., & Peiris, S. J. (2013). Inclusive growth: Measurement and determinants. *IMF Working Paper No. 13/135*. International Monetary Fund
- Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets*. Princeton University Press.
- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725.
- Azam, M., & Prakash, N. (2015). A distributional analysis of public-private wage differential in India. *Labour*, 29(4), 394-414. <https://doi.org/10.1111/labr.12068>
- Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.
- Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Columbia University Press.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436-455.
- Bourguignon, F. (2004). The poverty-growth-inequality triangle. *Indian Council for Research on International Economic Relations (ICRIER) Working Paper No. 125*. New Delhi: ICRIER.
- Chamarbagwala, R. (2006). Economic liberalization and wage inequality in India. *World Development*, 34(12), 1997-2015. <https://doi.org/10.1016/j.worlddev.2006.02.010>
- Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- Mincer, J. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research.
- Mukherjee, D. (2007). Post-reform trends in wage differentials: A decomposition analysis for India. *Indian Journal of Labour Economics*, 50(4), 667-684.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.
- Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources*, 23(3), 279-295.

- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.
- Organisation for Economic Co-operation and Development (OECD). (2015). *In it together: Why less inequality benefits all*. OECD Publishing. <https://doi.org/10.1787/9789264235120-en>
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62(4), 659-661.
- Ministry of Statistics and Programme Implementation, Government of India. Periodic Labour Force Survey (PLFS), Annual Reports, various years.

APPENDIX: DATA PROVENANCE AND VARIABLE VERIFICATION STATUS

Core demographic and geographic variables (gender, age, marital status, general education level, sector, state, household type, religion, social group, usual principal activity status) were verified against value distributions consistent with published NSS/PLFS code lists. Industry (NIC) and occupation (NCO) codes, and the monthly earnings fields for wage/salary and self-employment income, were identified with high confidence from value-range and magnitude checks (median earnings for employed men, for example, are consistent with published PLFS aggregates), pending final cross-validation against MoSPI's official schedule layout. Enterprise size, precise weekly hours of work, and detailed formal/informal employment sub-classifications require further verification and were therefore not used in the reported models; confirming these fields is the priority for the next iteration of this analysis and would additionally support construction of a true hourly wage measure.