A new turning point for women: artificial intelligence as a tool for reducing gender discrimination in hiring *

Elena Pisanelli†

Abstract

Can artificial intelligence (AI) perform better than human recruiters? Can AI help reduce gender disparities in employment outcomes? I analyze whether firms' adoption of AI has a causal effect on their probability of hiring female managers, using data on the 500 largest firms by revenues in Europe and the US, and a staggered difference-in-differences approach. I find that, despite the concerns that the existing literature prompts about AI fairness, firms' use of AI cause, on average, an increase by $\sim$2% in the hiring of female managers. This result is driven by the use of screening AI, while the effect of predictive AI cannot be claimed statistically different from zero. This result is best explained by AI leading to a reduction in gender discrimination in hiring. In fact, I find firms' use of AI to be correlated with a reduction in gender discrimination lawsuits.

Keywords: artificial intelligence, gender, hiring, discrimination
JEL Codes: J71, M51

*The author gratefully acknowledges Klarita Gërxhani; Arnout van de Rijt; Emilio J. Castilla; Giovanni Mastrobuoni for the insights and precious feedback. William Foley; Amanda Haraldsson; Mena Lüsse; Jonathan Winter; Guido Pialli; Luca Favero for the fruitful discussions.

†European University Institute, Florence, IT, elena.pisanelli@eui.eu

1 Introduction

Gender inequality in the labor market persists even if the market changes. Even if women have moved into paid labor market and have increasingly embraced managerial careers, they continue to be underrepresented in leadership positions (Roth, 2007; Hewlett et al., 2010). The literature has identified gender discrimination as one of the major factors contributing to the reproduction of gender inequality in firms' managerial pools (Eagly, 1997; Ridgeway and Bourg, 2004; Ridgeway, 2011). To address gender discrimination in hiring, firms have started to rely on AI, considering the real promise that hiring algorithms can help mitigate recruiters' stereotypes and ensure objectivity to the hiring process (Langenkamp et al., 2019). However, existing research prompts concerns about this so-called meritocratic procedure, since AI selects the best applicant to hire by relying on existing data on the local labor market and firms established hiring preferences (O' Neil, 2016; DeBrusk, 2018). For example, AI algorithms may evaluate a female worker who suffers from negative performance evaluations by her employer because of discrimination rather than meritocratic evaluation as low-performing, non-employable, or not promotable (Cowgill and Tucker, 2020; Gebru, 2020; Cowgill, 2019; O' Neil, 2016).

How comes that a technology aimed at getting rid of human discrimination in decision-making processes results in discriminating women in hiring? The existing literature that aims to explain whether and how AI algorithms may present biased outcomes well documents the sources and consequences of biases in AI, however, it has a tendency to especially focus on pre-hiring processes, that is, it studies what happens during the hiring process before firms hire workers (Datta et al., 2015; Cowgill, 2019; vanEsch et al., 2019; Lambrecht and Tucker, 2019). It leaves, thus, open questions on the effect of AI on key employment outcomes, such as whether and how AI affects the underrepresentation of women in certain jobs or the gender wage gap. This paper, therefore, aims to exploit unique firms' data to provide a causal answer to the above open question by looking at the effect of firms' adoption of AI on their probability to hire female managers and thus reduce female under-representation in firms' top-level positions.

My argument is that the existing under-representation of women in managerial positions arise from discrimination based on status, defined as "a ranking in a hierarchy that is socially recognized and typically carries with it the expectation of entitlement to certain resources" (Ball et al., 2001, p.161). Firms' use of AI can perpetuate or reduce such status-based discrimination, according to how AI selects the best candidate to hire. One the one hand, if firms use predictive AI — algorithms that predict the best job applicant to hire based on incumbent employees — they may perpetuate the existing gender disparities in managerial positions. Predictive AI can identify exiting female under-representation in firms' incumbent managerial pool as performance-based rather than discrimination-based and thus translate the existing gender discrimination in managerial positions into biased hiring outcomes (Cowgill and Tucker, 2020; Gebru, 2020; Daugherty et al., 2019; Cowgill, 2019; O' Neil, 2016). On the other hand, if firms use screening AI — AI algorithms that screen job applicants' resumes and select the best applicant to hire based on the information on the resumes — they may reduce the existing gender disparities in managerial positions. Screening AI can help firms overcome status-based discrimination since it can be taught to ignore data about race, gender, sexual orientation, and other characteristics that are not relevant to the hiring decision (Daugherty et al., 2019; Silberg and Manyika, 2019).

To test this argument, I analyze whether firms' use of AI in hiring has a causal effect on their probability to hire female managers, using data on the 500 largest companies by revenues in Europe and the US over 8 years and employing a staggered difference-in-differences approach.

I find that, despite the concerns that the existing literature prompts about algorithmic fairness (Cowgill and Tucker, 2020; Gebru, 2020; Cowgill, 2019; O' Neil, 2016), firms' use of AI cause, on average, an increase by $\sim$2% in the hiring of female managers. This result is driven by the use of screening AI, while the effect of predictive AI cannot be claimed statistically different from zero. This result is best explained by AI leading to a reduction in gender discrimination in hiring. In fact, I find firms' use of AI to be correlated with a reduction in gender discrimination lawsuits.

2 AI in hiring

Scholars have largely documented the association between gender and career outcomes. Men, for example, are more likely than women to supervise workers of the other sex and to dominate the top-level positions in their companies (Reskin and Bielby, 2005). A substantial number of studies have, thus, sought to explain why such gender disparities in the labor market persist. Key findings in the literature include statistical discrimination and status-based discrimination by employers (Correll and Benard, 2006), and prevalence of women in queues for low pay, low-status jobs (Fernandez and Mors, 2008) as major factors contributing to reproduce gender inequality in the labor market. Considering the main findings of the existing literature that studies the persistence of gender disparities in managerial positions (Roth, 2007; Stone, 2007; Hewlett et al., 2010), we know gender discrimination exists based on human agents' hiring choices. For example, employers hire, on average, more men than women in managerial positions because the gender stereotype that associates higher emotionality with women than men is incompatible with employers' stereotype of the agentic (ambitious, dominant, aggressive) manager (Brescoll, 2016).

In recent years, technological advancements in data collection and processing have paved the prospect of using technology to overcome such conscious or unconscious stereotypes in hiring. Firms, thus, have adopted AI under the promise to make the hiring process not only more efficient and effective but also supportive of workplace equality (Bildfell, 2019; Langenkamp et al., 2019; Newman et al., 2020). Such promise is based on the belief that, unlike human recruiters who may be subject to unconscious prejudice, AI can promote workplace equality because it is driven by hard data and mathematical evidence (Bildfell, 2019; Newman et al., 2020). In fact, AI can have the potential to reduce stereotypes and improve labor market inequalities (Cowgill and Tucker, 2020) if it is trained to focus on candidates' job-relevant characteristics while discarding their demographic indicators when evaluating their future performance if hired (Bildfell, 2019; Daugherty et al., 2019). This means that AI can enable firms to relieve hiring decisions from stereotypes and prejudices, and to improve fairness of the overall hiring process (Newman et al., 2020).

While this line of reasoning is possible, it seems inconsistent with the existing research on algorithmic fairness. Datta et al. (2015), and Lambrecht and Tucker (2019), in fact, show that in settings where AI algorithms advertise job positions, historically discriminated-against groups are less likely to see desirable job advertisements, even when the job advertisement seems to be gender neutral in principle. To understand the mechanisms behind AI algorithms' amplification of human biases, Cowgill (2019) develops a theoretical model of decision-making, comparing in a counterfactual way AI algorithms' decisions with human decisions. Cowgill (2019) points to discriminatory patterns in employers and AI algorithms' programmers' decisions as the primary mechanism leading to AI algorithms amplifying existing biases. More precisely, Cowgill (2019) argues that without enough noise and inconsistency in human decisions, AI algorithms cannot correct for human biases, learning to discriminate.

My theoretical account of AI argues against this strand of research which claims that AI can inherit and even amplify existing gender stereotypes and discrimination in the labor market and that major technological advances, such as AI, can show how technology itself may cause power imbalances and (un) intended consequences that can harm historically discriminated against social groups (Gebru, 2020; O' Neil, 2016). I claim that, conditional on a transparent and accountable use, AI can actually represent a powerful tool for firms to reduce demographic disparities in labor market outcomes and promote gender equality. Because of the above concerns about algorithmic fairness, both scholars and practitioners have recognized the importance of developing techniques to ensure that AI makes fair decisions (Silberg and Manyika, 2019). One major class of approaches, for example, aims at reducing AI's ability to predict job candidates' gender and ethnicity and at minimizing correlations between those demographic attributes and AI's hiring outcomes (Silberg and Manyika, 2019).

Since AI can be taught to minimize the correlation between gender and work-related characteristics, firms' use of AI in hiring can lead to a reduction in gender disparities in employment outcomes. In this scenario, thanks to an accurate evaluation of job applicants' quality, done through the analysis of a rich set of job candidates' work-related information, AI can help firms reduce employers' discrimination against women based on status (defined as in Ball et al. (2001)). Because status-based discrimination explains most of the underrepresentation of women in managerial positions (Stone, 2007) and has contributed to gender gaps in the labor market (Crittenden, 2001), a reduction of such type of discrimination would make AI a tool for healing the persistent gender inequality at firms' top-level positions.

3 Data and Method

3.1 Data

I perform the analyses in this paper on those European and American firms that entered Fortune Global 500 in 2021, that is, those European and American companies ranked as the top 500 corporations worldwide as measured by revenue. The main reason to focus on such specific firms is their financial and organizational similarity, that allows me to study the effect of firms' use of AI on their probability of hiring female managers by considering firms that use AI in hiring and their most similar counterfactual firms that do not use it. To do so, I rely on the Orbis database by Bureau van Dijk to collect, for the years going from 2013 to 2021, data on where each firm's headquarters are based, the industry in which the firm operates, firm's productivity, total assets, return on equity, profit margin, net income, number of employees, and share of female directors. Data on firms' use of AI in hiring are, instead, much harder to get. In order to know whether firms use AI in hiring and when they have adopted it, I manually extracted information on the use of AI from firms' public available annual integrated reports, which contain information on the strategies that firms undertake to reach sustainability and social goals. Figure 1 shows an example of the information regarding firms' use of AI in hiring on firms' annual integrated reports.

I assign each firm to the AI treatment ($D_i = 1$) if its annual integrated report contains evidence of the use of AI in hiring. The paper relies on one main *assumption*; those firms providing evidence of using AI in hiring are those firms being transparent and deploying a "good" AI, checked for hidden biases. For instance, Amazon does not provide any information on its use of AI in hiring and it was one of the first companies that was under fire in 2018 because its hiring algorithm was reportedly discriminating against women.

Using technology to promote diversity

In an effort to ensure that we are presenting the Group and all of its career opportunities in the fairest way and to the widest possible pool of potential candidates, we have recently introduced a language analytics application to identify gender tone and eliminate unconscious bias. This tool will be fully integrated in to the existing human resources and communication system framework to make everyday impartiality an efficient and instinctive reality.

Using advanced analytics, the platform strategically matches students with employers. The internX platform also provides the students with support in the form of skills assessments, mock interviews and soft skills training, so that students are prepared when an organization reaches out and expresses interest.

Figure 1: Information on the use of AI in hiring on firms' annual integrated reports

My dependent variable is firms' probability of hiring female managers. I collect this variable from the Orbis database, from the years going from 2013 to 2021. The variable is a dummy that takes on value 1 if the firm has hired at least one female manager in the reference year, zero otherwise.

The identification strategy, presented in Section 3.2, relies on two stages: (i) I match, based on observable characteristics, treated and control firms using Mahalanobis distance matching; (ii) I perform a staggered difference-in-differences estimation with multiple time periods, relying on Callaway and Sant'Anna (2021)'s approach, because not all treated firms in my dataset have adopted AI in hiring in the same year.

Table 1 reports descriptive statistics of all variables over the full sample and the restricted matched sample.

Table 1: Summary statistics

	Full sample		Restricted matched sample	
	Mean	SD	Mean	SD
Probability of female hire	0.026	0.05	0.024	0.042
Use of AI	0.107	0.309	0.298	0.458
Log assets	25.262	1.469	25.748	1.368
Log productivity	24.406	0.824	24.349	0.528
Net income (millions)	12.357	18.577	12.629	13.382
Profit margin	13.286	14.108	15.353	11.996
Return on equity	27.986	67.858	16.954	12.365
Log employees	11.278	1.115	11.489	0.63
Share of female directors	0.052	0.044	0.053	0.039
N	1,621		531	

The Appendix reports the distribution of the covariates used for matching treated and control firms with Mahalanobis distance matching (for each year between 2013 and 2021). Existing research shows that the probability that firms introduce innovation in their processes or products significantly depends on firms' size (Parisi et al., 2006). I, therefore, match treated and control firms — before estimating through staggered difference-in-differences the effect of AI on firms' probability to hire female managers — on measures of firms' size, that is, total assets and number of employees (Damanpour, 1992). Further, the probability that firms adopt innovations both within their processes and products is strongly associated with firms' profits and financial resources (Parisi et al., 2006). I, therefore, match treated and control firms also on productivity, profit margin and return on equity. Last, since existing studies show that the relationship between innovation and the above variables is not the same across countries and industries (Damanpour, 1992), I match treated and control firms on industry (4 digits NACE code) and country.

3.2 Empirical strategy

3.2.1 Mahalanobis distance matching

The first step of the empirical strategy resorts to using Mahalanobis distance matching to balance the treated and control firms on observable covariates. Relying on Mahalonobis distance matching before the difference-in-differences estimate allows to account for the systematic dynamic differences between those firms which use AI in hiring and those which do not.

For each treated firm, with treatment D_i defined as the use of AI in hiring ($D_i = 1$), I find all available untreated firms ($D_i = 0$) with the most similar — in terms of Mahalanobis distance — variables $\{x\}$ that may determine differences in firms' adoption of AI in hiring (firms' size, firms' productivity, profit margin, return on equity, industry, and country). Equation 1 presents the econometric specification of the Mahalanobis distance definition.

$$d(u, v) = (u - v)^T C_{OR}^{-1}(u - v) \tag{1}$$

With u and v values of $\{x^T, \hat{q}(x)\}^T$, where x are the observable covariates and $\hat{q}(x)$ is the estimated log odds against exposure to treatment; and C_{OR} sample covariance matrix of $\{x^T, \hat{q}(x)\}$ in the control group (Rosenbaum and Rubin, 1985).

3.2.2 Staggered difference-in-differences on matched firms

I estimate, for matched firms, the effect of using AI on firms' probability of hiring female managers through a staggered difference-in-differences technique with multiple time periods, relying on Callaway and Sant'Anna (2021)'s approach.

The setup comprises multiple treatment groups defined by each year of AI adoption. More precisely, equation 2 presents the econometric specification of the estimated group-time

ATT for firms who are members of each treated group g — with g defined as the year of AI adoption by each firm — at each time period t, as in Callaway and Sant'Anna (2021):

$$ATT(g,t) = \mathbf{E}[Y_t(g) - Y_t(0)|G_g = 1] \tag{2}$$

With Y_t firms' probability of hiring female managers for each year t.

When estimating $ATT(g,t)$, I weight each observation by the weight generated with Mahalanobis distance matching, in order to condition the estimate of the $ATT(g,t)$ on the observable covariates.

In order for the estimated ATT to be valid and reliable, a series of assumptions should be imposed (Callaway and Sant'Anna, 2021).

Assumption 1: Limited treatment anticipation. The assumption states that firms should not anticipate treatment by any period. The assumption is very likely to hold in the context of this paper, since it is unlikely that firms increase the hiring of female managers in sight of AI adoption. However, I account for potential anticipation of the treatment by allowing for anticipatory behavior and imposing conditional parallel trends in pre-treatment periods, making the parallel trend assumption (discussed in the next paragraph) stronger.

Assumption 2: Conditional parallel trends based on a never-treated group. as Callaway and Sant'Anna (2021) suggest, I compare treated firms with never-treated ones, rather than with not-yet-treated firms, since I have a sizeable group of firms that do not participate in the treatment in any period (around 70% of the firms in my restricted matched sample). The assumption imposes that, conditional on covariates, the average outcomes for the firms first treated in group g (with g year of AI adoption) and for the never-treated firms would have followed parallel paths in the absence of the treatment. Section 4 provides evidence of the validity of the conditional parallel trend assumption.

4 Results

4.1 Balancing of observable covariates

Before discussing the $ATT(g, t)$ results, estimated through the staggered difference-in-differences approach, this section reports evidence regarding the balance of the observable covariates x among treated and control firms before and after matching. In particular, Figure 2 presents the balancing achieved after the Mahalanobis distance matching.

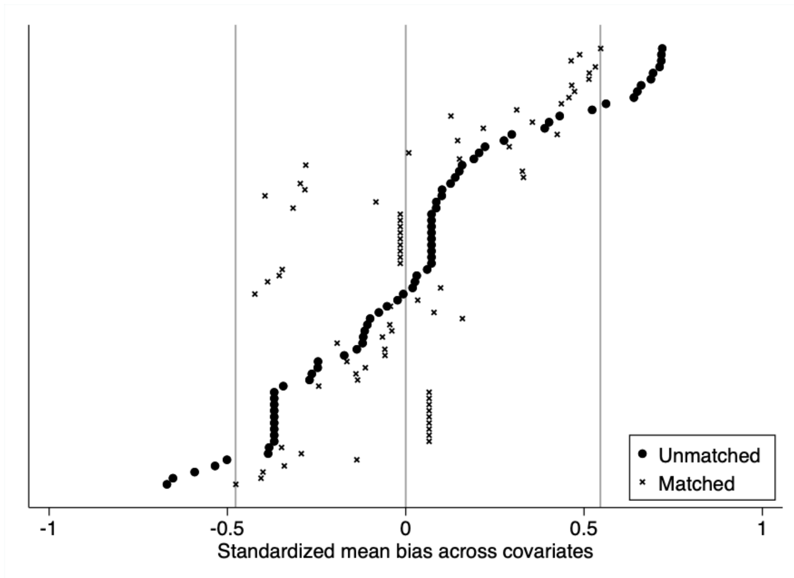


Figure 2: Balancing of the observable covariates achieved after the Mahalanobis distance matching.

Figure 2 shows the standardized mean bias for all covariates before and after matching, that is the difference of the means in the treated and non-treated firms as a percentage of the square root of the average of the sample variances in the treated and non-treated groups (Rosenbaum and Rubin, 1985). The mean absolute standardized bias across covariates after matching is 21.6, which is smaller than the absolute standardized mean bias across covariates before matching (29). As Figure 2 shows, matching reduced the standardized mean bias to less than ~0.5 for all covariates. Refinement is desirable, but matching has done well at balancing the treated firms and their control counterparts, adjusting reliably for all the covariates.

4.2 Effect of AI on firms' probability of hiring female managers

This section presents the average treatment effect for the treated ($ATT(g,t)$) of firms' use of AI on their probability of hiring female managers. Figure 3 shows the graphical representation of the standardized $ATT(g,t)$ estimate. Table 2 reports the event study and simple weighted average standardized estimates of the $ATT(g,t)$.

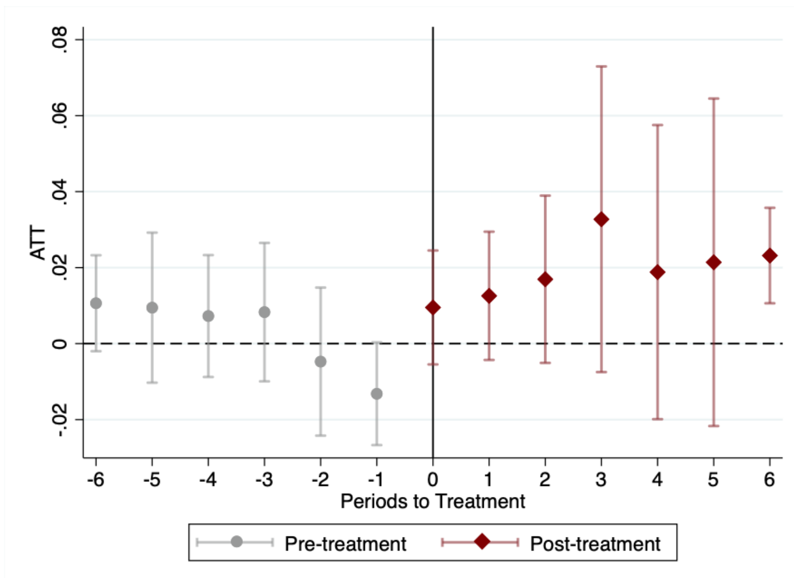


Figure 3: Staggered difference-in-differences results. Event study

Table 2: Staggered difference-in-differences results.

	Probability of hiring a female manager	
	(1)	(2)
	Simple weighted average	Event study
Firms' use of AI	0.017**	
	(0.007)	
Pre-treatment average		0.003
		(0.002)
Post-treatment average		0.019**
		(0.008)
Managers' age	✓	✓
Total assets	✓	✓
Number of employees	✓	✓
Productivity	✓	✓
Profit margin	✓	✓
Net income	✓	✓
Country	✓	✓
Industry	✓	✓
Return on equity	✓	✓
Share of female directors	✓	✓

Standard errors in parentheses, clustered at firm level
*** p<0.01, ** p<0.05, * p<0.1

As Figure 3 and Table 2 show, the conditional parallel trend assumption is satisfied, since the pre-treatment average estimates of the ATT and the estimates of the ATT in each year before treatment ($t-6$ to $t0$) are statistically equal to zero.

The estimated simple weighted average ATT shows that firms' use of AI in hiring causes, on average, an increase by $\sim$2% in the probability of hiring a female manager.

The estimated results suggest that, even if the existing literature prompts concerns about algorithmic fairness and argues that AI can increase gender discrimination in hiring because human stereotypes may be carried over to AI itself, firms' use of AI can actually increase women's representation in managerial positions. AI can, therefore, help to reduce the persistent under-representation of women in managerial positions and gender inequality in the labor market.

4.3 Potential channels

4.3.1 Type of AI

The type of AI firms use in hiring can best explain my results. One major problem in investigating this potential mechanism is the lack of information on firms' annual integrated reports about what type of AI firms use in hiring. In order to address this problem, I resort to a basic assumption. When firms decide to adopt AI in hiring, they face two macro categories of AI: (i) predictive AI, which predicts the best job candidate to hire based on the characteristics of incumbent employees; or (ii) screening AI, which predicts the best job candidate to hire through screening job applicants' resumes and determining their fit to the job based on the information in the resumes (O' Neil, 2016; DeBrusk, 2018; Daugherty et al., 2019; Cowgill and Tucker, 2020). Now, what do these types of AI mean for hiring? On the one hand, those firms that use screening AI are the ones making higher-quality hires. Thanks to AI's ability to process information at high volume and speed, firms can consider much more relevant information on job candidates' skills, compared to the amount of data human recruiters can handle and analyze (Langenkamp et al., 2019; Black and van Esch, 2020). It follows that firms using screening AI make higher-quality hires than otherwise equal firms, since they can exploit all available information on job candidates' resumes to determine candidates' fit to the job, independently of candidates' gender. On the other hand, those firms that use predictive AI are the ones keeping constant or even worsening the quality of their hires. Since AI bases its decision on the best candidate to hire on the characteristics of the incumbent workforce, AI keeps hiring job candidates who resemble as much as possible the existing workers (O' Neil, 2016; DeBrusk, 2018; Bogen, 2019). Since the focus of this paper is on hiring managers, it follows from the above reasoning that those firms using screening AI for hiring new managers will end up hiring managers who perform better than those hired by firms using predictive AI. As a result, firms using screening AI will show better performance in investment, financial, and organizational practices than firms using predictive AI, because of higher performing managers (Bertrand and Schoar, 2003).

I, therefore, assign firms to screening AI or predictive AI based on their performance along six different measures, collected from the Orbis database by bureau van Dijk, that — as the literature shows (see for instance Bertrand and Schoar (2003)) — are affected by managers' performance: cash flow, cash holdings, value of acquisitions, dividends, solvency ratio, and investments. Firms' performance is an index of how firms perform well along the above six measures after the year in which they adopt AI. Firms are assigned to screening AI if they have adopted AI in hiring and their performance is above the median, they are assigned to predictive AI otherwise. Figure 4 shows the distribution of firms' performance after the adoption of AI across percentiles. Table 3 shows the summary statistics of the variables used for assigning firms to screening and predictive AI for the companies in the restricted matched sample on which I performed the main analyses.

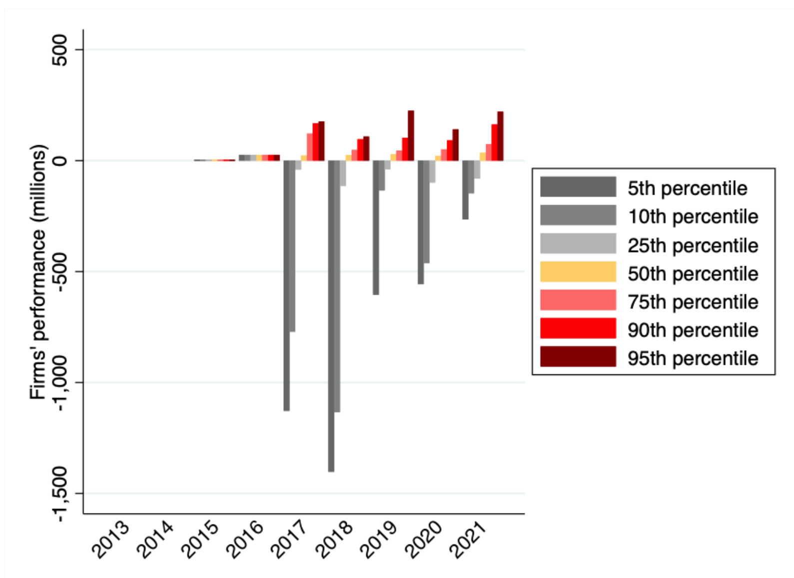
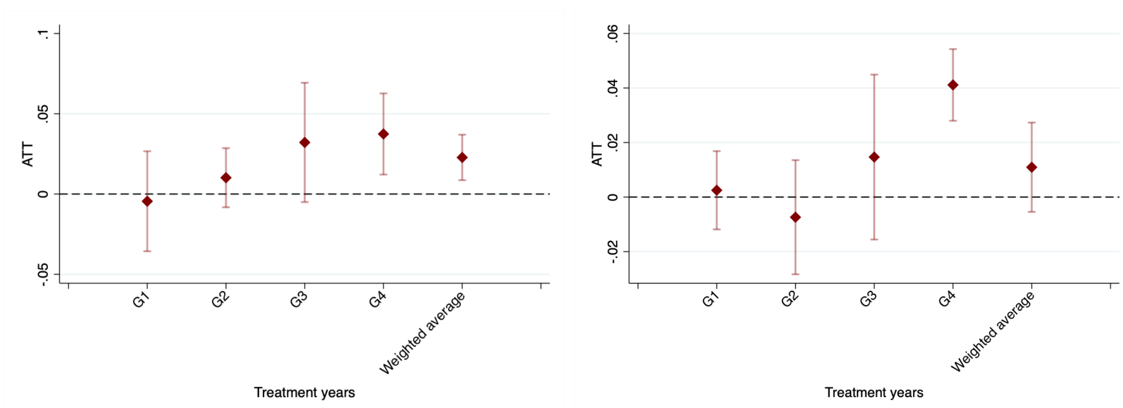


Figure 4: Distribution of firms' performance after the adoption of AI across percentiles.

Table 3: Summary statistics of the variables used for assigning firms to AI types.

	Mean	SD
Firms' performance after the adoption of AI (millions)	-274.644	646.922
Firms' use of screening AI	0.754	0.432
Firms' use of predictive AI	0.246	0.432
Cash flow (millions)	23.105	24.155
Cash holdings (millions)	-7.633	14.801
Acquisitions (millions)	-2.914	10.359
Investments (millions)	1.385	0.631
Solvency ratio	31.441	14.556
Dividends (millions)	-14.999	6.236
N	284	

Figure 5 shows the graphical results of the estimated effect of using each type of AI on firms' probability of hiring female managers, by each year of treatment. Table 4 reports the event study and simple weighted average standardized estimates of the effect of using each type of AI on firms' probability of hiring female managers.



(a) Screening AI

(b) Predictive AI

Figure 5: Staggered difference-in-difference results by type of AI. Group estimates.

Table 4: Staggered difference-in-differences results by type of AI

	Probability of hiring a female manager	
	(1)	(2)
	Simple weighted average	Event study
Firms' use of screening AI	0.019**	
	(0.008)	
Firms' use of predictive AI	0.008	
	(0.009)	
Pre-treatment average (screening AI)		-0.000
		(0.002)
Post-treatment average (screening AI)		0.020*
		(0.011)
Pre-treatment average (predictive AI)		0.015***
		(0.004)
Post-treatment average (predictive AI)		0.012
		(0.008)
Managers' age	✓	✓
Total assets	✓	✓
Number of employees	✓	✓
Productivity	✓	✓
Profit margin	✓	✓
Net income	✓	✓
Country	✓	✓
Industry	✓	✓
Return on equity	✓	✓
Share of female directors	✓	✓
Innovative activity (investments in tangible and intangibles)	✓	✓

Standard errors in parentheses, clustered at firm level
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The increase in the probability of hiring female managers is driven by firms using screening AI and firms' use of predictive AI cannot be claimed statistically different from zero. The event study shows that when firms' use of predictive AI has a positive effect on the probability of hiring female managers, this is because of an already existing positive pre-trend. Firms using predictive AI tend to hire new workers who are as similar as possible to the incumbent ones (O' Neil, 2016; DeBrusk, 2018; Bogen, 2019), therefore, the use of predictive AI does not affect firms' probability of hiring female managers because those firms hiring more women are the ones that already have a more diverse managerial pool.

4.3.2 Reduction of gender discrimination in hiring

The increase in the probability of hiring female managers caused by firms' use of screening AI can be explained by AI reducing gender stereotypes in hiring. In fact, thanks to an accurate evaluation of job applicants' quality, done through the analysis of a rich set of job candidates' work-related information in applicants' resumes, AI can help firms reduce employers' discrimination against women based on status (defined as in Ball et al. (2001)) (Daugherty et al., 2019; Langenkamp et al., 2019; Black and van Esch, 2020).

To test this potential mechanism, I collected using Google API information on the court cases in which lawsuits against firms for gender discrimination in hiring were filed between 2013 and 2021. In order to make sure that the legal system is common to all firms and because of public availability of lawsuits related documentation, I collected information only on firms in the US. Figure 6 and Table 5 show that the effect of using AI on US firms' probability of hiring female managers is very close to the ATT estimated on firms in both Europe and the US. Restricting the sample to US firms does not undermine my main results.

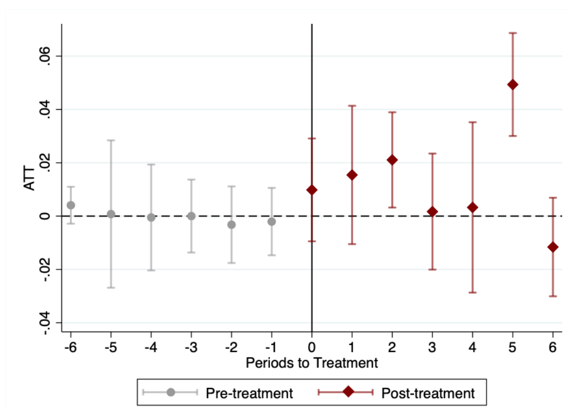


Figure 6: Staggered difference-in-differences results. Event study. US firms

Table 5: Staggered difference-in-differences results. US firms.

	Probability of hiring a female manager	
	(1)	(2)
	Simple weighted average	Event study
Firms' use of AI	0.013**	
	(0.006)	
Pre-treatment average		-0.000
		(0.002)
Post-treatment average		0.013**
		(0.006)
Managers' age	✓	✓
Total assets	✓	✓
Number of employees	✓	✓
Productivity	✓	✓
Profit margin	✓	✓
Net income	✓	✓
Country	✓	✓
Industry	✓	✓
Return on equity	✓	✓
Share of female directors	✓	✓

Standard errors in parentheses, clustered at firm level
*** p<0.01, ** p<0.05, * p<0.1

Figure 7 shows the graphical results of the effect of using AI on firms' probability of being sued for gender discrimination in hiring. Table 6 reports the event study and simple weighted average standardized estimates of the effect of using AI on firms' probability of being sued for gender discrimination in hiring.

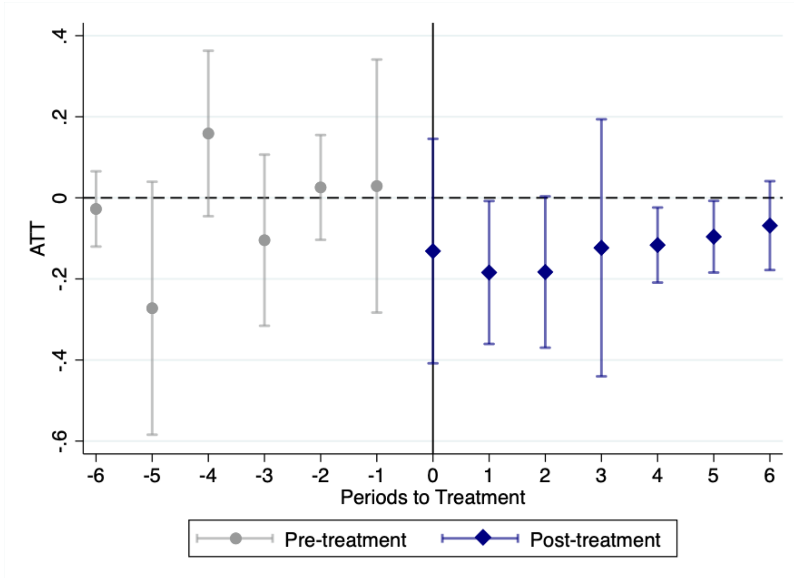


Figure 7: Staggered difference-in-differences results. Event study. US firms' lawsuits

Table 6: Staggered difference-in-differences results. US firms' lawsuits.

	Probability of being sued for gender discrimination in hiring	
	(1)	(2)
	Simple weighted average	Event study
Firms' use of AI	-0.155*	
	(0.089)	
Pre-treatment average		-0.032
		(0.034)
Post-treatment average		-0.129**
		(0.062)

Standard errors in parentheses, clustered at firm level
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Using AI leads to a reduction in the probability of firms being sued for gender discrimination in hiring. This result provides evidence of AI helping firms getting rid of human gender stereotypes in hiring, resulting in an increase in the probability of hiring female managers.

5 Policy implications

The positive association between gender equality and economic growth is well documented in the existing literature. An increase in female representation at firms' top-level positions can promote female participation in the workforce, reduce the gender wage gap and improve fertility and households well-being (Cohen and Huffman, 2007; Hensvik, 2014; Profeta, 2020).

Furthermore, existing scholars have estimated that women's purchasing power is constantly increasing; women cover more than US$20 trillion in worldwide spending, more than 50% of traditional male products, such as automobiles and consumer electronics, and outweigh men in financial plannings (for a discussion on this see for example Profeta (2020)). It follows that if using AI leads to an increase in firms' probability of hiring female managers in both Europe and the US, female participation in the labor market can further increase and women's purchasing power can expand further. Such improvement in female economic empowerment would have fundamental impacts on female social empowerment, allowing, for instance, more and more women to undertake divorce, improving their well-being and reducing domestic violence.

6 Conclusion

This study used a staggered difference-in-differences approach to test whether firms' use of AI may affect their probability of hiring female managers. An increase in firms' probability of hiring female managers can be attributed to the use of screening AI. Because of AI's ability to process large amounts of information on job applicants' skills without correlating job-relevant outcomes to gender, the decrease in gender discrimination in hiring is striking. The experimental design allows me to estimate the effect of firms' use of AI on their probability of hiring female managers and answer to the open question on whether and how firms' adoption of AI affects key employment outcomes. This paper adds to the lack of micro-level evidence on the impact of firms' use of AI on the persistent female under-representation in managerial positions. It is also the first quasi-experimental evaluation of AI on key employment outcomes.

References

Akerlof, G.A. (1997) Social Distance and Social Decisions. *Econometrica.* 65(5): 1005–1027.

Albert, E.T. (2019) AI in talent acquisition: a review of AI applications used in recruitment and selection. *Strategic HR Review.* 18(5): 215–221.

Arrow, K.J. (1973) The Theory of Discrimination. O. Ashenfelter, A. Rees, eds. *Discrimination in Labor Markets* (Princeton University Press, Princeton), 3–33.

Ball, S., C. Eckel, P. J. Grossman, W. Zame (2001) Status in markets. *The Quarterly Journal of Economics*, 116(1): 161–188.

Berger, J., D. G. Wagner (1997) Gender and Interpersonal Task Behaviors: Status Expectations Accounts. *Sociological Perspectives.* 40(1): 1–32.

Bertrand, M., A. Schoar (2003) Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics.* 118(4): 1169–1208.

Bildfell, C. (2019) Hiring Algorithms in the Canadian Private Sector: Examining the Promise of Greater Workplace Equality. *Canadian Journal of Law and Technology.* 17(2): 115.

Black, J. S., P. van Esch (2020). AI-enabled recruiting: What is it and how should a manager use it?. *Business Horizons.* 63(2): 215 – 226.

Bogen, M. (2019). All the ways hiring algorithms can introduce bias. *Harvard Business Review.* 6, 2019.

Brathwaite, T., A. Vij, J. L. Walker. (2017) Machine learning meets microeconomics: The case of decision trees and discrete choice. *arXiv preprint.* 1711(04826).

Breines, W. (2006) *The trouble between us: An uneasy history of white and black women in the feminist movement* (Oxford University Press, New York).

Brescoll, V. L. (2016) Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders. *The Leadership Quarterly.* 27(3): 415–428.

Brescoll, V. L., T. G. Okimoto, A. C. Vial (2018) You've come a long way. . . maybe: How moral emotions trigger backlash against women leaders. *Journal of Social Issues.* 74(1): 144–164.

Callaway, B., P. H. Sant'Anna (2021) Difference-in-Differences with multiple time periods. *Journal of Econometrics.* 225(2): 200–230.

Castilla, E. J. (2008) Gender, race, and meritocracy in firmal careers. *American Journal of Sociology.* 113:1479–1526.

Castilla, E. J. (2015) Accounting for the Gap: A Firm Study Manipulating firmal Accountability and Transparency in Pay Decisions. *firm Science.* 26(2):311-333.

Cerulli, G. (2019) Data-driven sensitivity analysis for matching estimators. *Economics Letters.* 185(2019): 108749.

Cohen, P. N., M. L. Huffman (2007) Working for the woman? Female managers and the gender wage gap. *American sociological review.* 72(5): 681–704.

Correll, S. J., S. Benard (2006) Biased estimators? Comparing status and statistical theories of gender discrimination. S.R. Thye, E.J. Lawler, eds. *Advances in group processes* (Emerald Group Publishing Limited), 89–116.

Correll, S. J., S. Benard, I. Paik (2007) Getting a Job: Is There a Motherhood Penalty?. *American Journal of Sociology.* 112(5): 1297–1338.

Cowgill, B. (2019) Bias and productivity in humans and machines. Working paper, Columbia Business School.

Cowgill, B., C. E. Tucker (2020) Algorithmic fairness and economics. Working paper, Columbia Business School.

Crittenden, A. (2001) *The Price of Motherhood: Why the Most Important Job in the World Is Still the Least Valued* (Metropolitan Books, New York).

Cuddy, A., S. T. Fiske, P. Glick (2004) When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues.* 60(4): 701–718.

Cunningham, J., T. Macan (2007) Effects of applicant pregnancy on hiring decisions and interview ratings. *Sex Roles.* 57(7): 497–508.

Damanpour, F. (1992) firmal size and innovation. *firm studies*, 13(3): 375–402.

Datta, A., M. C. Tschantz, A. Datta (2015) Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies.* 2015(1): 92–112.

Daugherty, P. R., H. J. Wilson, R. Chowdhury (2019) Using artificial intelligence to promote diversity. *MIT Sloan Management Review.* 60(2).

DeBrusk, C. (2018) The risk of machine-learning bias (and how to prevent it). *MIT Sloan*

Management Review (online). https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it/

Eagly, A. H. (1997) Sex differences in social behavior: Comparing social role theory and evolutionary psychology. *American Psychologist.* 52(12): 1380–1383.

Fernandez, R. M., M. L. Mors (2008) Competing for jobs: Labor queues and gender sorting in the hiring process. *Social Science Research.* 37(4): 1061–1080.

Frankish, K., W. M. Ramsey (2014) *The Cambridge Handbook of Artificial Intelligence* (Cambridge University Press, Cambridge).

Gebru, T. (2020) Race and gender. M. D. Dubber, F. Pasquale, S. Das, eds. *The Oxford handbook of ethics of AI.* (Oxford University Press, New York), 251–269.

Gloor, J. L., M. Morf, S. Paustian-Underdahl, U. Backes-Gellner (2020) Fix the Game, Not the Dame: Restoring Equity in Leadership Evaluations. *J*ournal of Business Ethics. 161(3): 497–511.

Goldin, C., C. Rouse (2000) Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American economic review.* 90(4): 715–741.

Hensvik, L. E. (2014) Manager Impartiality: Worker-Firm Matching and the Gender Wage Gap. *ILR Review.* 67(2):395–421.

Hewlett, S.A., K. Peraino, L. Sherbin, K. Sumberg (2010) The sponsor effect: Breaking through the last glass ceiling. *Harvard Business Review.*

Kahneman, D., S. P. Slovic, A. Tversky (1982) *Judgment under uncertainty: heuristics and biases* (Cambridge University Press, Cambridge).

Lambrecht, A., C. E. Tucker (2019) Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science.* 65(7): 2966–2981.

Langenkamp, M., A. Costa, C. Cheung (2019) Hiring Fairly in the Age of Algorithms. Working paper. https://ssrn.com/abstract=3723046 or http://dx.doi.org/10.2139/ssrn.3723046.

Newman, D. T., N. J. Fast, D. J. Harmon (2020) When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *firmal Behavior and Human Decision Processes.* 160: 149–167.

O' Neil, C. (2016) *Weapons of math destruction: How big data increases inequality and threatens democracy* (Penguin Random House LLC, New York).

Parisi, M. L., F. Schiantarelli, A. Sembenelli (2006) Productivity, innovation and RD: Micro evidence for Italy. *European Economic Review*, 50(8): 2037–2061.

Phelps, E. S. (1997) The Statistical Theory of Racism and Sexism. *The American Economic Review.* 62(4): 659–661.

Profeta, P. (2020) *Gender equality and public policy: Measuring progress in Europe* (Cambridge University Press, Cambridge).

Reskin, B. F. (2001) Employment discrimination and its remedies. Berg, I., A. L. Kalleberg, eds. *Sourcebook of Labor Markets* (Springer, Boston), 567–599.

Reskin, B. F., D. Bielby (2005) A sociological perspective on gender and career outcomes. *Journal of Economic Perspectives.* 19(1): 71–86.

Ridgeway, C. L. (1997) Interaction and the conservation of gender inequality: Considering employment. *American Sociological Review.* 62(2): 218–235.

Ridgeway, C. L. (2011) *Framed by Gender: How Gender Inequality Persists in the Modern World* (Oxford University Press, New York).

Ridgeway, C. L., C. Bourg (2004) Gender as Status: An Expectation States Theory Approach. A. H. Eagly, A. E. Beall, R. J. Sternberg, eds. *The psychology of gender* (The Guildford Press, New York), 217–241.

Ridgeway, C. L., S. J. Correll (2004) Motherhood as a status characteristic. *Journal of Social issues.* 60(4): 683–700.

Robert, L. P., C. Pierce, L. Marquis, S. Kim, R. Alahmad (2020) Designing fair AI for managing employees in firms: a review, critique, and design agenda. *Human-Computer Interaction.* 35(5-6): 545-575.

Rosenbaum, P. R. (2002) Sensitivity to Hidden Bias. P.R. Rosenbaum, eds. *Observational Studies* (Springer, New York), 71-104.

Rosenbaum, P. R., D. B. Rubin (1985) Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician.* 39(1): 33-38.

Roth, L. M. (2007) Women on Wall Street: Despite Diversity Measures, Wall Streets Remains Vulnerable to Sex Discrimination Charges. *Academy of Management Perspectives.* 21(1): 24–35.

Silberg, J., J. Manyika (2019) Notes from the AI frontier: Tackling bias in AI (and in humans). *McKinsey Global Institute.* 1(6).

Stone, P. (2007) *Opting Out? Why Women Really Quit Careers and Head Home* (University of California Press, Berkeley).

vanEsch, P., S. J. Black, J. Ferolie (2019) Marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior.* 2019(90): 215–222.

Vella, F. (1994) Gender Roles and Human Capital Investment: The Relationship between Traditional Attitudes and Female Labor Market Performance. *Economica.* 61(242): 191–211.

Appendix

Distribution of the observable covariates