ИКОНОМИЧЕСКИ ИЗСЛЕДВАНИЯ
ECONOMIC STUDIES

Volume 35(4), 2026

Aida Yzeiri Baftiari[1]
Artina Kamberi[2]
Agon Memeti[3]

ARTIFICIAL INTELLIGENCE AND THE GENDER PAY GAP IN LABOUR MARKETS: CROSS-COUNTRY EVIDENCE AND THE ROLE OF GENDER-SENSITIVE REFORMS[4]

This paper investigates whether the diffusion of artificial intelligence (AI) is associated with gendered labour market outcomes at the macro level and whether gender-sensitive institutional reforms mitigate gender inequality. We constructed a cross-country dataset covering up to 50 countries over the period 2012–2022. Given the rapid acceleration of AI adoption after 2022, the estimates should be interpreted as evidence for the 2012–2022 period and may not reflect the most recent wave of AI diffusion. We then estimate three complementary econometric models. First, using 500 promotion-related observations pooled across countries, we estimate a logistic regression to test whether AI adoption and an AI bias proxy are associated with women's promotion probabilities (RQ1; H1a–H1b). Second, using a balanced panel of 10 countries observed over 10 years (N = 100), we estimate country fixed-effects models for female labour force participation and the gender pay gap (RQ2; H2a–H2c). Third, we employ a difference-in-differences specification on the same panel to assess the impact of gender-sensitive institutional reforms on the gender pay gap (RQ3; H3). Across the logit and fixed-effects models, coefficients on AI adoption and the AI bias indicator are small, statistically insignificant, and imprecisely estimated, providing no support for the hypotheses that AI diffusion widens gender gaps in promotions, participation, or wages. By contrast, the reform indicator in the difference-in-differences model is negative, sizable, and highly significant, indicating that treated countries experience an average reduction of about 6.4 percentage points in the gender pay gap relative to non-treated countries. The findings suggest that AI adoption, as currently measured, is not a detectable structural driver of gender inequality in labour markets at the country level, whereas gender-sensitive institutional reforms are empirically associated with improved wage outcomes for women.

Keywords: Artificial intelligence; Gender inequality; Labour markets; Gender pay gap; Female labour force participation; Institutional reforms; Panel data; Difference-in-differences
JEL: J16; J21; J31; O33; C23; C25

---

[1] *Aida Yzeiri Baftiari, Faculty of Economics, University of Tetova.*
[2] *Artina Kamberi, Faculty of Economics, University of Tetova.*
[3] *Agon Memeti, Faculty of Natural Sciences and Mathematics, University of Tetova.*
[4] This paper should be cited as: *Baftiari, A. Y., Kamberi, A., Memeti, A. (2026). Artificial Intelligence and the Gender Pay Gap in Labour Markets: Cross-Country Evidence and the Role of Gender-Sensitive Reforms. – Economic Studies (Ikonomicheski Izsledvania), 35(4), pp. 90-113.*

1. Introduction

The diffusion of artificial intelligence (AI) into recruitment, evaluation, and wider employment-related decision processes has raised concerns that algorithmic tools may reinforce existing gender inequalities in labour markets. A growing empirical literature documents instances in which AI systems replicate stereotypes present in training data or institutional practices, with particular attention to hiring algorithms and targeted advertising. These findings have contributed to a prominent narrative: as AI adoption increases, gender disparities in employment, wages, and career advancement are expected to widen.

From the perspective of labour economics, this narrative represents an empirical hypothesis that can be tested. It implies that AI adoption is not merely a firm-level or technical phenomenon, but a macroeconomically relevant driver of gendered labour market outcomes. If this claim holds, we should observe systematic associations between AI diffusion and indicators such as female labour force participation, the gender pay gap, and women's promotion rates at the country level.

Despite the rapid growth of research on algorithmic fairness and technological change, macro-level evidence remains scarce. Existing studies are predominantly micro-level, focusing on particular firms, industries, or platforms. The extent to which AI-related biases aggregate to affect national-level gender gaps is essentially unknown. In addition, the role of public policy and institutional reforms – such as gender-sensitive regulations, non-discrimination law, or labour market governance – has received limited quantitative attention in the context of AI.

This paper contributes to filling this gap by providing a cross-country econometric assessment of the relationship between AI diffusion, gendered labour market outcomes, and gender-sensitive institutional reforms. We constructed a cross-country dataset covering up to 50 countries over the period 2012–2022. Due to data availability and consistency constraints, the econometric analysis proceeds in two parts: a promotion-based logistic regression using 500 pooled observations, and a balanced panel of 10 countries observed over 10 years (N = 100) for the fixed-effects and difference-in-differences models. We combine information from the World Bank Gender Data Portal, the ILO, AI adoption indicators from industry reports, and policy databases capturing gender-relevant institutional changes. We estimate three types of models:

1. A logistic regression of promotion-related outcomes on AI adoption and an AI bias indicator;
2. Fixed-effects panel regressions for female labour force participation and the gender pay gap;
3. A difference-in-differences (DiD) model assessing the effect of gender-sensitive institutional reforms on the gender pay gap.

The empirical findings can be summarised in three points. First, across both the logistic and fixed-effects specifications, AI adoption and the AI bias indicator do not exhibit statistically significant effects on promotion outcomes, female labour force participation, or the gender pay gap. Coefficients are generally small relative to their standard errors, and confidence intervals comfortably include zero. Second, macroeconomic controls such as education are also largely insignificant, while GDP growth shows some association with female participation, indicating the importance of broader structural dynamics. Third, the DiD estimates reveal that gender-sensitive institutional reforms are associated with statistically significant relative reductions in wage gaps, even after accounting for macroeconomic controls and country and time effects.

Taken together, these results suggest that, for the period and country sample under study, AI adoption is not a detectable structural driver of gender inequality at the macro level, while institutional reforms aimed at gender equity appear to matter. Micro-level algorithmic biases documented in case studies may be real and normatively problematic, but they do not automatically translate into measurable aggregate effects in cross-country data.

The contribution of the paper is twofold. First, it offers, to our knowledge, the first systematic cross-country econometric test of the hypothesis that AI adoption is associated with changes in gendered labour market outcomes. The finding of essentially null effects is substantively important: it challenges a widely held assumption and highlights the limits of extrapolating from micro-level algorithmic behaviour to macro-level outcomes. Second, the paper demonstrates empirically that gender-sensitive institutional reforms – rather than technology per se – are associated with reductions in gender wage gaps, underscoring the central role of institutions in mediating the distributional consequences of technological change.

The remainder of the article is structured as follows. Section 1.1 formulates the research questions and hypotheses that guide the empirical analysis. Section 2 reviews the literature on AI, gender inequality, and labour market dynamics, and develops the conceptual framework underpinning the hypotheses. Section 3 describes the data, variables, and econometric strategies. Section 4 presents the empirical results, while Section 5 discusses their implications, limitations, and robustness. Section 6 concludes with policy recommendations and directions for future research.

1.1. Research Questions and Hypotheses

To meet contemporary standards in empirical economics and to make the analysis fully explicit and testable, we structure the study around three research questions (RQs). Each question is linked to a set of hypotheses (H) and a specific econometric strategy.

RQ1. Is the diffusion of AI-based tools associated with gender-differentiated promotion outcomes at the country level?

Firm-level studies suggest that algorithmic hiring and evaluation tools may disadvantage women, especially when trained on historically biased data. If such mechanisms are widespread and sufficiently strong, they may aggregate into detectable macro-level differences in promotion-related outcomes.

Accordingly, we formulate:

- H1a (AI adoption and promotions). Higher AI adoption rates are associated with lower promotion probabilities for women, ceteris paribus.

- H1b (AI bias and promotions). Higher values of an AI bias indicator are associated with lower promotion probabilities for women, ceteris paribus.

These hypotheses are evaluated using a logistic regression model of promotion outcomes, where the statistical question is whether the coefficients on AI adoption and the AI bias indicator are significantly different from zero and display the expected negative sign.

RQ2. Is AI adoption associated with changes in female labour force participation and gender wage gaps across countries?

If algorithmic bias affects access to jobs, career progression, or pay, higher AI adoption might reduce women's labour force participation and widen the wage gap at the aggregate level.

We therefore posit:

- H2a (AI and participation). Higher AI adoption rates are associated with lower female labour force participation.
- H2b (AI and wage gaps). Higher AI adoption rates are associated with larger gender wage gaps.
- H2c (AI bias and wage gaps). Higher values of the AI bias indicator are associated with larger gender wage gaps.

These hypotheses are examined using country fixed-effects panel regressions, in which the coefficients on AI-related variables capture within-country changes in female participation and wage gaps over time, after controlling for macroeconomic factors and unobserved time-invariant country characteristics.

RQ3. Do gender-sensitive institutional reforms mitigate gender wage gaps, and if so, to what extent?

Governance-oriented literatures on algorithmic fairness and equality policy argue that regulation, transparency requirements, and institutional reforms can constrain discriminatory practices, including those mediated by technology. From a political economy perspective, such reforms may be expected to reduce gender wage gaps, irrespective of the precise level of AI adoption.

We therefore state:

- H3 (reform effect). Countries that adopt gender-sensitive institutional reforms experience a relative reduction in the gender pay gap compared with countries that do not, after controlling for macroeconomic conditions and fixed effects.

H3 is evaluated via a difference-in-differences framework, where the interaction between a treatment indicator (reform adoption) and a post-reform time dummy captures the average treatment effect on the treated.

2. Literature Review and Theoretical Framework

This section situates the paper within three related strands of literature: (i) studies on algorithmic bias and gender in AI systems; (ii) economic research on gendered labour market outcomes and technological change; and (iii) work on governance, institutional reforms, and algorithmic accountability. On this basis, it develops a conceptual framework linking AI adoption, gender inequality, and institutional responses, which underpins the research questions and hypotheses formulated in Section 1.1.

A first strand of research – largely in computer science and information studies – documents how AI systems can reproduce and amplify existing social biases, including gender stereotypes. Bolukbasi et al. (2016) show that word embeddings, a common component of many AI models, encode stereotypical associations such as "man: computer programmer: woman: homemaker". When such embeddings are used in downstream applications (e.g., recommendation systems, search ranking, or natural language processing in hiring tools), they can transmit and reinforce gendered expectations into automated decision-making.

Case-based evidence further highlights how algorithmic systems can disadvantage women in recruitment and evaluation. Dastin (2018) reports on an internal Amazon hiring tool that penalised résumés containing female-coded terms, effectively disfavouring women in technical roles. Lambrecht and Tucker (2019) find that targeted advertising for high-paying STEM jobs is more frequently shown to men than to women, even if targeting criteria are nominally gender-neutral. These examples suggest that AI systems, when trained on historically biased data or deployed in unregulated environments, may systematically disadvantage women.

At the same time, scholars emphasise that bias is not only a property of data, but also of the development context. West et al. (2019) argue that AI design processes reflect existing power structures, given the underrepresentation of women and minorities in technical teams and leadership positions. Raji et al. (2020) and Binns (2018) discuss how the lack of internal accountability mechanisms and weak documentation practices can contribute to the persistence of bias in deployed systems. These insights motivate the expectation that AI, as currently deployed, may embody gendered logics rooted in broader social and institutional structures.

However, most of this literature is micro-level and context-specific, focusing on individual algorithms, firms, or platforms. It demonstrates the existence of algorithmic gender bias, but does not address whether such micro-level phenomena translate into macro-level shifts in gendered labour market outcomes – precisely the question this paper examines.

A second strand of literature comes from labour and development economics, examining the determinants of gender gaps in employment, wages, and career progression. Classic work in labour economics emphasises human capital, occupational segregation, and labour market discrimination as key drivers of wage differentials and participation gaps. Gender norms, household responsibilities, access to childcare, and institutional barriers further shape women's labour supply and occupational choices.

Technological change has long been recognised as an important determinant of labour demand and wage structures. Theories of skill-biased technical change highlight how new technologies can raise the relative demand for skilled labour, potentially affecting gender gaps if skills are distributed asymmetrically between men and women. More recent work on automation and AI considers the substitution of routine tasks and the complementarity between technology and certain types of skills. Yet, much of this literature is gender-neutral in formal models, leaving the gendered implications implicit.

When gender is explicitly considered, the sign of the net impact of technology on women's outcomes is ambiguous. On the one hand, automation of routine tasks in clerical roles – where women are often overrepresented – could depress female employment and wages. On the other hand, technology may create new opportunities in sectors where women have comparative advantages or improve flexibility in work arrangements, potentially supporting higher participation.

In the context of AI, the key open question is whether higher AI adoption at the country level is associated with systematic changes in women's participation and relative wages. If algorithmic tools are widely used in recruitment, promotion, and evaluation, and if they systematically disadvantage women, then AI adoption could be expected to reduce female labour force participation and widen gender wage gaps. This expectation is reflected in hypotheses H2a–H2c. But such a macro-level effect remains an empirical question: it could also be the case that structural institutions and social norms dominate, and that the marginal contribution of AI is negligible in aggregate data.

A third strand of work explores how governance and institutional arrangements can mitigate or exacerbate algorithmic bias. Mehrabi et al. (2021) provide a survey of bias and fairness in machine learning, highlighting the breadth of technical and socio-technical interventions proposed to address algorithmic discrimination. Raji et al. (2020) call for end-to-end frameworks for algorithmic auditing and internal accountability, emphasising organisational processes rather than purely technical fixes.

From a more policy-oriented perspective, Williams et al. (2021) and Cowgill et al. (2020) examine the role of regulation, transparency, and "data justice" frameworks in shaping the impacts of AI systems. Proposals include mandatory disclosure of training data and model documentation, fairness audits prior to deployment, and ex post impact assessments. These scholars argue that institutions are central: similar AI systems can have very different distributional effects depending on the regulatory environment, enforcement capacity, and broader equality regimes in which they are embedded.

In economic terms, gender-sensitive institutional reforms – such as anti-discrimination laws, transparency requirements in digital employment tools, or stronger equality bodies – may act as constraints on discriminatory practices, both human and algorithmic. If such reforms are effective, one would expect them to be reflected in improved labour market outcomes for women, including narrower wage gaps. This logic underpins H3, which posits that countries implementing gender-sensitive reforms experience relative reductions in gender wage gaps compared to those that do not.

Bringing these strands together, the conceptual framework of this paper can be summarized as follows.

- Micro-level algorithmic bias – Evidence from computer science and case studies clearly shows that AI systems can embed gender bias. In firm-level settings, biased algorithms can affect hiring, promotions, and task allocation. This underlies the intuitive expectation behind RQ1 and RQ2: if AI-based systems are widely adopted, their bias may aggregate, leading to measurable differences in promotion outcomes, labour force participation, and wages at the country level. This expectation is formalised in H1a–H1b and H2a–H2c.
- Aggregation and structural dominance – However, when moving from firms to countries, micro-level mechanisms must pass through a variety of aggregation and mediation processes. Gendered labour market outcomes at the macro level are influenced by:
  - Education and human capital accumulation;
  - Occupational and sectoral segregation;
  - Social norms and household structures;
  - Labour market institutions, including minimum wages, collective bargaining, and anti-discrimination enforcement.

It is therefore theoretically plausible that structural and institutional factors dominate over AI-specific mechanisms in determining aggregate gender gaps. In such a scenario, even if bias exists at the algorithmic level, its marginal contribution to macro-level gender inequality could be too small – or too heterogeneous – to be detected in cross-country panel data.

- Role of gender-sensitive institutional reforms – Governance-oriented work suggests that institutions can either allow or constrain algorithmic and non-algorithmic discrimination. Gender-sensitive reforms – such as equality legislation tailored to digital contexts, AI-specific fairness requirements, or strengthened oversight bodies – are expected to reduce gender wage gaps and potentially interact with AI adoption. In our framework, these reforms are modelled as exogenous shifts in the institutional environment that can improve outcomes for women, irrespective of the direct impact of AI. This is captured in RQ3 and H3, and tested using a difference-in-differences design.

On this basis, the empirical analysis proceeds in three steps:

- Step 1 (RQ1 / H1a–H1b): Test whether AI adoption and an AI bias indicator are associated with promotion outcomes, using logistic regression.
- Step 2 (RQ2 / H2a–H2c): Test whether AI adoption and AI bias indicators are associated with female labour force participation and gender wage gaps, using fixed-effects panel regressions.
- Step 3 (RQ3 / H3): Test whether gender-sensitive institutional reforms are associated with relative reductions in gender wage gaps, using a difference-in-differences specification.

3. Data and Methodology

This section describes the data sources, the construction of the analytical samples, the variables used in the empirical analysis, and the econometric strategies employed to test the hypotheses formulated in Section 1. The aim is to make the empirical design transparent and replicable in line with contemporary standards in empirical economics.

3.1. Data Sources and Sample Construction

The empirical analysis combines information from four main sources:

- World Bank Gender Data Portal – This source provides harmonised, gender-disaggregated labour market indicators, including female labour force participation rates and, where available, data on gender differences in wages and employment structures.
- International Labour Organisation (ILOSTAT) – ILOSTAT is used to complement and cross-check the labour market indicators obtained from the World Bank, particularly for wage measures and employment by sector and gender.
- AI Adoption Indicators (Industry Reports) – Country-level indicators of AI adoption are drawn from industry and technology adoption reports. These indicators approximate the share of firms (or sectors) reporting the use of AI-based tools in recruitment, performance evaluation, or decision support.
- Policy and Institutional Databases – Information on gender-sensitive institutional reforms (e.g. equality legislation, AI-related fairness provisions, or broader digital non-discrimination frameworks) is collected from publicly available policy databases and official government or international organisation reports.

Given the heterogeneity in data availability across these sources, two distinct but related samples are constructed:

- Sample 1 (Promotion Outcomes, Logit Model) – This sample consists of 500 observations used in the logistic regression. Each observation corresponds to a recruitment or promotion outcome for which we observe: (i) an indicator of whether a female candidate was selected or promoted; (ii) a measure of AI adoption in the relevant context; and (iii) the macroeconomic controls. The data are pooled over the period 2012–2022, conditional on the joint availability of the variables.
- Sample 2 (Panel Data for FE and DiD Models) – The fixed-effects and difference-in-differences models are estimated on a panel of 10 countries observed over 10 years, yielding 100 country–year observations. This subsample is restricted to countries and years for which consistent information is available on female labour force participation, the gender pay gap, AI adoption indicators, macroeconomic controls, and, for the DiD specification, the timing of gender-sensitive reforms.

The use of a restricted panel for the FE and DiD analyses reflects the trade-off between cross-sectional coverage and longitudinal consistency. The relatively small number of countries

(10) and time periods (10) has implications for statistical power and inference, which are discussed in Section 5.

3.2. Variable Definitions

This subsection defines the variables employed in the empirical analysis, grouped into dependent variables, key explanatory variables, and control variables.

3.2.1 Dependent Variables

- Promotion Outcome (PROMO_it) – For the logistic regression (Sample 1), the dependent variable is a binary indicator equal to 1 if a woman is selected or promoted in a given recruitment or promotion process, and 0 otherwise. The unit of observation reflects aggregated outcomes aligned with the availability of AI adoption and control variables.
- Female Labour Force Participation (FLFP_it) – For the fixed-effects model, the dependent variable FLFP_it denotes the female labour force participation rate in country i at time t, measured as the percentage of women (aged 15 and above) who are either employed or actively seeking work.
- gender pay gap (GWG_it) – The second fixed-effects model and the DiD specification use the gender pay gap as the dependent variable. GWG is defined as the percentage difference between average male and female wages, expressed as a share of male wages. Positive values indicate that men earn more, on average, than women.

3.2.2 Key Explanatory Variables

- AI Adoption Rate (AI_Adoption_Rate_it) – AI_Adoption_Rate_it measures, for each country–year, the proportion of firms (or sectors, depending on source) that report using AI-based tools in their operations. While coverage and measurement precision vary across sources, the variable is treated as an approximate indicator of the diffusion of AI-related technologies within the economy.
- AI Bias Indicator (AI_Bias_Indicator_it) – AI_Bias_Indicator_it is a composite index intended to proxy the extent to which the institutional and technological environment in a country may give rise to algorithmic bias. It is constructed from information on:
  - the presence or absence of explicit AI fairness or non-discrimination guidelines;
  - the strength of general anti-discrimination enforcement;
  - the degree of transparency and documentation required for algorithmic systems;
  - whether equality bodies have an explicit mandate regarding digital or AI-driven discrimination.

Higher values of AI_Bias_Indicator_it are interpreted as signalling a higher potential for algorithmic bias due to weaker institutional constraints. Given the complexity of the concept and the limitations of available data, this indicator should be understood as an approximate institutional proxy rather than a direct measure of algorithmic behaviour.

- Gender-Sensitive Reform Indicator (Treat_i and Post_t) – For the DiD analysis, we identify countries that have introduced gender-sensitive reforms in the relevant period (e.g. legislation targeting gender equality in digital or AI-related contexts, or significant reforms strengthening equality institutions).
  - Treat_i is a binary variable equal to 1 for countries that implemented such reforms during the sample period and 0 otherwise.
  - Post_t is a binary variable equal to 1 for years after the reform implementation (for treated countries) and 0 otherwise.

The interaction term Treat_i × Post_t captures the treatment effect in the DiD model.

3.2.3 Control Variables

- GDP Growth (GDP_Growth_it) – GDP_Growth_it is the annual real GDP growth rate (%) for country i at time t. It captures the cyclical and structural macroeconomic conditions that may influence labour market outcomes independently of AI.
- Education Years (Education_Years_it) – Education_Years_it measures the average years of schooling in the working-age population. It controls human capital accumulation, which is known to affect labour force participation and wages.
- Additional Sectoral Controls (where available) – In extended specifications, sectoral composition variables (e.g. employment share in services, manufacturing, or high-tech sectors) can be included to capture structural differences in the economy. In the reported main regressions, the focus remains on the core variables above.

3.3. Handling of Missing Data and Sample Restrictions

Given the heterogeneity in reporting across countries and years, not all variables are available for the full 50-country, 2012–2022 window. For the panel-based analyses (FE and DiD), we impose the following restrictions:

- Countries must have at least a minimum number of consecutive observations for FLFP, GWG, AI_Adoption_Rate_it, GDP_Growth_it, and Education_Years_it;
- Countries must have clear information on the timing of gender-sensitive reforms (for DiD).

After applying these restrictions, the FE and DiD models are estimated on a balanced panel of 10 countries observed over 10 years (N = 100). No imputation is used for the core dependent and key explanatory variables in the reported models; observations with missing values are dropped listwise. For the logistic regression, the sample size N = 500 reflects the pooled availability of promotion outcomes and AI-related variables across countries and years.

3.4. Econometric Specifications

The empirical strategy is aligned with the research questions and hypotheses set out in Section 1.1 and the conceptual framework in Section 2.

3.4.1 Logistic Regression for Promotion Outcomes (RQ1; H1a–H1b)

To assess whether AI diffusion and the AI bias indicator are associated with promotion outcomes, we estimate a binary response model of the form:

$$Pr(PROMOit = 1 \mid Xit) = \Lambda(\beta 0 + \beta 1 AI\_Adoption\_Rateit + \beta 2 AI\_Bias\_Indicatorit + \beta 3 GDP\_Growthit + \beta 4 Education\_Yearsit)$$

where PROMO_it = 1 if a female candidate is selected or promoted in observation i,t and 0 otherwise, X_it denotes the vector of explanatory variables, and Λ(·) is the logistic cumulative distribution function.

The coefficients β1 and β2 are of primary interest:

H1a predicts β1 < 0 (higher AI adoption reduces the probability of promotion for women).

H1b predicts β2 < 0 (higher AI bias increases the likelihood of women being disadvantaged in promotion).

Estimation is performed by maximum likelihood. Model fit is assessed via the log-pseudolikelihood, pseudo R^2, and Wald χ^2 tests, as reported in Tables 1 and 2. Because the logistic regression uses pooled data rather than a panel with fixed effects, unobserved heterogeneity at the country or firm level is not explicitly controlled; this limitation is acknowledged in the interpretation of results.

3.4.2 Fixed-Effects Panel Regressions for FLFP and GWG (RQ2; H2a–H2c)

To examine the relationship between AI diffusion and aggregate gender gaps, we estimate country fixed-effects models for female labour force participation and the gender pay gap:

$$Yit = \alpha i + \gamma t + \beta 1 AI\_Adoption\_Rateit + \beta 2 AI\_Bias\_Indicatorit + \beta 3 GDP\_Growthit + \beta 4 Education\_Yearsit + \varepsilon it$$

where Y_it is either FLFP_it or GWG_it; α_i are country fixed effects (capturing time-invariant, unobserved characteristics such as culture, long-run institutions, geography); γ_t are time fixed effects (capturing global shocks and common trends); and ε_it is an idiosyncratic error term.

The coefficients β1 and β2 test H2a–H2c:

H2a predicts β1 < 0 in the FLFP model (higher AI adoption reduces women's participation).

H2b predicts β1 > 0 in the GWG model (higher AI adoption widens the wage gap).

H2c predicts β2 > 0 in the GWG model (higher AI bias is associated with larger wage gaps).

Country fixed effects address omitted variable bias from unobserved, time-invariant country characteristics. Time fixed effects capture common shocks. Estimation is by ordinary least squares (OLS) with the fixed-effects (within) transformation.

The reported results (Tables 3–6) include within, between, and overall R^2 values, as well as F-tests for the joint significance of the regressors. Given the limited number of countries (10) and degrees of freedom (F(4,9) in the FE models), the estimates must be interpreted with caution. In the reported regressions, standard errors are based on conventional FE inference; the small number of clusters is a limitation for robust clustering and is explicitly recognised in the discussion.

3.4.3 Difference-in-Differences for Institutional Reforms (RQ3; H3)

To evaluate whether gender-sensitive institutional reforms are associated with changes in the gender pay gap, we employ a difference-in-differences specification:

$$\begin{aligned} GWGit = \alpha i + \gamma t + \delta(Treati \times Postt) + \theta 1AI\_Adoption\_Rateit \\ + \theta 2AI\_Bias\_Indicatorit + \theta 3GDP\_Growthit \\ + \theta 4Education\_Yearsit + vit \end{aligned}$$

where GWG_it is the gender pay gap; Treat_i is 1 for treated (reform-adopting) countries and 0 otherwise; Post_t is 1 for years after reform implementation and 0 otherwise; δ measures the average treatment effect on the treated (ATT); α_i and γ_t are country and time fixed effects; and ν_it is the error term.

H3 predicts $\delta < 0$: treated countries should experience a relative reduction in the gender pay gap after the introduction of gender-sensitive reforms, compared to untreated countries.

Estimation is again carried out by FE-OLS. The key identifying assumption of the DiD design is that, absent the reforms, treated and controlled countries would have followed parallel trends in the wage gap. Given the small number of countries, formal pre-trend tests are difficult to perform; instead, we rely on qualitative assessment and robustness checks discussed in Section 5. The strong and negative estimated coefficient on the interaction term (Table 7) should therefore be interpreted as suggestive but not definitive causal evidence.

3.5. Estimation, Inference, and Diagnostics

All regressions are estimated using standard econometric software. For transparency, we report:

- Sample sizes (N) and number of groups (countries);
- R^2 measures (within, between, overall) for panel models;
- Wald χ^2 (logit) and F-statistics (FE and DiD) for overall model significance;
- Coefficient estimates, standard errors, t- or z-statistics, p-values, and 95% confidence intervals for all regressors.

Given the modest number of clusters (10 countries), there is a trade-off between using cluster-robust standard errors and preserving degrees of freedom. In the tables reproduced in Section 4, standard errors correspond to the conventional FE and DiD implementations for this sample size. As a result, statistical inference – particularly for marginally significant coefficients – should be interpreted cautiously.

Multicollinearity checks were performed by examining pairwise correlations between regressors and variance inflation factors in auxiliary regressions. No severe collinearity was detected among the core explanatory variables. Nevertheless, the wide confidence intervals on AI-related coefficients suggest that measurement error and limited variation are important concerns, in line with the discussion in Section 5.

4. Empirical Results

This section presents the empirical results of the three econometric specifications described in Section 3: (i) the logistic regression for promotion outcomes (RQ1; H1a–H1b); (ii) the fixed-effects panel regressions for female labour force participation and the gender pay gap (RQ2; H2a–H2c); and (iii) the difference-in-differences (DiD) model for gender-sensitive institutional reforms (RQ3; H3). For each model, we first summarise the main coefficients and model fit statistics, and then discuss their implications in terms of the research questions and hypotheses.

4.1. Logistic Regression: AI and Promotion Outcomes (RQ1; H1a–H1b)

Table 1 reports the estimated coefficients of the logistic regression model for the promotion outcome, while Table 2 summarises the overall model fit.

Table 1. Logistic Regression Results for Promotion Outcome (PROMO_it)

Variable	Coefficient (β)	Std. Error	z-statistic	p-value	95% CI (Lower)	95% CI (Upper)
AI_Adoption_Rate	0.1917	0.3972	0.48	0.629	-0.5867	0.9701
Education_Years	-0.0051	0.0452	-0.11	0.910	-0.0936	0.0834
GDP_Growth	0.0377	0.0314	1.20	0.229	-0.0237	0.0992
AI_Bias_Indicator	-0.4608	0.3165	-1.46	0.145	-1.0811	0.1595
Constant	0.0824	0.5996	0.14	0.891	-1.0929	1.2577

Table 2. Model Fit for Logistic Regression

Statistic	Value
Log Pseudolikelihood	-344.6769
Pseudo R^2	0.0055
Observations (N)	500
Wald $\chi^2(4)$	3.83
Prob > χ^2	0.4293

The coefficients on the AI-related variables are not statistically significant at conventional levels:

AI_Adoption_Rate has a positive coefficient (0.1917), but the p-value (0.629) and wide confidence interval (−0.5867, 0.9701) indicate that we cannot reject the null hypothesis of no association between AI adoption and the probability of a woman being promoted or selected.

AI_Bias_Indicator is negative (−0.4608), which would be consistent with a higher bias being associated with a lower chance of promotion for women, but the p-value (0.145) remains above 0.10, and the confidence interval (−1.0811, 0.1595) again includes zero.

The control variables are also insignificant:

Education_Years is essentially zero (−0.0051, $p = 0.910$), suggesting no detectable relationship between average education and promotion probabilities in this specification.

GDP_Growth is positive (0.0377, $p = 0.229$) but not statistically significant.

Pseudo R^2 is very low (0.0055), and the Wald test does not reject the joint null that all coefficients are zero (Prob > $\chi^2 = 0.4293$), implying that the model has limited explanatory power.

Implications for RQ1 and H1a–H1b

- H1a predicted that higher AI adoption rates would be associated with lower promotion probabilities for women ($\beta_1 < 0$).
- H1b predicted that higher values of the AI bias indicator would be associated with lower promotion probabilities for women ($\beta_2 < 0$).

The estimated coefficients do not support these hypotheses:

- For AI adoption, the coefficient is positive and insignificant.
- For AI bias, the coefficient is negative but insignificant.

Thus, in the available data, there is no statistically robust evidence that AI adoption or the AI bias indicator significantly affects promotion outcomes for women at the aggregate level. RQ1 is therefore answered in the negative: AI diffusion, as measured here, does not appear to be associated with gender-differentiated promotion outcomes in a way that is detectable by this model.

4.2. Fixed-Effects Results: Female Labour Force Participation (RQ2; H2a)

Table 3 presents the fixed-effects estimates with female labour force participation (FLFP_it) as the dependent variable, and Table 4 summarises the model fit.

Table 3. Fixed-Effects Regression Results for Female Labour Force Participation (FLFP_it)

Variable	Coefficient (β)	Std. Error	t-statistic	p-value	95% CI (Lower)	95% CI (Upper)
AI_Adoption_Rate	0.5323	6.3770	0.08	0.935	-13.8934	14.9580
GDP_Growth	-1.2306	0.5256	-2.34	0.044	-2.4196	-0.0416
Education_Years	-0.4989	0.7592	-0.66	0.528	-2.2163	1.2186
AI_Bias_Indicator	-3.2927	5.9421	-0.55	0.593	-16.7346	10.1492
Constant	60.1668	10.8541	5.54	0.000	35.6131	84.7206

Table 4. Model Fit for FLFP Fixed-Effects Regression

Statistic	Value
R^2 (within)	0.0858
R^2 (between)	0.0001
R^2 (overall)	0.0713
Observations (N)	100
Countries (groups)	10
F(4, 9)	2.25
Prob > F	0.1439
Corr(u_i, Xb)	-0.0808

The AI-related variables again do not exhibit significant coefficients:

AI_Adoption_Rate has a coefficient of 0.5323 with a p-value of 0.935 and a very wide confidence interval (−13.8934, 14.9580). This indicates extreme imprecision: the data are consistent with a large negative, a large positive, or zero effect of AI adoption on female labour force participation.

AI_Bias_Indicator is negative (−3.2927, p = 0.593) with a similarly wide confidence interval (−16.7346, 10.1492), again precluding any meaningful inference.

The only statistically significant coefficient is:

GDP_Growth: −1.2306, p = 0.044, with a 95% confidence interval excluding zero (−2.4196, −0.0416). This suggests that higher GDP growth is associated with lower female labour force participation, a pattern that may reflect structural or cyclical dynamics (for example, growth driven by male-dominated sectors). However, a detailed exploration of this mechanism is beyond the scope of this paper, and the result should be interpreted cautiously given the limited sample size.

Overall R^2 and within R^2 values are modest, which is not unusual for panel models with limited regressors and complex underlying processes.

<u>Implications for RQ2 and H2a</u>

- H2a predicted that higher AI adoption rates would be associated with lower female labour force participation ($\beta_1 < 0$).

The FE estimates do not support this hypothesis:

- The sign of β_1 is positive (0.5323), not negative.
- The coefficient is far from statistically significant ($p = 0.935$).

Thus, within this panel of 10 countries over 10 years, no evidence is found that AI adoption is associated with changes in female labour force participation. RQ2, with respect to participation, is therefore answered in the negative.

4.3. Fixed-Effects Results: gender pay gap (RQ2; H2b–H2c)

Table 5 shows the FE estimates with the gender pay gap (GWG_it) as the dependent variable, while Table 6 reports model fit statistics.

Table 5. Fixed-Effects Regression Results for gender pay gap (GWG_it)

Variable	Coefficient (β)	Std. Error	t-statistic	p-value	95% CI (Lower)	95% CI (Upper)
AI_Adoption_Rate	0.8100	3.1381	0.26	0.802	-6.2889	7.9089
GDP_Growth	0.4081	0.2537	1.61	0.142	-0.1659	0.9821
Education_Years	0.0984	0.3948	0.25	0.809	-0.7946	0.9914
AI_Bias_Indicator	-3.5787	2.7796	-1.29	0.230	-9.8666	2.7093
Constant	19.9499	6.9630	2.87	0.019	4.1986	35.7012

Table 6. Model Fit for GWG Fixed-Effects Regression

Statistic	Value
R^2 (within)	0.0617
R^2 (between)	0.0356
R^2 (overall)	0.0514
Observations (N)	100
Countries (groups)	10
F(4, 9)	5.02
Prob > F	0.0209
Corr(u_i, Xb)	-0.161

For the wage gap, neither AI adoption nor the AI bias indicator is statistically significant:

AI_Adoption_Rate: $\beta_1 = 0.8100$, $p = 0.802$, with a 95% confidence interval spanning from −6.2889 to 7.9089. This suggests no detectable relationship between AI adoption and the magnitude of the gender wage gap.

AI_Bias_Indicator: $\beta_2 = -3.5787$, $p = 0.230$, with a confidence interval (−9.8666, 2.7093) that includes zero and even implies a potential narrowing of the wage gap at higher values of the bias indicator – contrary to theoretical expectations. This lack of robustness and the inverted sign further illustrate the imprecision of the estimates.

Control variables are likewise insignificant:

GDP_Growth: positive but only marginally close to significance ($p = 0.142$).

Education_Years: close to zero and insignificant ($p = 0.809$).

The F-statistic indicates that at least some regressors jointly explain variation in GWG (Prob > F = 0.0209), but no individual coefficient (other than the constant) reaches conventional significance. The modest R^2 values again reflect the complexity of the underlying processes and the limited number of regressors.

Implications for RQ2 and H2b–H2c

- H2b predicted that higher AI adoption would be associated with a larger gender pay gap ($\beta_1 > 0$).
- H2c predicted that higher AI bias would be associated with a larger gender pay gap ($\beta_2 > 0$).

The FE results do not provide support for these hypotheses:

- For H2b, the coefficient on AI adoption is positive but extremely imprecise and statistically insignificant.
- For H2c, the coefficient on the AI bias indicator is negative and insignificant, contradicting the hypothesised direction.

Therefore, within this panel, we do not find empirical evidence that AI adoption or AI-related bias are associated with changes in the gender pay gap at the country level. RQ2, with respect to wage gaps, is likewise answered in the negative.

4.4. Difference-in-Differences Results: Gender-Sensitive Reforms (RQ3; H3)

The DiD specification evaluates whether the introduction of gender-sensitive institutional reforms is associated with changes in the gender pay gap. Table 7 presents the coefficient estimates, while Table 8 reports key model statistics.

Table 7. Difference-in-Differences Results for gender pay gap (GWG_it)

Variable	Coefficient	Std. Error	t-statistic	p-value	95% CI (Lower)	95% CI (Upper)
AI_Adoption_Rate	1.2419	3.1205	0.40	0.700	-5.8172	8.3011
GDP_Growth	0.4434	0.2365	1.88	0.094	-0.0915	0.9784
Education_Years	0.1825	0.3674	0.50	0.631	-0.6488	1.0137
AI_Bias_Indicator	-2.8440	2.6907	-1.06	0.318	-8.9307	3.2427
post_treatment	2.8985	0.4397	6.59	0.000	1.9039	3.8931
did_interaction	-6.4384	0.8729	-7.38	0.000	-8.4129	-4.4638
Constant	17.3228	6.2805	2.76	0.022	3.1152	31.5304

Table 8. Model Fit for DiD Regression

Statistic	Value
Observations (N)	100
Countries (groups)	10
R^2 (within)	0.1277
R^2 (between)	0.0503
R^2 (overall)	0.0770
F(6, 9)	16.96
Prob > F	0.0002

The coefficients on the AI-related variables again remain insignificant:

AI_Adoption_Rate: 1.2419, $p = 0.700$, with a wide confidence interval (−5.8172, 8.3011).

AI_Bias_Indicator: −2.8440, $p = 0.318$, with a confidence interval crossing zero.

The key results concern:

post_treatment: 2.8985, $p < 0.001$. This coefficient reflects the average change in the gender pay gap in the post-reform period across all countries, treated and controlled. Its positive and significant value indicates that, absent further controls, wage gaps tended to increase over time in the sample as a whole.

did_interaction (Treat_i × Post_t): −6.4384, $p < 0.001$, with a confidence interval entirely below zero (−8.4129, −4.4638). This is the DiD estimator of the average treatment effect on the treated (ATT).

The negative and significant DiD coefficient implies that countries which implemented gender-sensitive reforms experienced a 6.4 percentage point lower wage gap (or smaller increase) relative to what would be expected given the general upward trend in wage gaps. In other words, reforms are associated with a significant relative improvement in wage outcomes for women compared to non-reforming countries.

Implications for RQ3 and H3

- H3 predicted that countries adopting gender-sensitive institutional reforms experience a relative reduction in the gender pay gap ($\delta < 0$).

The DiD estimates provide strong support for H3:

- The estimated δ (did_interaction) is large in magnitude, negative, and highly significant.
- This result holds after controlling for AI adoption, AI bias, GDP growth, education, and fixed effects.

Subject to the usual DiD assumptions – particularly parallel trends in the absence of reforms – the evidence suggests that gender-sensitive reforms are associated with meaningful reductions in gender wage gaps.

5. Discussion

This section interprets the empirical findings in light of the research questions and theoretical expectations developed in Sections 1 and 2. It proceeds in four steps. First, it synthesizes the main empirical results. Second, it discusses why macro-level data may fail to detect algorithmic bias, even when micro-level evidence suggests its presence. Third, it examines the role of gender-sensitive institutional reforms suggested by the DiD estimates. Finally, it outlines the main limitations of the study and their implications for future research.

5.1. Summary of Empirical Findings

The empirical analysis yields a clear and internally consistent pattern. Across three econometric approaches – logistic regression, fixed-effects panel models, and difference-in-differences – the following results emerge:

- No significant AI effect on promotion outcomes (RQ1; H1a–H1b) – The logistic regression model shows that neither AI adoption nor the AI bias indicator has a statistically significant association with women's promotion outcomes. The coefficients are small relative to their standard errors, and the confidence intervals include zero with substantial margins on both sides. The model's low pseudo R^2 and non-significant Wald test further indicate limited explanatory power. Thus, the data do not support the hypothesis that greater AI adoption or higher AI bias systematically reduces women's promotion probabilities at the aggregate level.
- No significant AI effect on female labour force participation or wage gaps (RQ2; H2a–H2c) – The fixed-effects regressions for female labour force participation and the gender pay gap similarly provide no evidence that AI adoption or the AI bias indicator are significant predictors of changes in these outcomes. Coefficients on AI-related variables remain statistically insignificant with wide confidence intervals. The sign patterns are inconsistent with the directional hypotheses in some cases, particularly for the AI bias indicator in the wage gap model. The only notable coefficient is the negative and significant association between GDP growth and female labour force participation; while interesting, this result lies outside the core focus of AI and gender and is therefore interpreted cautiously.
- Significant effect of gender-sensitive reforms on wage gaps (RQ3; H3) – In contrast, the difference-in-differences specification shows that countries implementing gender-sensitive institutional reforms experience a statistically significant relative reduction in gender wage gaps compared to non-reforming countries. The DiD interaction term is negative, large in magnitude (approximately −6.4 percentage points), and highly significant. This result holds after controlling for AI adoption, AI bias, macroeconomic conditions, and fixed effects, and remains robust in terms of sign and significance across reasonable variations of the specification.

Taken together, these findings imply that, in the period and countries examined, AI adoption and AI-related institutional proxies are not detectable macro-level drivers of gender inequality, while gender-sensitive institutional reforms are associated with meaningful improvements in wage outcomes for women.

5.2. Why Macro-Level Data May Not Detect Algorithmic Bias

The absence of significant coefficients on AI adoption and AI bias indicators does not contradict the micro-level literature documenting instances of algorithmic discrimination. Rather, it suggests that micro-level biases do not straightforwardly scale into macro-level patterns that can be captured in cross-country panel data. Several mechanisms can explain this discrepancy.

5.2.1 Aggregation Effects and Heterogeneity

Algorithmic systems are deployed in specific firms, sectors, and occupations. Bias in one firm's hiring algorithm may disadvantage women competing for a particular type of job, but have a negligible impact on aggregate female labour force participation or national wage gaps. When outcomes are aggregated to the country level, idiosyncratic biases may average out, especially if AI penetration is uneven across sectors or if biased and unbiased systems coexist. Furthermore, the direction of AI's impact may differ across sectors. For example, algorithmic tools in high-wage sectors may disadvantage women, while tools in other sectors (e.g. services) may be neutral or even beneficial. When such heterogeneous effects are aggregated, the net effect can be close to zero. The insignificant AI coefficients and wide confidence intervals in our models are consistent with this possibility.

5.2.2 Dominance of Structural and Institutional Determinants

Female labour force participation and gender wage gaps are shaped by a wide range of structural factors: educational attainment, occupational segregation, fertility and childcare arrangements, tax and benefit systems, labour market regulation, collective bargaining, and social norms. These determinants evolve slowly and exert strong, pervasive influences on gender outcomes. By contrast, AI adoption in employment processes is relatively recent and non-uniform, particularly in the time window used here (2012–2022). It is therefore plausible that the marginal contribution of AI to aggregate gender gaps is small in magnitude compared to these structural forces. Even if bias exists at the algorithmic level, its incremental effect may be too modest to shift country-level indicators in a statistically detectable way.

5.2.3 Measurement Error and Proxy Limitations

A further explanation lies in the measurement of the AI-related variables. AI adoption at the country–year level is approximated using available indicators that may not precisely capture the intensity or specific use of AI in hiring and promotion processes. Similarly, the AI bias indicator is a composite proxy for the institutional environment rather than a direct measure of algorithmic behaviour. From an econometric perspective, classical measurement error in explanatory variables tends to bias coefficient estimates towards zero. Thus, even if AI adoption has a true but moderate effect on gender outcomes, noisy measurement may attenuate the estimated coefficients sufficiently that they appear insignificant in our regressions. The unusually wide confidence intervals around the AI coefficients are consistent with this interpretation.

5.2.4 Temporal and Diffusion Lags

Finally, the period studied may be too early in the diffusion of AI for its effects to manifest strongly at the macro level. Adoption of advanced AI systems in recruitment and human resource management has accelerated only in recent years, and the time horizon for such systems to affect lifetime earnings and labour force participation is long. If AI's impact is cumulative or non-linear, a longer time series or more recent data might be necessary to detect its influence on aggregate gender outcomes. In particular, early years in the 2012–2022 window capture a pre-generative-AI environment and substantially lower penetration of AI-enabled HR tools, which increases heterogeneity and can attenuate estimated average effects in panel regressions.

5.3 The Role of Gender-Sensitive Institutional Reforms

While AI-related variables show no significant influence on gender outcomes, the DiD estimates reveal a robust association between gender-sensitive institutional reforms and reductions in wage gaps. This result aligns with the governance and political economy literature emphasising that institutions, not technology alone, shape distributional outcomes. The reforms captured in our treatment indicator typically include one or more of the following elements:

- stronger equality legislation or enforcement mechanisms with updated coverage of digital and AI-mediated discrimination;
- enhanced transparency and documentation requirements for algorithmic decision systems;
- targeted policy initiatives aimed at reducing gender wage gaps or promoting equal opportunities.

The negative and significant DiD coefficient suggests that, relative to non-reforming countries, those that implemented such measures experienced a substantive relative improvement in women’s wage outcomes, even as global trends moved in the opposite direction (as indicated by the positive and significant post-treatment coefficient). From an economic perspective, this supports a view in which AI operates within an institutional envelope. In contexts with weak governance and limited equality infrastructure, algorithmic tools may reproduce existing structural biases. In contexts with stronger institutional safeguards, their potential harms can be mitigated, and broader efforts towards gender equality may dominate any technological effects. Our results indicate that, at the macro level, institutional reforms are empirically more salient than AI adoption in explaining cross-country variation in wage gaps over time.

It is important to emphasise that the DiD result, while statistically strong, is still subject to the usual identification assumptions – particularly parallel trends in the absence of treatment. Given the relatively small number of treated and control countries, these assumptions cannot be tested as rigorously as in larger samples. Nonetheless, the magnitude, sign, and robustness of the DiD coefficient make it a meaningful and policy-relevant finding.

6. Conclusion

This paper set out to examine whether the diffusion of artificial intelligence (AI) is associated with measurable changes in gendered labour market outcomes at the macro level, and whether gender-sensitive institutional reforms alter this relationship. The evidence pertains to the 2012–2022 period; given the speed of AI diffusion since 2023, future updates using post-2022 data are necessary before drawing conclusions about the current AI wave. The analysis was organised around three research questions and three sets of hypotheses. RQ1 and H1a–H1b asked whether AI adoption and an AI bias indicator were associated with gender-differentiated promotion outcomes. RQ2 and H2a–H2c extended this logic to macroeconomic gender gaps in female labour force participation and wages. RQ3 and H3 shifted the focus from technology to institutions, asking whether gender-sensitive reforms are empirically associated with narrower gender wage gaps.

The empirical results for RQ1 and RQ2 are striking in their consistency. Across the logistic regression for promotion outcomes and the fixed-effects models for female labour force participation and the gender pay gap, the coefficients on AI adoption and the AI bias indicator are statistically insignificant and imprecisely estimated. In the promotion model, the estimates do not reveal any systematic relationship between AI-related variables and women's likelihood of being selected or promoted. In the panel regressions, AI adoption shows neither a robust negative association with female participation nor a robust positive association with wage gaps. The AI bias indicator similarly fails to display the expected positive correlation with gender wage gaps and, in some specifications, even takes on the opposite sign, though without statistical significance. Taken together, these results imply that H1a, H1b, H2a, H2b, and H2c are not supported by the data.

The rejection of these hypotheses is not a trivial outcome. Theoretical and micro-level studies have often suggested that AI systems replicate or amplify pre-existing gender biases, particularly in hiring and evaluation. If such mechanisms were pervasive, large in magnitude, and homogeneous across settings, one might expect to observe at least modest correlations between AI diffusion and aggregate gender indicators. The absence of such correlations in the present analysis suggests that, over the period and countries studied, micro-level algorithmic biases do not translate into macro-level shifts that are detectable in cross-country panel data. The empirical relationship between AI diffusion and gendered labour outcomes appears to be weak, heterogeneous, or heavily mediated by structural and institutional factors.

This interpretation is reinforced by the deep contrast between the AI-related null results and the strong finding associated with RQ3 and H3. The difference-in-differences estimates show that gender-sensitive institutional reforms are associated with a statistically significant relative reduction in gender wage gaps, even after controlling for AI adoption, AI bias proxies, macroeconomic conditions, and fixed effects. While the post-treatment dummy captures a general tendency for wage gaps to increase over time across all countries, the negative and sizeable interaction term indicates that reform-adopting countries deviate from this trend: conditional on common dynamics, they exhibit substantially narrower wage gaps than would otherwise be expected. In contrast to the non-significant AI coefficients, the reform effect is both economically and statistically meaningful.

The combined pattern of findings points to a deeper conclusion concerning the political economy of AI and gender inequality. The failure of H1 and H2, together with the confirmation of H3, suggests that AI adoption, at least in its current forms and measurement, is not an independent structural determinant of gender inequality in labour markets. Instead, AI appears as one element embedded within a broader institutional and structural environment that continues to be the primary determinant of aggregate gender outcomes. Institutions, not technology per se, are where the statistically visible leverage lies: when legal and policy frameworks change in favour of gender equality, wage gaps respond; when AI adoption increases without corresponding institutional shifts, the data do not reveal a consistent pattern in gendered outcomes.

This does not imply that algorithmic bias is irrelevant. Rather, the findings indicate that its effects are likely to be local, sector-specific, and mediated by context, so that they are either too small or too heterogeneous to generate clear macro-level correlations in the available data. The null results for AI can thus be read as evidence of a decoupling between the popular narrative of AI as a central driver of gender inequality and the observable behaviour of country-level labour market indicators. In contrast, the robust association between reforms and wage gaps shows that where the institutional environment is reshaped in a gender-sensitive direction, the distributional consequences are sufficiently strong and coherent to be visible even in relatively small samples.

From the standpoint of the research questions, the conclusion is therefore threefold. First, regarding RQ1, the analysis finds no empirical support for the claim that AI adoption or AI-related bias proxies are associated with gender-differentiated promotion outcomes in a manner that is detectable in aggregate data. Second, regarding RQ2, there is no evidence that AI diffusion at the country level systematically influences female labour force participation or the gender pay gap; structural inequalities appear to be driven by forces other than the measured dimensions of AI adoption. Third, regarding RQ3, the evidence indicates that gender-sensitive institutional reforms are correlated with meaningful reductions in gender wage gaps, even in the presence of general trends moving in the opposite direction. In other words, the data reject the centrality of AI as a macro-driver of gender inequality, but they affirm the centrality of institutions.

Interpreted as a whole, the study suggests that current debates may overstate the independent macroeconomic role of AI in shaping gender inequality while understating the continuing importance of institutional design and enforcement. AI may well embody gendered logics at the micro level, but its aggregate effects are neither deterministic nor autonomous; they are filtered through, and often overshadowed by, existing labour market structures and equality regimes. For scholars and policymakers alike, the implication is that research and interventions aimed at gender equality in the age of AI should focus less on technology as an isolated causal agent and more on the institutions that govern how technology is developed, deployed, and constrained.

Acknowledgments

The authors would like to thank colleagues at the University of Tetova for their valuable discussions and feedback during the development of this research.

Funding

This research received no external funding.

Contribution of individual authors

Aida Yzeiri Baftiari conceptualized the study and contributed to the literature review and theoretical framework.

Artina Kamberi designed the econometric methodology, contributed to data collection, dataset construction, and empirical analysis.

Agon Memeti supervised the research process, and contributed to the interpretation of results and final revision of the manuscript.

Conflict of Interests

The authors declare no conflict of interest.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors did not use any generative AI or AI-assisted technologies in the writing process. The authors take full responsibility for the content of the publication.

References

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. – In: Friedler, S. A., Wilson, C. (eds.). Proceedings of the 1st Conference on Fairness, Accountability and Transparency (FAT 2018), pp. 149-159. Proceedings of Machine Learning Research, 81.

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. – Advances in Neural Information Processing Systems, 29, pp. 4349-4357.

Cowgill, B., Dell'Acqua, F., Matz, S. (2020). The managerial effects of algorithmic fairness activism. – AEA Papers and Proceedings, 110, pp. 85-90.

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Retrieved from https://www.reuters.com.

Lambrecht, A., Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. – Management Science, 65(7), pp. 2966-2981.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A. (2021). A survey on bias and fairness in machine learning. – ACM Computing Surveys, 54(6), pp. 1-35.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. – In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)*, Association for Computing Machinery, pp. 33-44. https://doi.org/10.1145/3351095.3372873.

West, S. M., Whittaker, M., Crawford, K. (2019). Discriminating systems: Gender, race and power in AI. AI Now Institute.

Williams, R. (2021). Rethinking administrative law for algorithmic decision making. – Oxford Journal of Legal Studies, 42(2), pp. 468-494.