INTERNATIONALE HOCHSCHULE

www.iu.de

IU DISCUSSION PAPERS

Human Resources

Assessing Wage Inequality with Machine Learning: Approaches for Measuring the Adjusted Gender Pay Gap

OLIVER PLÜGHAN

KATHARINA-MARIA REHFELD

IU Internationale Hochschule

Main Campus: Erfurt

Juri-Gagarin-Ring 152

99084 Erfurt

Telefon: +49 421.166985.23

Fax: +49 2224.9605.115

Kontakt/Contact: kerstin.janson@iu.org

Autorenkontakt/Contact to the author(s):

Oliver Plüghan

ORCID-ID: 0009-0000-9271-9722

Email: oliver@plueghan.com

Prof. Dr. Katarina-Maria Rehfeld

ORCID-ID: 0000-0003-4366-1282

IU Internationale Hochschule - Fernstudium

Juri-Gagarin-Ring 152

99084 Erfurt

Email: katharina-maria.rehfeld@iu.org

IU Discussion Papers, Reihe: Human Resources, Vol. 6, No. 4 (MÄR 2026)

ISSN: 2750-0721

DOI: https://doi.org/10.56250/4118

Website: https://repository.iu.org

Assessing Wage Inequality with Machine Learning: Approaches for Measuring the Adjusted Gender Pay Gap

Oliver Plüghan

Katharina-Maria Rehfeld

ABSTRACT:

This paper investigates the methodological performance of Ordinary Least Squares (OLS) regression and Random Forest machine learning algorithms in measuring adjusted gender pay gaps. The research is motivated by the European Union's Pay Transparency Directive (2023/970), which mandates that employers report adjusted gender pay gaps. While Oaxaca-Blinder Decomposition and the underlying OLS regression have served as the industry standard for gap estimation, this paper examines whether machine learning approaches can better capture complex, nonlinear compensation relationships. Using synthetic datasets with controlled discrimination parameters, the study compares both methods across two sample sizes and multiple discrimination scenarios. Key findings demonstrate that both methods successfully distinguish between occupational segregation and direct wage discrimination at large sample sizes. However, at smaller sample sizes, Random Forest exhibits substantial instability whereas OLS remains slightly more stable. A methodological adjustment, training Random Forest on the larger population before applying predictions to subsets substantially improves small-sample performance. The paper concludes that OLS regression remains preferable for formal regulatory compliance due to its interpretability and stability, while Random Forest can serve as a complementary validation tool for large-scale analysis.

KEYWORDS:

Gender Pay Gap, Pay Transparency, OLS Regression, Random Forest, Wage Discrimination, Unexplained Wage Gap, Adjusted Gender Pay Gap

JEL CLASSIFICATION: J16, J31, J71, M52, C13, C45

AUTHORS

Oliver Plüghan *completed a Bachelor of Engineering in Engineering Management at IU International University and is employed in the Human Resources department of a global shipping and logistics company, working in the field of compensation and benefits. In his current position, he plays an active role in implementing the EU Pay Transparency Directive and supports the development and evaluation of compensation structures, with a strong interest in transparency, the gender pay gap, and the usage of Artificial Intelligence.*

Prof. Dr. Katharina-Maria Rehfeld *is a Professor of Human Resource Management at IU International University of Applied Sciences, specializing in International Human Resource Management, Digital HR, personnel and leadership development, and organizational behavior. Drawing on over twelve years of professional experience in Asia—including roles at the German Chamber of Commerce Abroad in Shanghai and Allianz China Life—she combines academic expertise with practical insights, which also inform her work as a trainer and coach for international executives.*

1. Gender Pay Gap Persistence and Measurement

Understanding the adjusted gender pay gap requires clarifying how wage inequality is measured. The gender pay gap refers to the difference between the average gross hourly earnings of male and female employees, expressed as a percentage of male earnings (Eurostat, 2025). In the European Union, the unadjusted gap stood at 12.0 percent in 2023, meaning women earned about 88 cents for every euro earned by men. While the gap declined from 16.4% in 2012 to 12.9% in 2020, progress has slowed in recent years (Eurostat, 2025). Germany shows a larger gap, standing at 18% in 2023 and declining slightly to 16% in 2025 (Destatis, 2025).

The adoption of Directive 2023/970 in May 2023 represents an important development in EU pay transparency policy. Entering into force on 7 June 2023 and requiring implementation by June 2026, the Directive introduces mandatory pay transparency obligations and gender pay gap reporting across the EU (Directive 2023/970, 2023). A key requirement is the measurement and reporting of the adjusted gender pay gap to identify wage discrimination.

The unadjusted gap measures average earnings differences across the workforce and reflects broader labor market patterns such as women's concentration in lower-paid sectors (Leythienne & Ronkowski, 2018, p. 6). The formula is:

$$\frac{Mean\ (hourly)\ wage\ of\ men\ -\ Mean\ (hourly)\ wage\ of\ women}{Mean\ (hourly)\ wage\ of\ men} \tag{1}$$

Measuring the adjusted gender pay gap requires more sophisticated methods. Econometric approaches, particularly Ordinary Least Squares (OLS) regression within the Oaxaca-Blinder framework, have long been the standard for adjusted pay gap estimation (Castagnetti et al., 2017, p. 3). OLS coefficients are easily interpretable, allow statistical testing, and provide confidence intervals (Bzovsky et al., 2022, pp. 1715-1716).

However, OLS assumes linear relationships between predictors and wages (Nokeri, 2022, p. 13), which may not reflect complex organizational pay structures —such as nonlinear relationships between performance ratings and compensation. In contrast, machine learning methods such as Random Forests offer an alternative by capturing nonlinear relationships and interactions (Zare, 2025, p. 2), as they construct multiple decision trees and aggregates their predictions (Biau & Scornet, 2015, p. 1).

Pay transparency can also reduce information asymmetries between employers and employees (Baggio & Marandola, 2023, p. 166). When employees become aware of pay disparities, they may negotiate higher wages or leave the organization (Baggio & Marandola, 2023, p. 181).

This paper therefore examines how OLS regression and Random Forest models perform in measuring the adjusted gender pay gap and how these approaches support HR communication, stakeholder trust, and decision-making.

This article combines a literature review with empirical analysis using synthetic datasets and computational modeling. Literature searches were conducted using the EBSCO academic database with search terms including "gender pay gap," "pay transparency," "ordinary least squares," and "Random Forest." Additionally, the AI-powered research tool Perplexity was used in research mode with prompts aimed at identifying literature on regression-based gender pay gap estimation.

AI Assistance: Perplexity AI supported literature identification, phrasing improvements, and grammar checks, while all analysis and content were produced by the authors.

2. Gender Pay Gap

Leythienne & Ronkowski (2018) describe the explained portion of the unadjusted gender pay gap as the part attributable to observable differences between men and women, such as occupational distribution or tenure, noting that this binary framing does not capture non-binary gender identities.

The unexplained portion represents the residual gap that remains after accounting for observable characteristics (Leythienne & Ronkowski, 2018, p. 15).

However, explained differences do not necessarily imply the absence of discrimination. Occupational segregation may reflect historical barriers, promotion bias, or discriminatory career guidance (Blau & Kahn, 2017, p. 832; Leythienne & Ronkowski, 2018, p. 6). Likewise, the unexplained component may capture discrimination but also omitted variables or measurement error (Leythienne & Ronkowski, 2018, p. 11).

This distinction has practical implications. If pay gaps mainly reflect occupational segregation, organizations may focus on recruitment and promotion policies to increase women's representation in higher-paid roles (International Labour Organization, 2018, pp. 18-19). A large unexplained gap, by contrast, indicates potential pay disparities within comparable roles.

This paper focuses on the unexplained component, referred to as the Adjusted Gender Pay Gap (APG), a terminology also used by the German Federal Statistical Office (Mischler, 2021, p. 5). The Unadjusted Pay Gap (UPG) includes both components.

2.1 REASONS FOR THE GENDER PAY GAP

Gender pay gaps arise from multiple factors at individual, organizational, and labor market levels and discrimination may be reflected in either the unadjusted or also in the adjusted gender pay gap, depending on the underlying reasons.

Negotiation and Behavioral Differences

Gender differences in negotiation behavior contribute to wage disparities. Experimental evidence indicate that male employees obtain starting salaries about 7.6% higher than female employees (Dittrich et al., 2014, p. 862). The authors attribute this to differences in negotiation approaches rather than bargaining ability (Dittrich et al., 2014, p. 872). More recent research suggests that women are more likely to negotiate pay but less likely to succeed, which Kray and Lee (2024) attribute to direct discrimination against women.

Occupational Segregation

Occupational segregation refers to the unequal distribution of men and women across industries and occupations. Women are more often concentrated in lower-paid sectors such as education and social services, while men dominate higher-paid fields like STEM and finance (Bishu & Alkadry, 2016, p. 74). The OECD distinguishes between horizontal segregation (different occupations) and vertical segregation (different hierarchical positions) (OECD, 2025, p. 4). Women are underrepresented in management positions (Blau & Kahn, 2017, p. 826).

Research also suggests that psychological traits may play a role. Bertrand (2011) finds that women are on average more risk-averse than men, which may affect earnings in high-reward environments (p. 1547).

Because the adjusted gender pay gap compares employees in similar roles, occupational segregation primarily affects the unadjusted gap.

Hiring Discrimination

A meta-analysis of more than 70 audit experiments shows that discrimination often penalizes applicants entering gender-atypical occupations (Galos & Coppock, 2023). Women face disadvantages in male-dominated high-wage occupations, while men experience similar penalties in female-dominated fields, reinforcing occupational segregation.

Flexibility & Career Interruptions

Another explanation relates to job structures that reward long or inflexible working hours. Goldin (2014) argues that the gender pay gap would decline substantially if firms did not disproportionately reward employees who can work long hours (p. 1092).

Research also shows that the gender pay gap often emerges after the birth of the first child. Using administrative data, Kleven et al. (2019) demonstrate that earnings trajectories diverge sharply after childbirth due to reduced working hours and career interruptions among women (p. 10).

These mechanisms produce substantial economic consequences. In Germany, women earn approximately €670,000 less than men over their working lives (Bönke et al., 2020, p. 29), and lifetime wage expectation gaps may reach about €500,000 (Kiessling et al., 2024, p. 4).

At the macroeconomic level, gender equality could significantly increase economic output. The McKinsey Global Institute estimates that gender parity in economic participation could increase global GDP by USD 28 trillion (Das et al., 2019, p. 7).

2.2 EU PAY TRANSPARENCY DIRECTIVE

The EU Pay Transparency Directive combines gender pay gap reporting with enforcement mechanisms and joint pay assessments.

The Directive identifies transparency as a key instrument for achieving pay equity (Directive 2023/970, 2023, Recital 11). Employers must conduct joint pay assessments when gender pay gaps exceed 5% and cannot be justified by objective criteria (Directive 2023/970, 2023, Art. 10).

Article 9 requires employers with at least 100 employees to report gender pay gap information by categories of workers performing the same work or work of equal value (Directive 2023/970, 2023). This structure requires adjusted pay gap calculations to determine whether observed differences are justified.

The reporting requirement addresses historical information asymmetries between employers and employees (Ceballos et al., 2022, p. 4). Article 7 further strengthens transparency by granting employees access to pay information within their worker category (Directive 2023/970, 2023).

If a pay gap of 5% or more cannot be justified by objective criteria, employers must conduct a joint pay assessment and implement corrective measures (Directive 2023/970, 2023, Art. 10). Failure to comply may result in fines (Directive 2023/970, 2023, Art. 23).

The Directive also shifts the evidentiary burden toward employers, who must demonstrate that pay differences within comparable roles are based on objective criteria (Directive 2023/970, 2023, Art. 18).

Evidence from existing policies suggests transparency can reduce wage gaps. For example, pay transparency in the United Kingdom increased women's wages by about 1.6 percentage points relative to men's wages (Blundell, 2021).

3. Measuring the Adjusted Gender Pay Gap

Adjusted gender pay gaps are typically estimated using regression analysis, most commonly log-linear models that relate wages to variables such as education, experience, and job characteristics (Leythienne & Ronkowski, 2018, p. 15).

3.1 OAXACA-BLINDER DECOMPOSITION

The Oaxaca-Blinder decomposition is a common method for analyzing wage differences (Hübler, 2003, p. 557). Separate regression models are estimated for male and female employees, and differences in characteristics and coefficients are decomposed to identify explained and unexplained components (Leythienne & Ronkowski, 2018, pp. 8-9).

The explained component reflects differences in observable characteristics, while the unexplained component remains after controlling for these factors (Blinder, 1973; Oaxaca, 1973).

However, results depend on the choice of reference coefficients and rely on linear regression assumptions (Jann, 2008, p. 461).

3.2 OLS-REGRESSION

Ordinary Least Squares (OLS) regression estimates the adjusted gender pay gap directly as the coefficient on a gender variable while controlling for relevant characteristics (Redmond, 2019, p. 570).

The standard model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \tag{2}$$

The gender coefficient can be interpreted as a percentage wage difference. For example, $\beta_{female} = -0.08$ indicates that women earn approximately 7.7% less than comparable men (Redmond, 2019, pp. 570-573).

OLS offers transparency and statistical inference through significance tests and goodness-of-fit measures (Cameron, 2022, p. 224). However, it assumes linear relationships between predictors and wages (Cameron, 2022, p. 218). In practice, wage growth varies across career stages (Deming, 2023), and gender penalties may differ across organizational levels (Ciminelli et al., 2021).

Despite these limitations, this paper uses pooled OLS regression as the primary method because it provides clear and interpretable estimates suitable for organizational communication.

3.3 MACHINE LEARNING: RANDOM FORESTS

Random Forests represent a machine learning method that constructs multiple decision trees and aggregates their predictions (Biau & Scornet, 2015, p. 1). Unlike OLS, Random Forests do not impose a predefined functional form and can capture nonlinear relationships and interactions (Louppe, 2014, p. 26).

Decision trees partition the data based on employee characteristics, leading to predicted salary values (Sarker, 2021; IBM, 2023).

Random Forest analysis requires selecting a reference gender. The model is trained on the reference group and predicts counterfactual salaries for all employees (Flachaire & Picard, 2025, p. 2). The adjusted gender pay gap can then be calculated by comparing predicted and actual salaries.

However, Random Forests suffer from the “black box” problem because their predictions are less interpretable than regression coefficients (Louppe, 2014, p. 2). This may complicate communication in organizational contexts.

Therefore, this paper uses Random Forests as a complementary validation method. OLS and Oaxaca-Blinder provide transparent and established estimates, while Random Forests help identify nonlinear patterns that may not be captured by regression models.

4. Methodology

This section outlines the methodological approach used to compare OLS regression and Random Forest performance in measuring adjusted gender pay gaps. It explains the construction of synthetic datasets with known discrimination parameters, the OLS regression specification, and the Random Forest implementation using a counterfactual approach.

4.1 ESTABLISH A DATABASE

To compare OLS regression and Random Forest approaches, a synthetic dataset was constructed with controlled characteristics, known parameters, and systematically varied sources of wage inequality. This design allows both methods to be tested against known wage-setting mechanisms and gender wage penalties while making assumptions and limitations transparent.

The foundational datasets consist of 1,000 and 100 fictional employee records, each with a unique ID and the following attributes (created with Microsoft Excel):

- Gender: randomly distributed at around 50 percent male and 50 percent female using randbetween(0,1), with 0 coded as male and 1 as female
- Age: randomly generated between 22 and 62 years
- Department: eight occupational departments reflecting common organizational structures
- Education: four levels, from High School to PhD
- Years of Experience: derived from age using (Age – 22) as a proxy for labor market tenure
- Job Grade/Title: five hierarchical levels from Specialist to Executive
- Performance Rating: scale from 1 to 5 representing annual performance outcomes

Salaries were generated using a controlled compensation function combining linear and nonlinear components to reflect realistic organizational pay practices. Salary depends linearly on department, education, and job grade/title, and nonlinearly on years of experience and performance, reflecting stronger early-career salary growth and exponential rewards for exceptional performance (Deming, 2023).

Each attribute also includes random error to capture unexplained variation in compensation due to unmeasured factors, ensuring that salary cannot be perfectly predicted from observable characteristics, as in real company settings.

Baseline Dataset (dataset_0)

The baseline dataset includes only the compensation mechanisms described above and no deliberate gender-based pay differentiation. Gender has no systematic effect on salary; any observed pay gap results only from random error and possible differences in gender distribution across departments and job grades. By design, both the unadjusted and adjusted gender pay gap should be close to zero, allowing validation that the methods correctly detect the absence of discrimination.

Biased Representation Dataset (dataset_0_biased)

To test whether the methods can distinguish occupational segregation from wage discrimination, the baseline dataset was modified to create structural gender imbalance across job grades. Female representation was increased in lower grades (Specialist and Manager) and reduced in higher grades (Senior Manager, Director, Executive). All other employee characteristics and salaries remained unchanged; only gender assignments were modified.

This dataset is intended to produce an unadjusted gender pay gap clearly different from zero, due to women's concentration in lower-paid roles, while keeping the adjusted gender pay gap near zero. It therefore tests whether OLS regression and Random Forests correctly attribute the gap to occupational segregation rather than direct wage discrimination.

Biased Representation Dataset incl. Adjusted Gender Pay Gap of x percent (dataset_x-x_percent)

To evaluate methodological performance under direct wage discrimination, the baseline dataset was modified so that female employees received deliberate salary reductions relative to comparable male employees. This introduces a known discrimination mechanism resembling real-world within-role wage penalties.

In the first discrimination dataset (dataset_0-3_percent), each female employee received a random salary reduction between 0% and 3%. This is expected to generate an adjusted gender pay gap of roughly 1.5%, the midpoint of the range.

Additional datasets were created with expected adjusted gender pay gaps between -1% and 15%, using random penalty ranges of 0-3%, 3-5%, 6-10%, 12-15%, 0-15%, and -1-5%. These scenarios allow an assessment of whether methodological performance changes across different discrimination levels.

The same employee attributes were also used to create a smaller dataset in order to test method reliability with limited sample sizes.

4.2 MODEL FOR OLS

The OLS regression model estimates the relationship between employee compensation and measured job and demographic characteristics, with particular emphasis on isolating the gender effect on salary. The model follows the formula introduced in the previous section (Wooldridge, 2019, p. 47):

$$ln(w) = \beta_0 + \beta_{female} x_{female} + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \quad (3)$$

Where:

- ln(w) = natural logarithm of employee salary
- β_0 = constant
- x_{female} = binary indicator variable (1 for female, 0 for male)
- β_{female} = regression parameter on gender, representing the natural logarithm of the adjusted gender pay gap
- β_i = regression parameters for the control variables
- x_i = control variables (e.g. experience, performance)
- ϵ = random error term

Data Preparation and Variable Construction

The analysis begins by loading the data into Python using the pandas library (McKinney, 2010). Salary is transformed into its natural logarithm to create the dependent variable ln_salary required for the log-linear model.

Gender is converted into a binary indicator variable, and categorical variables are transformed into dummy variables using one-hot encoding.

Model Estimation: Pooled OLS Regression

Two OLS models are estimated from the pooled dataset:

1. Pooled model with gender
2. Objective pooled model without gender, enabling gender-neutral salary prediction

$$\ln(w) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \quad (4)$$

The regression coefficient on the female variable represents the log-difference in salary between female and male employees. To convert this into a percentage gap, the following transformation is used:

$$APG = e^{\beta_{female}} - 1 \quad (5)$$

Statistical significance is assessed using the p-value associated with the gender coefficient. If the p-value exceeds 0.05, the adjusted gender pay gap is treated as not statistically different from zero (Cameron, 2022, p. 224). If the p-value is below 0.05, the gender-based salary difference is considered statistically significant.

Additionally, the adjusted R^2 value is calculated to show the proportion of log-salary variance explained by objective characteristics, excluding gender. Higher R^2 values suggest that the model captures more of the true wage-setting relationship, whereas lower values may indicate nonlinear patterns or omitted variables (Wooldridge, 2019, pp. 199-200).

Calculation of Unadjusted Gender Pay Gap

For comparison, the unadjusted gender pay gap is calculated directly from the raw data. This measure captures the difference between average female and average male salaries without controlling for job-related characteristics (Leythienne & Ronkowski, 2018, p. 6):

$$Unadjusted\ Gender\ Pay\ Gap\ (\%) = \frac{\overline{w_m} - \overline{w_f}}{\overline{w_m}} \tag{6}$$

where $\overline{w}_g$ denotes the mean wage for gender group g.

The unadjusted gap is calculated using non-log salary values.

To support interpretation and validation, the OLS model produces two outputs: a predictions file with individual-level predicted salaries and a coefficient table listing regression coefficients and p-values. The coefficient table helps identify which factors most strongly influence compensation.

4.2 MODEL FOR RF

The Random Forest implementation begins with the same preprocessing steps used for OLS to ensure comparability. Salary is log-transformed, gender is recoded as a binary indicator, and categorical variables are converted into dummy variables. This produces an identical feature set across both methods.

The Random Forest model is built with deliberately selected hyperparameters to balance predictive flexibility and comparability with linear regression. A total of 1,000 trees (n_estimators = 1000) is used to reduce prediction variance through averaging and improve stability, especially relative to smaller default settings (Breiman, 2001; Probst & Boulesteix, 2017). Maximum tree depth is restricted to 8 (max_depth = 8) to limit model complexity, memory use, and processing time (Pedregosa et al., 2011).

The model is trained on the male subset of the data, following the counterfactual framework (Leythienne & Ronkowski, 2018, p. 11). It learns the wage-setting function from male employees' characteristics and wages, then applies this function to all employees to generate gender-neutral salary predictions.

Unlike OLS, the Random Forest explicitly excludes the gender variable. This is necessary for the counterfactual logic: the model should learn compensation patterns only from objective characteristics such as department, job title, education, experience, and performance.

Adjusted Gender Pay Gap Calculation via Counterfactual Predictions

The adjusted gender pay gap is then calculated using a counterfactual approach similar to the Oaxaca-Blinder decomposition, though operationalized differently. First, the trained Random Forest generates salary predictions for all employees.

These predictions represent what female employees would earn if they were compensated according to the gender-neutral wage function learned from male employees, given their actual characteristics.

The difference between actual and predicted salaries is then compared across gender groups. The difference between the gender means of these gaps represents the adjusted gender pay gap in log form; exponentiation converts it into percentage terms. This approach is similar to Takács & Vincze (2019), though it does not use classification trees (CART) but directly compares gender prediction results.

If the Random Forest generates identical predictions for equivalent workers regardless of gender, the adjusted gap approaches zero. If female employees systematically earn less than predicted, or male employees more than predicted, the resulting gap captures unexplained pay differences after accounting for observed characteristics.

An additional practical benefit is that, once salaries are predicted, this logic can also be applied in a simple Excel spreadsheet to subsets of data using predictions from the full male-trained model.

The two methods are then evaluated by comparing their estimates with the actual constructed pay gap.

A detailed employee-level file is also exported, including both OLS and Random Forest salary predictions. This enables validation at the individual level and helps assess whether the models produce plausible results.

To address small-sample constraints in Random Forest estimation, a methodological adaptation is used for the n=100 sample. Instead of training the model on the small male subset, the Random Forest is trained on the full population equivalent (n=1,000) and then applied to the n=100 subset. This assumes that the underlying compensation mechanism remains stable across organizational subsets.

5 Results

.

This section presents the results of Ordinary Least Squares (OLS) regression and Random Forest (RF) models across three scenarios: no discrimination, occupational segregation, and direct wage discrimination. The analysis uses synthetic datasets with known pay gaps, allowing a direct comparison of how accurately each method identifies gender-based pay differences. Two sample sizes (1,000 and 100 employees) are considered to assess performance under different data conditions.

5.1 RESULTS INTRODUCTION

The baseline dataset represents a scenario without intentional discrimination, where salaries are determined only by job-related factors. Any observed gender pay gap is therefore due to random variation.

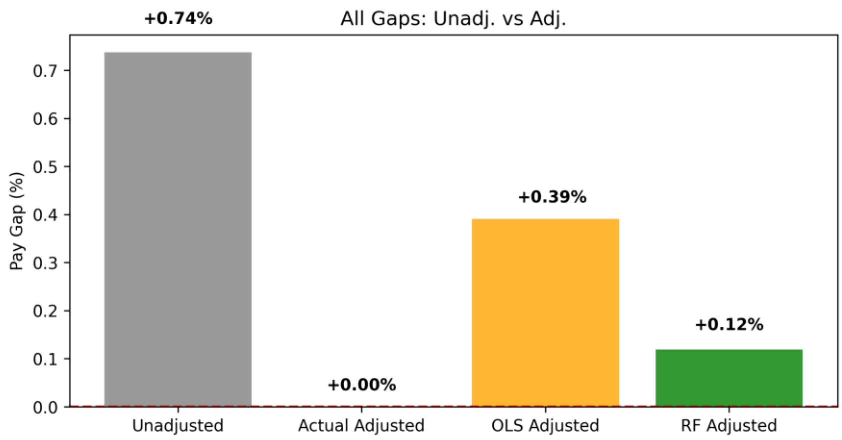
Both OLS and RF correctly identify the absence of discrimination. Differences between large and small samples are minor and reflect normal statistical variation. This confirms that both methods do not falsely detect discrimination when none exists.

5.2 BASELINE DATASET (NO INTENTIONAL DISCRIMINATION)

The baseline dataset (dataset_0) represents a controlled scenario with no deliberate gender-based wage discrimination. Compensation is determined only by legitimate job-related factors, departmental assignment, education, job grade, years of experience, and performance, plus random variation. Any observed gender pay gap therefore reflects random fluctuation rather than systematic discrimination.

This baseline serves two purposes: it tests whether both methods correctly identify the absence of discrimination and provides a reference point for evaluating discrimination scenarios.

Figure 1: Unadjusted and Adjusted Pay Gaps / dataset_0 / n=1000



Comparing n=1,000 and n=100 shows that, when no discrimination exists, both methods correctly estimate a pay gap close to zero percent. Differences across sample sizes are small and mainly reflect greater random variation in the smaller sample.

5.3 OCCUPATIONAL SEGREGATION WITHOUT WAGE DISCRIMINATION

To distinguish occupational segregation from direct wage discrimination, the baseline dataset was modified to create gender imbalance across hierarchical levels. Female representation was increased in lower-tier job grades and reduced in higher-tier positions.

All salaries, performance ratings, and employee characteristics remained unchanged; only gender assignments were altered. The resulting unadjusted gender pay gap therefore reflects occupational segregation rather than unequal pay within comparable roles.

This scenario tests whether both methods can separate structural gender concentration in lower-paid roles from direct wage discrimination.

Large Sample (n=1,000)

The unadjusted gender pay gap reaches **-10.37 percent**, far above the baseline value of 0.74 percent, due to the concentration of women in lower-paid positions.

The OLS-derived adjusted gender pay gap is **0.51 percent** (**p = 0.0795**), indicating that after controlling for hierarchy and other characteristics, female compensation is slightly higher than male compensation. However, the p-value exceeds the 0.05 threshold, so the difference is not statistically significant.

The RF estimate is **0.17 percent**, slightly smaller than the OLS estimate. Both adjusted estimates remain close to zero.

Small Sample (n=100)

In the small sample, the unadjusted gender pay gap is **-20.14 percent**.

OLS yields an adjusted gender pay gap of **-1.32 percent** (**p = 0.2287**), which is not statistically significant. The larger deviation relative to the n=1,000 case reflects higher sampling variability.

RF yields **-2.89 percent**, producing a larger divergence from OLS than in the large sample. This suggests that RF's flexible functional form becomes less stable when training data are limited.

5.4 WAGE DISCRIMINATION

Intended 1.5% APG

Female salaries are reduced by random amounts between 0 and 3 percent, creating an intended adjusted gender pay gap of about -1.50 percent.

At n=1,000, the unadjusted gap is -11.68 percent. OLS estimates an adjusted gap of -0.98 percent (p = 0.0008), underestimating the true value by 0.52 percentage points. RF estimates -1.32 percent, underestimating by 0.18 percentage points and performing more accurately.

At n=100, the unadjusted gap is -21.44 percent. OLS estimates -2.86 percent (p = 0.0118), while RF estimates -4.53 percent. Both overestimate the true gap, especially RF.

Intended 4% APG

Female salaries are reduced by 3 to 5 percent, creating an intended adjusted gap of about -4.00 percent.

At n=1,000, the unadjusted gap is -13.94 percent. OLS estimates -3.50 percent (p < 0.0001) and RF estimates -3.82 percent. Both are close to the true value.

At n=100, the unadjusted gap is -23.35 percent. OLS estimates -5.26 percent, while RF estimates -6.84 percent. Again, both overestimate, with RF deviating more strongly.

Intended 8% APG

Female salaries are reduced by 6 to 10 percent, creating an intended adjusted gap of about -8.00 percent.

At n=1,000, the unadjusted gap is -17.51 percent. OLS estimates -7.49 percent (p < 0.0001) and RF -7.80 percent. Both perform accurately.

At n=100, the unadjusted gap is -26.78 percent. OLS estimates -9.40 percent, while RF estimates -10.88 percent. Both overestimate, with RF again less stable.

Intended 13.5% APG

Female salaries are reduced by 12 to 15 percent, creating the largest discrimination scenario with an intended adjusted gap of about -13.50 percent.

At n=1,000, the unadjusted gap is -22.55 percent. OLS estimates -13.16 percent ($p < 0.0001$) and RF -13.42 percent. Both closely match the true value, with RF slightly closer.

At n=100, the unadjusted gap is -30.88 percent. OLS estimates -14.66 percent and RF -15.95 percent. Both overestimate, especially RF.

Intended 7.5 APG

This scenario applies random female salary penalties across the full 0 to 15 percent range, producing an average intended adjusted gap of about -7.50 percent.

At n=1,000, the unadjusted gap is -16.92 percent. OLS estimates -7.00 percent ($p < 0.0001$) and RF -7.36 percent. Both are close to the true value.

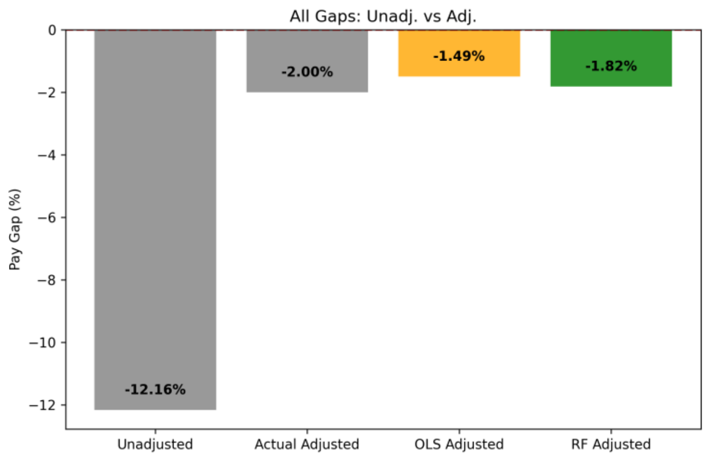
At n=100, the unadjusted gap is -25.17 percent. OLS estimates -8.30 percent, while RF estimates -9.22 percent. RF again shows greater overestimation.

Negative Discrimination Scenario

In this scenario, female salary adjustments range from -1 to +5 percent, so some women receive salary premiums and others penalties. The intended adjusted gap is about -2.00 percent.

At n=1,000, the unadjusted gap is -12.16 percent. OLS estimates -1.49 percent ($p < 0.0001$) and RF -1.82 percent.

Figure 2: Unadjusted and Adjusted Pay Gaps / dataset_-1-5_percent / n=1000



At n=100, the unadjusted gap is -21.96 percent. OLS estimates -3.72 percent ($p = 0.0011$) and RF -5.03 percent. RF again overestimates more strongly.

Small sample with large model

Small samples appear challenging for RF. In the original analysis, training RF only on the male subset of the n=100 sample led to unstable estimates because too few observations were available to learn the wage-setting function.

To address this, RF was trained on the male employees of the full population (**n=1,000**) and then applied to the **n=100** subset. This reflects the practical idea that employers may need to report adjusted gaps by worker category rather than for the full organization (EU Directive 2023/970, 2023, Art. 9), while assuming that the underlying compensation mechanism remains broadly similar across subsets.

This full-population training approach improved numerical stability. Table 1 compares OLS results with the initial RF estimates and the revised RF estimates based on full-population training.

Table 1: Adjusted Gender Pay Gaps for each model / n=100

n=100	OLS APG	initial RF APG	RF APG	Actual APG
dataset_0_s	-0,95%	-1,32%	-1,32%	0%
dataset_0_biased_s	-1,32%	-2,89%	-0,78%	0%
dataset_0-3_percent_s	-2,86%	-4,53%	-2,46%	-1,50%
dataset_3-5_percent_s	-5,26%	-6,84%	-4,82%	-4%
dataset_6-10_percent_s	-9,40%	-10,88%	-8,95%	-8%
dataset_12-15_percent_s	-14,66%	-15,95%	-14,13%	-13,50%
dataset_0-15_percent_s	-8,30%	-9,22%	-8,49%	-7,50%
dataset_-1-5_percent_s	-3,72%	-5,03%	-2,05%	-2%

5.5 SUMMARY

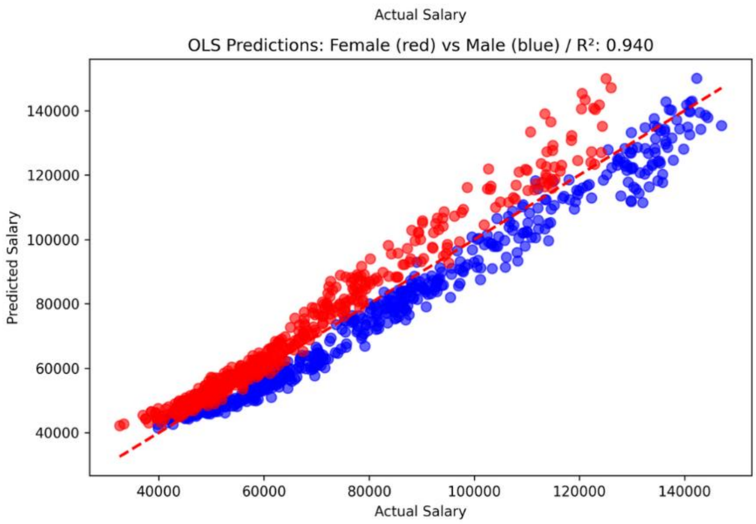
Across all datasets without wage discrimination, baseline and occupational segregation scenarios, both OLS therefore do not identify discrimination where none was introduced.

Across all datasets with deliberate wage discrimination, both methods produce p-values below 0.05 and reject the null hypothesis of a zero adjusted gender pay gap. This indicates sufficient statistical power to detect discrimination across the tested range.

A key result is the effect of sample size. At **n=1,000**, the mean absolute divergence between OLS and RF is only **0.32 percentage points**, and both methods recover the true pay gap with similar accuracy. At **n=100**, divergence rises to **1.27 percentage points**, showing that methodological choice becomes much more important in small samples.

Although both methods perform similarly in large samples, OLS shows systematic inaccuracies at the individual level because it linearizes relationships such as experience and performance.

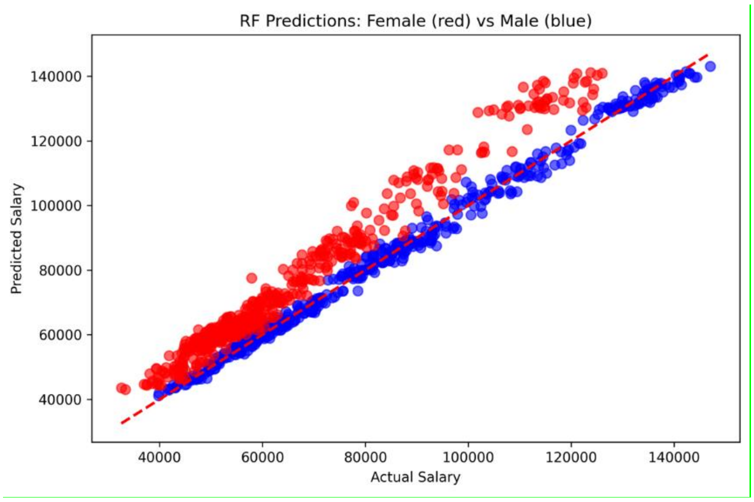
Figure 3: Predicted vs. Actual salaries from OLS model



As shown in Figure 3, OLS predictions display a clear curvilinear pattern rather than a random spread around the line. This indicates that the linear model cannot fully capture the nonlinear compensation structure in the data.

By contrast, RF aligns more closely with actual salaries and largely eliminates this curvature.

Figure 4: Predicted vs. Actual salaries from RF model



However, the RF regression line is centered around male salaries because the model is trained only on the male subset, whereas OLS uses the full sample and therefore centers more between male and female observations.

The average RF divergence of about **2.5 percentage points** across small-sample scenarios indicates systematic instability under limited data. Under these typical organizational conditions, OLS appears more stable.

The alternative RF approach, training on the full population and applying the model to subsets, substantially reduces this instability and improves consistency. Still, in practice smaller companies may not have access to a sufficiently large training sample. Under ordinary sample constraints, OLS therefore remains preferable.

Beyond predictive performance, OLS also offers an important practical advantage: interpretability. Unlike RF, which operates as a “black box,” OLS provides explicit regression coefficients showing how each variable contributes to salary determination (Louppe, 2014).

For individual-level wage prediction, the model excluding gender was used so that predicted salaries reflect only legitimate, non-discriminatory pay factors.

6 Discussion

Section 6 interprets the findings and compares them with the literature. It examines what the results imply for OLS and RF performance, and for their usefulness in organizational decision-making. In simple terms, this section explains what the results mean in practice and how organizations can use them.

6.1 RESULT DISCUSSION

The findings provide a methodological basis for estimating adjusted gender pay gaps by comparing OLS and RF using controlled datasets with known discrimination parameters. Both methods identify adjusted pay gaps and distinguish occupational segregation from direct wage discrimination, but their performance differs by sample size.

OLS produces regression coefficients that directly quantify the contribution of each variable to salary. The female coefficient can be translated into an intuitive percentage wage gap. For example, a coefficient of **-0.038** implies that women earn about **3.7%** less than comparable men. This transparency makes OLS easier for stakeholders and HR professionals to interpret and communicate. This makes OLS particularly useful for non-technical audiences who need clear and explainable results.

As a pooled regression, OLS also avoids the reference-group ambiguity that complicates counterfactual RF approaches. However, Figure 9 reveals a major limitation: OLS predictions show pronounced curvature, especially at lower salary levels. This reflects the nonlinear dependence of salary on experience and performance in the data. Because OLS imposes a linear functional form, it cannot fully capture such patterns (Nokeri, 2022, pp. 13-14). This is consistent with Dustmann & Meghir (2005), who found nonlinear relationships between experience and wages in German labor market data (p. 100).

RF shows a different pattern. Predicted salaries align much more closely with actual salaries across the distribution, suggesting that its tree-based structure captures nonlinearities that OLS misses (Louppe, 2014, p. 26).

At the same time, the adjusted gap estimates themselves are very similar at **n=1,000**. This suggests that although OLS predicts individual salaries less precisely, both methods recover the average gender wage penalty similarly when sample size is sufficient.

At n=1,000, OLS and RF produce very similar results, while at n=100, differences become much larger and RF is less accurate.

Han et al. (2021) report similar small-sample issues in health-related research and propose techniques such as class balancing, variable screening, and hyperparameter tuning. These options were not examined further here, but the pattern supports the finding that RF performs worse in small datasets when no mitigation is applied.

The article also introduces a solution that improves RF in small-sample settings: training the RF model on the full population and then applying it to smaller subsets. Under this approach, RF estimates for the n=100 subset become more accurate. This approach reflects a practical idea: using as much available data as possible to improve predictions in smaller groups.

In practice, this means that for category-specific analysis, for example within the Finance department, the predictive model could first be trained on all available employee data and then applied to the relevant subgroup. This uses the full information available while preserving the flexibility of machine learning.

A further advantage of RF is that, once trained, it can be used in simple tools such as Microsoft Excel. This may allow HR professionals to calculate category-specific adjusted pay gaps without specialized statistical software.

For formal compliance reporting and regulatory submission, however, OLS appears more suitable as the primary method. It is stable, interpretable, well established, and performs more reliably across sample sizes. Its coefficients quantify the gender wage penalty directly, while significance testing indicates whether observed gaps are distinguishable from zero.

For larger organizations and more detailed category-specific monitoring, RF trained on the full population may be useful as an additional check. In this sense, both methods can complement each other.

Overall, OLS is more interpretable and robust, while RF offers greater flexibility but less transparency.

6.2 REMEDIATION

The identification of gender pay gaps through OLS and RF is only the diagnostic step. The ultimate goal is remediation that reduces or removes identified disparities.

Because both methods generate individual-level predicted salaries representing what employees would earn absent discrimination, organizations can compare predicted and actual salaries to identify individual gaps for both male and female employees. This helps organizations move from analysis to concrete action.

This makes it possible to detect strong deviations from expected pay and prioritize targeted remedial action.

Under the EU Pay Transparency Directive, organizations must also ensure that the gender pay gap within each worker category does not exceed 5% (Directive 2023/970, 2023, Art. 10). After salary adjustments, however, the models must be rerun iteratively, since changes in salaries alter the coefficients or, in the RF case, the model outputs. Remediation therefore requires repeated updating.

Effective long-term remediation also depends on identifying the mechanisms behind the observed gaps. Different causes require different interventions.

Remediation should not focus on pay alone but also address structural issues such as occupational segregation. Organizations should also address occupational segregation, especially vertical segregation, in which women are overrepresented in lower-paid levels and functions (OECD, 2025, p. 4).

7 Conclusion

This section summarizes the key findings, discusses limitations, and identifies directions for future research.

7.1 SUMMARY

This paper examined how well OLS regression and Random Forest models measure the adjusted gender pay gap and how their outputs can be used by HR professionals and organizational leadership, including for remediation.

The findings show that both OLS and RF successfully distinguish between occupational segregation and direct wage discrimination. In all non-discrimination scenarios, both methods correctly fail to identify significant adjusted pay gaps. In all deliberate discrimination scenarios, ranging from 1.5 percent to 13.5 percent, both methods detect wage gaps.

A critical finding concerns sample size. At **n=1,000**, OLS and RF estimates converge closely, indicating that both methods recover the true discrimination parameter similarly despite their different assumptions. At **n=100**, RF becomes more unstable. Training RF on the full population and then applying it to subsets substantially improves this small-sample performance.

Regarding interpretability and organizational use, OLS provides explicit coefficients for major pay determinants such as experience, education, job grade, and performance. These coefficients help HR professionals diagnose wage drivers, plan remediation, and communicate results transparently. RF, by contrast, offers better predictive fit in large samples under nonlinearity but less interpretability. Still, once trained, RF can be used in offline spreadsheet-based monitoring, which may support ongoing compliance with Article 9–10 of the Directive (Directive 2023/970, 2023, Art. 9-10).

7.2 LIMITATIONS AND CRITICAL CONSIDERATIONS

This paper uses 16 synthetic datasets with predefined wage-setting functions. This allows validation against known discrimination parameters but simplifies real organizational complexity. Future research should test these findings using actual company data collected under EU Pay Transparency Directive compliance.

Real-world discrimination may also operate through more indirect mechanisms, such as gendered performance evaluation or biased valuation of experience, rather than the explicit salary penalties modeled here.

The article further assumes additive compensation effects, meaning each variable contributes independently to salary. Real pay systems likely contain important interactions, for example between experience and job role or between performance and seniority. Because these interactions were not built into the datasets, the extent to which RF’s interaction-capturing ability provides a real advantage cannot be fully assessed.

The analysis also assumes complete and reliable employee data, including salary, education, experience, and performance. In practice, organizations often face fragmented, incomplete, or subjective data, especially when measuring relevant experience or performance consistently.

Although synthetic data show that RF can capture nonlinearity missed by OLS, real organizational data do not reveal the true discrimination magnitude. As a result, methodological choice in practice must rest on theoretical and practical considerations rather than definitive empirical proof of superior accuracy.

Finally, OLS was used only for coefficient estimation, although it also produces predicted salary values that could be used to calculate adjusted pay gaps manually, similar to the RF spreadsheet approach. This prediction-based OLS approach was not explored and could be examined in future research.

This paper expands the methodological options in pay equity analysis by bringing Random Forest into direct comparison with traditional OLS regression for gender pay gap measurement in organizational contexts.

References:

Baggio, M. & Marandola, G. (2023). Employees' reaction to gender pay transparency: an online experiment. *Economic Policy, 38*, 161-188. https://doi.org/10.1093/epolic/eiac066

Bertrand, M. (2011). *New Perspectives on Gender. Amsterdam, Boston, Heidelberg:North-Holland, an Imprint of Elsevier*, 1543-1590.

Biau, G. & Scornet, E. (2015). *A Random Forest Guided Tour.* https://arxiv.org/abs/1511.05741

Bishu, S. & Alkadry, M. (2016). *A Systematic Review of the Gender Pay Gap and Factors That Predict It.* Administration & Society. https://doi.org/10.1177/0095399716636928

Blau, F. D. & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *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. *The Journal of Human Resources, 8*(4), 436-455. https://doi.org/10.2307/144855

Blundell, J. (2021). *Wage responses to gender pay gap reporting requirements*. Centre for Economic Performance, LSE.

Bönke, T., Glaubitz, R., Göbler, K., Harnack, A., Pape, A. & Wetter, M. (2020). *Die Entwicklung und Prognose von Lebenserwerbseinkommen in Deutschland.* https://doi.org/10.17169/refubium-26748

Bonaccolto-Töpfer, M. & Briel, S. (2022). The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics; Volume 78* (2022) 102223. https://doi.org/10.1016/j.labeco.2022.102223

Breiman, L. (2001). Random Forests: Random forests. *Machine Learning, 45*, 5-32. https://doi.org/10.1023/a:1010933404324

Bzovsky S., Phillips M., Guymer R., Wykoff C., Thabane L. & Bhandari M. (2022). The clinician's guide to interpreting a regression analysis. *Eye (Lond). 2022;36*(9):1715-1717. https://doi.org/10.1038/s41433-022-01949-z

Cameron, A. C. (2022). Analysis of Economics Data: An Introduction to Econometrics.

Castagnetti, C., Rosti, L., & Töpfer, M. (2017). The convergence of the gender pay gap: An alternative estimation approach. Stuttgart: Universität Hohenheim, Fakultät Wirtschafts- und Sozialwissenschaften.

Ceballos, M., Masselot, A., & Watt, R. (2022). Pay Transparency across Countries and Legal Systems. CESifo Forum, 23(2), 3-11.

Ciminelli, G., C. Schwellnus & B. Stadler (2021). Sticky floors or glass ceilings? The role of human capital, working time flexibility and discrimination in the gender wage gap. OECD Economics Department Working Papers, No. 1668, OECD Publishing, Paris, https://doi.org/10.1787/02ef3235-en

Das, S., Kotikula, A. & Carranza, E. (2019). Gender-Based Employment Segregation; Understanding Causes and Policy Interventions. World Bank, Washington, DC. https://doi.org/10.1596/31510

Deming, D. (2023). *Why Do Wages Grow Faster for Educated Workers?* SSRN Electronic Journal; ISSN 1556-5068. https://doi.org/10.2139/ssrn.4483673

Destatis. (2025). Gender pay gap.

https://www.destatis.de/DE/Themen/Arbeit/Verdienste/Verdienste-GenderPayGap/_inhalt.html

Directive 2023/970. (2023). Directive (EU) 2023/970 of the European Parliament and of the Council

of 10 May 2023 to strengthen the application of the principle of equal pay for equal work or work of equal value between men and women through pay transparency and enforcement mechanisms. https://eur-lex.europa.eu/eli/dir/2023/970/oj

Dittrich, M., Knabe, A. & Leipold, K. (2014). Gender Differences in Experimental Wage Negotiations. *Economic Inquiry, 52*(2), 862-873. https://doi.org/10.1111/ecin.12060

Dustmann, C. & Meghir, C. (2005). Wages, Experience and Seniority. *The Review of Economic Studies, 72*(1), 77-108.

Eurostat. (2025). EU gender pay gap 2023. https://www.statista.com/statistics/1203158/gender-pay-gap-in-europe/

Flachaire, E. & Picard, B. (2025*). Decomposing Inequalities using Machine Learning and Overcoming Common Support Issues.*

Galos, D. & Coppock, A. (2023). *Gender composition predicts gender bias: A meta-reanalysis of hiring discrimination audit experiments.* Science Advances, 9. https://doi.org/10.1126/sciadv.ade7979

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *The American Economic Review, 104*(4), 1091-1119. https://doi.org/10.1257/aer.104.4.1091

Hübler, Olaf. (2003). *Geschlechtsspezifische Lohnunterschiede*. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung. 36. 539-559.

IBM. (2023). *What is random forest?* IBM Think Newsletter. https://www.ibm.com/think/topics/random-forest

International Labour Organization. (2018). *Global wage report 2018/19: What lies behind gender pay gaps.* ILO Publications. https://www.ilo.org/publications/global-wage-report-201819-what-lies-behind-gender-pay-gaps

Jann, B. (2008). The Blinder-Oaxaca Decomposition for Linear Regression Models. The Stata Journal: Promoting Communications on Statistics and Stata, 8, 453-479. https://doi.org/10.1177/1536867x0800800401

Kiessling, L., Pinger, P., Seegers, P. & Bergerhoff, J. (2024). *Gender differences in wage expectations and negotiation.* Labour Economics, 87. https://doi.org/10.1016/j.labeco.2024.102505

Kleven, H., Landais, C. & Søgaard, J. E. (2019). Children and Gender Inequality:Evidence from Denmark. *American Economic Journal: Applied Economics, Vol. 11 , No. 4* , Pp. 181-209.

Kray, L. J., Kennedy, J. A., & Lee, M. (2024). Now, women do ask: A call to update beliefs about the gender pay gap. *Academy of Management Discoveries, 10(*1), 11-37. https://doi.org/10.5465/amd.2022.0021

Leythienne, D. & Ronkowski. (2018). *A decomposition of the unadjusted gender pay gap using Structure of Earnings Survey data.* Publications Office of the European Union.

Louppe, G. (2014). *Understanding Random Forests: From Theory to Practice*. https://arxiv.org/abs/1407.7502

McKinney, W. (2010). *Pandas: Python data analysis library*. https://pandas.pydata.org/

Mischler, F. (2021). *The gender pay gap: A cause analysis based on the structure of earnings survey 2018*. Federal Statistical Office of Germany (Destatis).

Nokeri, T.C. (2022). *Introduction to Econometrics. In: Econometrics and Data Science*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-7434-7_1

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*(3), 693-709. https://doi.org/10.2307/2525981

OECD. (2025). *Gender gaps in paid and unpaid work persist*. OECD Publishing, Paris. https://doi.org/10.1787/25a6c5dc-en

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, E. (2011). *Scikit-learn: Machine learning in Python*. Journal of Machine Learning Research, 12, 2825-2830. https://pypi.org/project/scikit-learn/

Probst, P. & Boulesteix, A. (2017). *To tune or not to tune the number of trees in random forest?* Journal of Machine Learning Research. https://doi.org/10.48550/arxiv.1705.05654

Redmond, P. & Mcguinness, S. (2019). The Gender Wage Gap in Europe: Job Preferences, Gender Convergence and Distributional Effects. *Oxford Bulletin of Economics & Statistics, 81*(3), 564-587. https://doi.org/10.1111/obes.12282

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science, 2*(3). https://doi.org/10.1007/s42979-021-00592-x

Han, S., Williamson, B. & Fong, Y. (2021). Improving random forest predictions in small datasets from two-phase sampling designs. *BMC Medical Informatics and Decision Making, 21*(1), 1-9. https://doi.org/10.1186/s12911-021-01688-3

Takács, O. & Vincze, J. (2019). *The gender pay gap in Hungary: New results with a new methodology.* Budapest: Hungarian Academy of Sciences, Institute of Economics.

Wooldridge, J. M. (2019). *Introductory Econometrics: A Modern Approach. (4th edition)*. Cengage Learning.

Zare, M. (2025). *Forecasting market returns using machine learning: evidence from Random Forest models.* Applied Economics Letters, 1-5. https://doi.org/10.1080/13504851.2025.2567614