Do robots hurt humans?

Evidence from the dark side of workplace automation

Kuntara Pukthuanthong, Zhihua Wei, Aoran Zhang,* and *Yunfei Zhao*

Abstract

We investigate the influence of utilizing automated capital to replace human workers in U.S. public firms and find that firms that are more capable of replacing human capital with automated capital need better records on employee-related corporate social responsibility (employee CSR) and workplace safety. Our results remain robust after addressing the endogeneity concern. We further document that the negative link is stronger for firms adopting more aggressive financial policies, incurring higher labor-related costs, and locating in less religious areas. Our analysis also unveils an ethical conundrum. The compromise on employees' well-being among firms that are highly subjective to workplace automation comes at the cost of reputation damage; however, such firms are associated with better financial performance. Overall, we document that workplace automation creates a moral hazard: firms expect to increase productivity with automated capital while paying less attention to current employees' welfare.

Keywords: Labor and finance; Workplace automation; Workplace safety; Corporate social responsibility; Artificial intelligence

JEL Classification: G31, G34, J23, J28, O33

Pukthuanthong is from University of Missouri, Columbia, MO 65211, Wei is from the School of Economics, Xiamen University, Building of Economics, Xiamen, Fujian, China, 361005, Zhang is from Department of Finance, Nottingham University Business School, University of Nottingham, Jubilee Campus, Nottingham, United Kingdom, NG8 1BB, Zhao is from Finance Department, John Molson School of Business, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada, H3G 1M8 Email addresses: pukthuanthongk@missouri.edu (pukthuanthong). finjoy@126.com (Z. Wei), aoran.zhang@nottingham.ac.uk (A. Zhang), yunfei.zhao@mail.concordia.ca (Y. Zhao) *Corresponding Author

1. Introduction

"*Tesla's problem: overestimating automation, underestimating humans.*"

IMD Professor Bettina Büchel, with Dario Floreano

In the recent decade, we have been living in an accelerated technological revolution. The rapid application and adaptation of cutting-edge technologies have profoundly changed our daily life and well-being. Artificial intelligence (AI), big data, and machine learning are typical examples of applying breakthrough technologies to human society. Nowadays, AI and machine learning have become more and more ubiquitous and have been permeating a variety of business activities and shaping up novel production processes. One pivotal impact of such advanced technologies can be found in the change in the workplace of human beings. The utilization of automated capital has already exhibited advantages in replacing human laborers to perform routine tasks (Zhang, 2019). However, with the development of cutting-edge technologies such as AI and machine learning, a host of non-routine jobs that are supposed to be performed by skilled workers can now be allocated to robots, such as customer service, financial and real estate agencies, radiology, etc.

According to the annual report from the International Federation of Robots (IFR), despite the COVID-19 pandemic, the installations of industrial and service robots keep increasing. In 2020, approximately 400,000 robots were installed in production processes and the provision of services worldwide. Also, the operational stock of industrial and service robots had increased by 10% in 2020 to a level of more than 3 million robots 1. Interestingly, robots may even be able to take over skilled tasks such as financial reporting and auditing in accounting firms (Forbes, 2021).

Replacing human workers with robots has tremendous financial benefits. Given that AI can be programmed and replace human labor, it can significantly benefit corporations by reducing fixed costs and, in turn, maximizing profits and firm value (*Fortune*, 2018). Moreover, according to *Fortune* (2018), the preponderance and reliance on AI will remain the same but will further accelerate. PwC has estimated that the overall economic value created by AI-related workplace replacement is substantial, with approximately $16 trillion to worldwide GDP by 2030.

Given the conspicuous economic benefits of utilizing AI-related automation, an emerging string of research on labor economics has highlighted the positive financial and policy implications of AI capital on employment as well as wages (see, e.g., Autor and Dorn, 2013; Goos, Manning, and Salomons, 2014; Frey and Osborne, 2017; Acemoglu and Restrepo, 2018 and 2019; Autor and Salomons, 2018; Graetz and Michaels, 2018; and Webb, 2020). Even so, studies have yet to explore the utilization of AI in the workplace from the perspective of corporate finance. The few notable exceptions are the research of Zhang (2019) and Bates, Du, and Wang (2020). Zhang (2019) focuses on routine-task laborers that are replaced by automated capital in the workplace, and the theoretical framework and accompanying empirical analysis show that firms with the option to replace routine-task workers with automated means are more capable of hedging against macro-level economic turmoil and in turn tend to have lower systematic risk. Additionally, Bates, Du, and Wang (2020) extend the horizon of workplace replacement by including non-routine tasks replace by AI-related automation.

With the development of advanced technologies such as AI, big data, and machine learning, replacing the jobs that are supposed to be done by skilled workers with automation has become feasible. Bates, Du, and Wang (2020) shed new light on how workplace replacement by automation can influence corporate financial policies in U.S. public firms. Their empirical evidence reveals that firms that can substitute routine and non-routine jobs tend to have more aggressive monetary policies. Specifically, such firms hold less precautionary cash, pay more dividends, and have higher financial leverage. They further explore the mechanism via which automation replacement affects corporate economic policies and document that the capability of substituting labor by automated capital can effectively reduce firms' operating leverage induced by fixed labor costs and increase firms' operating flexibility. The enhanced functional flexibility enables such firms to adopt more aggressive financial policies.

The abovementioned two studies involving workplace automation successfully manifest the benefits of advanced technologies on corporate finance. However, it is also necessary to recognize all possible shortcomings of adopting new technologies, especially within the corporate bound. Surprisingly, to our best knowledge, no study has pinned down the dark side of AI utilization from the corporate perspective. Nevertheless, anecdotal evidence has started unveiling some drawbacks associated with automated capital. Elon Musk, the CEO of Tesla (the famous electric vehicle company), considers Tesla's AI team one of the best in the world. Indeed, Tesla has extensively employed AI automation in manufacturing and has even applied AI automation in non-consumer products, such as in its energy software *Autobidder* (*Electrek,* 2021). However, Elon Musk has recognized the cost of AI, claiming that "*excessive automation was a mistake*" and that "*humans are underrated.*" The mistake came with the price that Tesla failed to keep the promise of 5,000 new electric cars per week. Conversely, the automated factory only produced less than half of the promised quantity (Büchel and Floreano, 2018).

The above fact suggests that overestimating AI automation can lead to overlooking the essence of human capital. Moreover, the overlook of human capital can trigger another cost. That is, employers relying on automated capital tend to downplay the well-being of their employees. Even though Tesla's factory is well known for the extensive use of AI, the factory has been grappled with a more dubious allegation to fame: problematic workplace safety. So far, Telsa has been the target of more investigations into workplace safety and has been fined more than its industry peers over the past half-decade. Specifically, the annual rate of serious injury (requiring an employee to take time off to recover) is higher than the auto-industry average (*Forbes*, 2019). The article in *Forbes* further states the reason for Tesla's poor workplace safety record. Unfortunately, compared to its peers, Tesla offered insufficient training for its employees to operate heavy machines to prevent injuries in an assembly line. This is because of Elon Musk's early vision of over-focusing on robots instead of human workers to manufacture electric vehicles.

It is worth noting that the economic consequences of productivity losses from poor practice on employee well-being are tremendous. According to Leigh (2011), in the U.S., injuries from the workplace are more than 3.5 million, with an enormous financial cost of $250 billion approximately. This considerable cost is even higher than all kinds of cancer combined. If workplace automation is associated with such a vast financial cost, it is worth bringing the attention of academia and practitioners. In this study, we conjecture that a moral hazard problem may exist in firms that are highly capable of substituting their workers with automated capital. As such firms are aware that they can eventually get rid of human labor, they may believe that it is unnecessary to devote resources to the well-being of their employees. Consequently, we hypothesize that firms with the option to replace human workers with AI tend to have lower performance in terms of employee-related corporate social responsibilities (employee CSR) and poorer records of workplace safety (an essential component of employee CSR) than their counterparts.

Based on a large sample of U.S. public firms between 1999 and 2016 and using an upgraded firm-level measure of the capability of replacing human workers with automated capital (*AI replacement* in short) based on the proxy at the industry-level adopted by Bates, Du, and Wang (2020), our empirical analysis suggests that firms with a high level of *AI replacement* are significantly associated with lower employee CSR performance as well as more workplace injuries and casualties. Our results are economically significant; for instance, a one-standard-deviation increase in *AI replacement* is associated with a 0.015 decrease in employee CSR performance. However, we find that the negative impact of *AI replacement* on CSR is only limited to employee-related CSR, and no statistical evidence suggests that AI replacement can influence CSR in other categories.

Our results remain robust after addressing the potential endogeneity concern. Specifically, we adopt many empirical strategies to mitigate the endogeneity concern. Firstly, we assume an instrumental variable (IV). Our IV is the level of agricultural mechanization in 1940 (*LAM 1940*) in a given U.S. state (the aggregate value of agricultural machinery ($ thousands) scaled by the population in 1940 in a given state), which captures a U.S. state's adaptivity to new technologies. The results from the IV analyses confirm our baseline findings. To further reinforce the causal interpretation of our baseline results, we design a quasi-natural experiment using the 2007 America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (*America COMPETES Act*) as a plausibly exogenous shock. The *America COMPETES Act* has been viewed as potentially the most notable U.S. government policy initiative on natural science and innovation of the new millennium (Furman, 2013; Gonzalez, 2015). This act has had a monumental impact on U.S. firms' innovation commitment through research and development, including the development of artificial intelligence. Based on this act, we perform the difference-in-differences estimation based on the propensity score matching routine (PSM-DID). The results from the PSM-DID analysis indicate that the negative link between *AI replacement* and poor employee CSR performance becomes even more vital. We interpret the results as further support of our baseline results.

Moreover, we perform additional analysis to discuss the transmission mechanisms via which *AI replacement* discourages workplace safety and overall employee CSR. Firstly, we decompose the employee CSR score into categories and find significant results in six categories. Specifically, firms with a high likelihood of replacing human workers with automated capital are associated with a significantly lower commitment to employee involvement and human capital development. However, such firms, on the other hand, have more concerns about the relationship with unions, health and safety, and workers' retirement benefits.

Secondly, union coverage and membership can effectively protect workers' rights in their workplace. Union inclusion can significantly increase firms' labor-induced operating leverage. Consequently, we find that union coverage and membership increase firms' substitution by using robotic capital. Moreover, we find that union inclusion can effectively attenuate the negative impact caused by *AI replacement* on employees' well-being.

Thirdly, according to Bates, Du, and Wang (2020), the positive link between aggressive financial policies and AI replacement drives the relaxed labor-induced operating leverage. Therefore, firms with higher labor-induced operating power are eager to eliminate human workers to increase their operating flexibility and focus more on acquiring AI capital while paying less attention to employee CSR. Indeed, our empirical evidence reveals that the negative link between *AI replacement* and employee CSR only exists among firms whose labor-induced operating leverage (measured by selling, general, and administrative expenses) is above the median value.

Fourthly, Bates, Du, and Wang (2020) indicate that firms adopt less conservative financial policies by reducing their precautionary cash holdings when they can substitute their labor with automated capital. In line with this rationale, we further find that employee well-being in firms with less precautionary cash holdings is even worse when they can utilize AI to replace their human workers. This channel is also consistent with Cohn and Wardlaw (2016) that financial constraints can hinder workplace safety.

Other than the channels from the firm level, we also document some tracks from the macro level (state and county). We find that state-level labor costs and difficulty of hiring can significantly affect the link between *AI replacement* and *Employee CSR*. Specifically, when the hiring costs are higher (higher mandatory minimum wage and bearish job market with a higher unemployment rate), firms are more willing to switch to automated capital for their production line to reduce the operating leverage stemming from human labor and, in turn, focus less on building employees' well-being. Additionally, we find that religious belief may have an ethical effect and can attenuate the negative impact of workplace automation on employees' well-being. By proxying the spiritual development by the percentage of Christians in a county's population, our analysis suggests that the negative link between workplace automation and employee CSR becomes weaker in more religious counties, implying that firms with a high likelihood of workplace automation are less likely to overlook their employees' well-being in the more religious county, in comparison to the less religious county.

Furthermore, we investigate whether firms capable of replacing human workers with automated capital while downplaying their employees' well-being can have negative reputational and financial implications. Indeed, in line with our expectations, such firms can incur significant reputation damage. Specifically, firms more susceptible to workplace automation are less likely to be recognized as the best employers in the U.S., according to Fortune's ranking of the top 100 employers. A one-standard-deviation increase in AI replacement reduces the likelihood of being ranked as one of the best employers by 11.5%. Surprisingly, such firms can achieve better firm value and accounting performance when they downplay employees' well-being. This evidence indicates that firms' unethical behavior is rewarded by higher firm performance, incurring a moral hazard problem.

Our paper contributes to the literature in three folds. First, our empirical evidence further enriches an emerging strand of studies on employees' well-being. Cohn and Wardlaw (2016), Cohn, Nestoriak, and Wardlaw (2017), Caskey and Ozel (2017), and Bradley, Mao, and Zhang (2021) indicate that corporate financial constraints, private equity ownership, and financial have a significant impact on corporate workplace safety. Our study extends the research by looking into a broader sense of the employee-related issue, the general employee-related corporate social responsibility practice, rather than only workplace safety. In particular, our study contributes to the available literature about the factors influencing CSR. To the best of our knowledge, our paper is the first attempt to investigate how the likelihood of workplace automation affects CSR, especially employee CSR. The previous literature indicates that a company's CSR can be influenced by family ownership (Block and Wagner, 2014; Oh et al., 2019), financial analysts' coverage (Adhikari, 2016), female directors (McGuinness et al., 2017), institutional ownership (Dyck et al. 2019; Li et al., 2021b), management skills of CEO (Chen et al. 2020a), cross-listing (Lu and Wang 2021), and so on. While our study focuses on employee CSR and how the likelihood of workplace automation can influence it.

More importantly, our research directly contributes to a nascent but growing string of research on the financial implications of workplace automation (see, e.g., Zhang, 2019; Bates, Du, and Wang, 2020; and Cheng, Lyandres, Zhou, and Zhou, 2021). As discussed above, existing studies by Zhang (2019) and, Bates, Du, and Wang (2020), Cheng et al. (2021) document the benefit of adopting industrial robots on corporate financing as increased optimal financial leverage and lower cost of borrowing. Nonetheless, in our research, we outline the dark side associated with workplace automation instead of further praising the utilization of robotic capital for replacing human capital. We argue that our findings have crucial policy implications on public firms' accounting and financial policies and corporate governance. Human intervention is indispensable to realize the benefits of robotic automation fully. Corporate executives are required to attach importance to this issue and to foster a healthy environment for the welfare of corporate employees. Otherwise, downplaying employees' welfare can trigger significant reputation damage (lower chance of being named as the best employer), financial cost (lower accounting performance and firm value), and social cost (poor record of workplace safety), which can offset the economic benefits of using automated capital a firm initially intends to. In this way, corporations can further maximize productivity and achieve sustainability through a more effective and efficient connection between artificial intelligence and human.

AI replacement is measured as the portion of a company's existing human workers susceptible to being taken over by robotic capital. Initially, Frey and Osborne (2017) and Bates, Du, and Wang (2020) computed it as the weighted average susceptibility to AI replacement across all occupations in a given industry to which a firm belongs. However, we upgrade the measure as the weighted average value across all business segments a firm operates in. Furthermore, we improve the proxy of AI replacement based on the one used by Frey and Osborne (2017) and Bates, Du, and Wang (2020) from the industry to the firm level. We believe the action at the firm level can better reinforce the causal interpretation.

The remainder of the paper is structured as follows. Section 2 introduces our data, sample, variable construction, and descriptive statistics. Section 3 shows the main empirical results. Section 4 discusses our strategies to mitigate endogeneity concerns. Section 5 investigates the transmission channels. Section 6 examines the actual costs of poor employee CSR practice. Section 7 concludes.

2. Hypothesis development

The conflict between shareholders and stakeholders is recognized as one of the most contested topics in corporate governance (Adams, Licht, and Sagiv, 2011). Based on the traditional view, as a modern corporation, the main goal is to maximize the shareholders' wealth. Any benefits to corporate stakeholders (such as corporate employees) at the expense of maximizing shareholders' wealth are considered counterproductive and, therefore, should not be encouraged (see, e.g., Tirole, 2001; Gelter, 2009; Ferrell et al., 2016; Leung et al., 2019). Indeed, existing literature has documented the detrimental effect of corporate social responsibility commitments that specifically benefit stakeholders. Servaes and Tamayo (2013), Di Giuli and Kostovetsky (2014), Masulis and Reza (2015), and Buchanan et al. (2018) empirically examine the influence of CSR on firm performance and find that more CSR activities can significantly impede firm performance, such as lower profitability (measured by ROA), lower stock performance as well as lower firm value (proxied by Tobin's q). They explain their finding that any benefits to corporate stakeholders from social responsibility can come at the direct cost of firm performance.

In terms of corporate employees, they are recognized as one of the major stakeholders of modern corporations. Thus, employee-related CSR is a pivotal component of corporate CSR. In line with the dark side of CSR, prior studies also reveal that protecting the well-being of corporate employees is costly. Simintzi, Vig, and Volpin (2015) indicate that enhanced protection of corporate employees comes at the expense of increased corporate restructuring costs and operating leverage. The empirical evidence from Serfling (2016) suggests that augmented labor rigidity is linked to increased earnings volatility. Favilukis, Lin, and Zhao (2020) illustrate that the rigidity of employees' wages is associated with not only higher operating leverage but also triggers higher credit risk. Like Simintzi, Vig, and Volpin (2015), Schmalz (2018) examines the effect of unionization and uncovers the negative impact of increased employee protection on operating leverage.

Based on the abovementioned shortcomings of commitments that specifically benefit corporate stakeholders (employees), firms may be incentivized to reduce their reliance on human workers. Suppose advanced technologies, such as AI-related workforce automation, allow firms to replace humans in routine and cognitive jobs. In that case, such firms may not be willing to invest in employee-related benefits. Thus, we postulate that, all else equal, firms with a higher portion of employees that are more susceptible to being replaced with automated capital are associated with lower employee-related CSR performance than their counterparts.

However, corporate stakeholders' well-being has been significantly highlighted in recent years. Specifically, in 2019, the Business Roundtable, consisting of many corporate executives of major U.S. firms, released a statement about a modern corporation's goal. The report advocates that CEOs in modern corporations should benefit all stakeholders (such as customers, employees, suppliers, and communities) rather than merely shareholders. Moreover, the statement was endorsed by 222 influential CEOs (Harrison et al., 2020)

Indeed, the benefits of accommodating the broad interests of stakeholders rather than merely shareholder-oriented have been underpinned by prior studies. Firms can strategically engage in CSR activities to reconcile the agency conflicts between stakeholders and shareholders (see, e.g., Porter & Kramer, 2006 and Benabou and Tirole, 2010). Through CSR commitments, firms can build superior corporate reputation and, in turn, leverage to augment financial performance (Van der Laan et al., 2008; Rettab et al., 2009; El Ghoul et al., 2011, 2017; Servaes and Tamayo, 2013; Ferrell et al., 2015; De Roeck et al., 2016; Leiva et al., 2016 Albuquerque et al., 2019; Cremers et al., 2019), achieve special announcement returns when making mergers and acquisitions (Deng et al., 2013), as well as boost corporate innovation (Flammer and Kacperczyk, 2016).

In terms of employee-oriented CSR, several studies indicate that firms can obtain significant benefits when their employees are better treated. Flammer and Kacperczyk (2016) provide empirical evidence that better employee-oriented CSR policies are associated with higher innovation performance. Such firms have more successfully granted patents and patent citations. They explain that better employee welfare fosters corporate innovation because their employees feel that they work in a secure and delighted working environment and, therefore, can be more productive on innovation commitment. Flammer and Kacperczyk (2019) show that better employee CSR practices can significantly prevent knowledge spillovers. Specifically, employees are less likely to join their rival firms. Even if they enter, they are less likely to disclose valuable knowledge to rival firms. Chunyu, Volpin, and Zhu (2022) illustrate that U.S. firms headquartered in the state mandating the paid sick policy tend to obtain higher labor productivity and firm profitability.

As AI technology is the outcome of technological innovation, firms can utilize AI-related technology as themselves innovation-oriented. Specifically, Babina et al. (2022) reveal that firms adopting AI technology are associated with high-quality human capital, and therefore such firms can achieve higher productivity and more product innovation. Specifically, among firms adopting AI technology, a significant portion of their workforce are well educated and trained and have more marketable skills. Thus, even if such firms can replace a substantial amount of their employees with automated capital, the remaining workforce should be fared very well to guarantee the success of such firms. Hence, it is unclear about the net effect of employee-oriented CSR, even if such firms choose to reduce their commitments to employees' well-being for potentially replaceable workers.

More importantly, our proposed proxy of AI replacement is a look-forward measure, which is like a real option capturing the potential for firms to replace their human workers with AI-related automation. As noted by Bates, Du, and Wang (2022), the initial investment outlay of substituting human workers constitutes a salient financial cost for firms. Achieving the potential level of workforce automation also requires financial capacity. Therefore, it is not the case that firms achieve workforce automation immediately when AI technology becomes available. Conversely, firms would balance the costs and benefits and exercise the option to replace human workers when optimal. Thus, given the significant benefits of better employee treatment specified by existing literature, such firms do not need to start downplaying their employees' welfare at a time when they have the potential to replace their employees.

Based on our above arguments, how a firm's capability of replacing human workers with automated capital would impact their employee-oriented CSR performance remains an empirical question.

3. Data and sample

In this section, we introduce the procedure of data collection and sample construction explicitly, followed by the summary statistics. We retrieve the data from a battery of databases. We exclude financials (SIC code between 6000 and 6999) and utility firms (between 4900 and 4999) from our sample. Our baseline sample consists of 3,746 publicly traded U.S. firms between 1999 and 2016. The period from 1999 to 2016 is selected because *AI replacement* data is from 1999, and the data on CSR is available only until 2016.

3.1. Data on CSR

We collect U.S. public firms' CSR information from the KLD database from 1999 to 2016, which is widely used among CSR-related studies (see, e.g., Jiao, 2010; El Ghoul et al., 2011; Attig et al., 2013; Deng et al., 2013; Jiraporn et al., 2014; Cahan et al., 2015; Lee, 2017). The KLD database identifies public firms' CSR based on various general information, such as financial statements, surveys, media reports, etc. KLD evaluates each firm's CSR performance annually and identifies the strengths and concerns in seven qualitative dimensions: community, corporate governance, diversity, employee, environment, human rights, and product quality and safety. Moreover, KLD provides concern ratings for six controversial business issues, which are alcohol, gambling, firearms, military, nuclear power, and tobacco. To calculate a firm's CSR performance, we exclude the six concern ratings as we want to only focus on corporate managers' discretionary choices on CSR, not firms' involvement in particular industries (Kim et al., 2014). Within a specific qualitative dimension, KLD contains a set of indicators for each strength and concern activity. The indicator is scored one if a firm meets the required assessment criteria. Otherwise, it is scored zero. The qualitative dimension score equals the total strength score minus the total concern score.

Furthermore, we exclude corporate governance rating from our CSR performance calculation, given that our proxy of CSR performance concentrates on the overall benefits of corporate stakeholders rather than merely on shareholders (Dhaliwal et al., 2011). Accordingly, a firm's CSR performance equals the aggregate value of the scores in the six qualitative dimensions. Nevertheless, as the raw values of CSR performance assign equal weight to individual factors and the number of individual factors varies, evaluating firms' CSR performance on fundamental values can be biased (El Ghoul et al., 2011; Deng et al., 2013). To cope with this issue, we construct adjusted CSR performance based on the raw CSR performance. Following the methodology of Deng et al. (2013), we divide the strength (concern) scores by the number of strength (concern) indicators for each dimension in each year to calculate the adjusted power (concern) scores. To obtain the aggregate strength (concern) scores, we sum up the adjusted strength (problem) scores across the six dimensions. Finally, a firm's adjusted CSR performance is measured as the difference between the adjusted total strength score and the adjusted total concern score. Our specific focus is the adjusted CSR performance on employees (*Employee CSR*). Another *CSR* is computed as the sum of the scores in the other five dimensions.

3.2. AI replacement measurement

We construct our key variable, *AI replacement*, based on the one used by Frey and Osborne (2017) and Bates, Du, and Wang (2020), which equals the percentage of a firm's current human labor force susceptible to automated capital replacement. Specifically, the proxy is created as the weighted average capability of a firm to substitute human capital with automated capital across all occupations in various industries the company operates in. Our proxy can be viewed as a time-varying and firm-specific measure for replacing human labor and robotic capital that dynamically constitutes a firm's production factor combination and corporate financial policies. It is also worth noting that we upgraded the proxy used by Frey and Osborne (2017) and Bates, Du, and Wang (2020) from an industry-specific measure to a firm-specific measure. The detail of the upgrade is described below.

The proxy of *AI replacement* is constructed using the data from two sources. The primary source is the one from Frey and Osborne (2017). Based on the most recent development in machine learning and industrial robotics, Frey and Osborne (2017) create a novel proxy that characterizes occupations by their susceptibility to AI-related automation. Specifically, they estimate the probabilities of *AI replacement* for 702 detailed works defined by the 2010 Standard Occupation Classification (SOC) code. A higher estimated likelihood implies that the job is very likely to be replaced by robots.

In comparison, a lower estimated likelihood indicates that the job is improbable to be taken over by robots. Therefore, the estimated probabilities can range from zero to one. The advantage of the proxy of Frey and Osborne is that it quantifies the updated influence of cutting-edge technologies, such as artificial intelligence, big data, and machine learning, on workplace evolution. Notably, the novelty of Frey and Osborne (2017)'s proxy includes not only the workplace replacement for both cognitive and manual routine tasks by automated capital but also focuses on the workplace substitution for non-routine tasks requiring skilled workers by AI-related automation, which has yet been adequately explored by existing literature on workplace automation. For instance, driving, medical diagnosis, financial and law services, etc. The second database used is the Occupational Employment Statistics (OES) program maintained by the Bureau of Labor. OES program contains the industry-level occupational employment and salary estimates annually from 1999 to the present. The OES program used 2000 SOC definitions from 1999 to 2009, while updates to the 2010 SOC definitions after 2009. To ensure the consistency of SOC definitions across our sample period, we use a crosswalk table available at the National Crosswalk Service Center to link the 2000 SOC codes to the 2010 ones. Industries are classified by the 3-digit Standard Industrial Classification (SIC) codes until 2001 and by 4-digit North American Industry Classification System (NAICS) codes from 2002. According to the 3-digit SIC classification, there are 376 unique industries from 1999 to 2001, while there are 290 unique 4-digit NAICS industries for 2002 and after. Finally, we construct the proxy of a firm's capability of replacing human labor with robotic capital (*AI replacement*) by incorporating the data obtained from the two databases. Specifically, the proxy at the industry level each year is shown as follows:

$$SLAC_{j,t} = \sum_{k=0}^{n} \mathrm{Prob_k} \times \frac{Emp_{j,k,t} \times Wage_{j,k,t}}{\sum Emp_{j,k,t} \times Wage_{j,k,t}} \tag{1}$$

where $Prob_k$ is the probability of workplace automation for occupation k, calculated by Frey and Osborne (2017). $Emp_{j,k,t}$ and $Wage_{j,k,t}$ represent the number of employees and the average annual salaries of human labor allocated to occupation k of industry j, in a given year t. Following Donangelo (2014), Zhang (2019), and Bates, Du, and Wang (2020), the weights are assigned to the portion of employees across all occupations in each industry using the annual salaries of human laborers in that occupation to reflect human labor's influence on corporate cash flows. Therefore, the proxy has been updated to the weighted average capability of workplace automation across all professions of industry j in year t. Holding $Prob_k$ constant, changes in *AI replacement* over time indicates the dynamics in the occupational distribution of employees of a given industry j.

Furthermore, in our paper, we upgrade the proxy from the industry to the firm level using the segment data from the Historical Segments Database in *Compustat*. The segment data reflects the industries a company operates in. Suppose a firm operates in m different industries. In that case, the firm-level AI replacement is the weighted average of all the number (m) of different industries (j)' *AI replacement* each year t. The value of AI replacement can range from zero to one. A higher value means the firm can substitute human workers with automated capital, while a lower value means the firm is less susceptible to workplace automation. Specifically, the firm-level *AI replacement* is shown as the following equation:

$$AI\ replacement_{i,t} = \sum_{j=0}^{m} SLAC_{j,t} \times Weight_{i,j} \qquad (2)$$

3.3. Control variables

We obtain the data to construct control variables from Compustat. Following prior studies on CSR, we include a vector set of firm-specific characteristics as control variables. The control variables in our baseline model are total long-term debt over total assets (Leverage), market value of assets scaled by the book value of assets (Tobin's q), Herfindahl-Hirschman Index (HHI), operating income scaled by total assets (ROA), cash balance scaled by total assets (Cash ratio), the logarithm value of the book value of total assets (Firm size), the number of years since the first appearance of a firm in Compustat (Firm age), and research and development expense over total assets (R&D). Detailed variable definitions are provided in Appendix A.

3.4. Summary statistics

We report the summary statistics in Table 1. We can see that the average *Employee CSR* is -0.03. Regarding our explanatory variable, *AI replacement*, the average level is 0.47, which is generally like the statistics reported by Bates, Du, and Wang (2020). Before moving on to our multivariate analysis, we review the correlation matrix presented in Table 2. The results in Table 2 raise little concerns about potential multicollinearity in our subsequent regression analysis.

[Please Insert Table 1 and Table 2 here]

4. Empirical results

4.1. Workplace automation and employee CSR

We employ the pooled OLS panel regression analysis methodology to empirically test the potential impact of AI replacement on Employee CSR. Our multivariate regression analysis starts from the baseline regression analysis on the relationship between workplace automation and employee-related CSR performance among U.S. public firms. The model is displayed as Equation (3):

$$\begin{aligned} &Employee\ CSR\ (Wokplace\ safety)_{jt} \\ &\qquad = \beta_1 \cdot AI\ replacement_{jt} + \beta_2 \cdot Controls_{jt} + Year\ FE + Industry\ FE \\ &\qquad + \varepsilon_{jt} \qquad (3) \end{aligned}$$

The dependent variable is the employee-related CSR performance at firm j in year t. The explanatory variable is *AI replacement*, measured by Equation (2). We also include the control variables that are measured in the prior year. In addition, we have year and industry-fixed effects to cope with the unobserved heterogeneity over time and across industries.

We present the empirical results using Equation (3) to examine the impact of *AI replacement* on *Employee CSR* in Table 3. In columns (1) and (2), the coefficients of *AI replacement* are negative and statistically significant at either 1% or 5% level when including the fixed effects. These results are economically significant. For instance, after controlling for firm fixed effects, a one-standard-deviation increase in *AI replacement* can reduce firms' commitment to *Employee CSR* by 0.020. We interpret the above empirical results as solid support for our conjecture that workplace automation can create a moral hazard. Since corporate managers are aware that their human capital can eventually be substituted with robots, they may think it is unnecessary to devote sufficient resources to maintain a healthy and safe working environment for their workers.

[Please Insert Table 3 here]

5. Endogeneity concern

Identifying a causal link between workplace automation with artificial intelligence and corporate employees' well-being may be subject to potential endogeneity concerns. It is possible that other unobserved firm-level factors are correlated with employee-related CSR/workplace safety and AI replacement. Thus, this may hinder the causal interpretation of the proposed relationship. Also, there might likely be a reverse causality that firms with poor records on employees' well-being are significantly different from other firms. Such firms are more likely to utilize AI-related technologies to replace human labor.

To further reinforce a casual interpretation of the impact of AI replacement on corporate employees' well-being, we design a quasi-natural experiment by using a plausibly exogenous shock and then perform the difference-in-differences estimation. Ideally, we should employ a natural experiment by using an exogenous shock involving policy initiatives directly related to the development of artificial intelligence. Hence, we carefully look through the legislation and executive orders of the U.S. government related to AI research and applications from the National Artificial Intelligence Initiative Office (www.AI.gov). The office website contains all federal laws and executive orders that aim to advance the study and application of AI. Unfortunately, we have found that all the AI policy initiatives were passed after 2016, which is beyond our sample period. Thus, relying on those policy initiatives to design the natural experiment is not feasible.

Nonetheless, after delving into U.S. federal laws and executive orders published in the past decades, we have found that an act passed in 2007 is related to AI research and application to a large extent, which is the 2007 America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (*America COMPETES Act*). This act became the federal law of the U.S. on August 9th, 2007. The act was one of the most prominent bipartisan legislative accomplishments over the past decades and has been recognized as one of the most monumental science and innovation policy initiatives in the new century (Furman, 2013; Gonzalez, 2015). This policy initiative aims to bolster U.S. firms' innovation via investment in research and development and, in turn, stimulate the U.S's competitiveness. Under the *America COMPETES Act*, several merit-based partnerships between the federal government and businesses are extended and created to provide fundamental support for innovation that can lead to transformative advances in manufacturing technologies and processes. In addition, the *America COMPETES Act* aims to incentivize U.S. firms in the private sector to support manufacturing in the U.S. via ameliorated performance, productivity, sustainability, and competitiveness. Specifically, this law encourages research areas on manufacturing and construction machines and equipment that include robotics, automation, and other artificially intelligent systems[1]. Since then, the achievement of deep learning research has been substantially advanced, and deep understanding has started becoming the new trend of artificial intelligence. Therefore, the development on artificial intelligence significantly boosts the real-world application of artificial intelligence in substituting skilled workers' non-routine jobs dramatically (Hinton, Osindero, and The, 2006).

To better identify the causal effect and eliminate potential selection bias, we apply a difference-in-differences estimation post propensity score matching routine (PSM-DID) (Rosenbaum and Rubin, 1983; Bertrand et al., 2004) to estimate the treatment effect of workplace automation more effectively on employees' welfare. We first perform the propensity score matching (PSM) routine. To facilitate this, we divide the sample firms into two groups (treatment and control). The treatment group includes firms whose *AI replacement* values are higher than the median, while the control group consists of firms whose *AI replacement* values are lower than the median. Precisely, we match firms with above median *AI replacement* with firms whose *AI replacement* values are lower than the median on the same control variables as

[1] https://www.govinfo.gov/content/pkg/PLAW-110publ69/pdf/PLAW-110publ69.pdf

our baseline model using a one-to-one nearest neighbor matching each year without replacement. Next, we perform the post-matching DID estimation based on the matched sample.[2] To balance the sample, we only include the three years before and after the event year (2007). With this, we estimate the following regressions:

$$Employee\ CSR_{\ jt} = \beta_1 \cdot AI\ replacement_{it} * Post_{it} + \beta_2 \cdot AI\ replacement_{it} + \beta_3 \cdot Post_{it} + \beta_4 \cdot Controls_{jt} + Year\ FE + Industry\ FE + \varepsilon_{jt} \quad (4)$$

with the variable *Employee CSR*, the control variables, and the included fixed effects being the same as those in equation (3). The dummy variable $Post_{it}$ equals one for the years after 2007 and equals zero otherwise. Our main variable of interest is the interaction between treated firms and the post publication period $(AI\ replacement_{it} \times Post_{it})$. The coefficient on the interaction term β_1 captures the incremental change in employee-related CSR from the pre to the post *America COMPETES Act* period for firms in the treatment group relative to the change for firms in the control group. A negative coefficient on β_1 reveals that after the passage of the *America COMPETES Act* and AI technology being advanced, employees' welfare deteriorates even more among firms with higher level of *AI replacement* compared to their counterparts.

Table 4 reports the empirical results of the PSM-DID estimation. We observe that the coefficients of the DID estimators are negative and statistically significant in Table 4. This is broadly consistent with our conjecture and further confirms the causal link between *AI replacement* and employees' well-being.[3] Further, as a placebo test, we employ the same PSM-DID strategy to examine the possible impact of AI replacement on CSR in other dimensions than employee-related.

---

[2] In unreported tests, we perform the balanced test (t-test) on our matching variables and find no significant differences in most matching variables between treated firms and controls. We conclude that the PSM routine successfully eliminates the differences between the two groups of firms.

[3] We also perform the PSM-DID estimations using our alternative workplace safety measures in our unreported tests. The results are broadly similar to those reported in Table 4.

[Please Insert Table 4 here]

6. Further analysis

In this section, we identify and discuss several possible mechanisms through which substituting labor force by automated capital may impede employees' well-being. We have pinpointed potential channels from both firm-level and state and county levels.

6.1. Decomposition of the employee CSR

According to the KLD database, the employee CSR score consists of 16 sub-components (both strengths and concerns). Specifically, the strengths include union density, cash profit sharing, employee involvement, access to health plans, supply chain labor standards, human capital development, labor management, controversial sourcing, and other strengths. The concerns include union relationship concerns, health & safety concerns, supply chain labor standards, child labor, labor management relations, retirement benefits concerns, and other concerns. We would like to figure out the poor employee-related CSR practice among firms that are more susceptible to workplace automation stems from which specific factors of the employee CSR score. Put differently, we aim to find out if a firm is more capable of replacing human workers with automated capital, which factors of the employee CSR the firm tends to overlook. We then run regression analyses for each of the 16 sub-components. Six of the 16 sub-components are significantly related to workplace automation. The six factors are union density, union relationship concerns, employee involvement, human capital development, workplace safety concerns, and retirement benefit concerns. The results for the six factors are reported in Columns 1-6 in Table 5.

Specifically, in Column 1, union density positively relates to *AI replacement*. When a firm is more likely to replace human employees with automated capital, the percentage of the employees covered by the union tends to increase as the workers are worried about their welfare and hope to seek protection from the union. In Column 2, the union relationship concerns are positively related to *AI replacement*, indicating that as firms with a higher likelihood of *AI replacement* tend to downplay the employees' well-being, they are more likely to incur conflicts with the union. In Column 3, the employee involvement is negatively related to *AI replacement*. When a firm is more likely to rely on an automated workforce, it is less likely to use employee stock ownership plans (ESOPs) or employee stock purchase plans (ESPPs) to incentivize employee involvement. The result in Column 4 suggests that firms with a higher likelihood of *AI replacement* are less likely to offer competitive benefits packages and performance incentives to develop human capital. About the result in Column 5 reveals that firms with higher probability of workplace automation tend to have more job accidents, injuries, fatalities, and employee mental health issues. In Column 6, the retirement benefits concerns are positively associated with *AI replacement*, indicating that firms that are more likely to utilize an automated workforce tend to underfund the pension plan or provide inadequate retirement benefits.

[Please Insert Table 5 here]

6.2. Union protection

It has been documented by Bradley, Mao, and Zhang (2021) that union protection is closely bound up with corporate employees' welfare. Particularly, they provide empirical evidence that union protection can significantly reduce corporate employees' injuries and casualties occurred in the workplace. As a union is established with the aim of unifying workers and advocating for better working conditions as well as higher wages, it should effectively monitor the employers to ensure their workers are fared well. However, when labor unions do not protect the workers, employers dare to attach less attention on workers' well-being if they can substitute those humans with robots. To investigate the role of union protection in the link between AI replacement and employees' well-being, we collect data on labor unions compiled by Hirsch and Macpherson (2003), available from the website of unionstats.gsu.edu. In this dataset, we can find data on union membership and collective bargaining agreement coverage. Subsequently, we construct two proxies to capture the union protecting role for their workers. The first one is *Cov_percent*, the percentage of a company's employees covered by the collective bargaining agreement. The second one is *Mem_percent*, the percentage of a company's employees having union membership.

We conjecture that if a company's susceptibility to workplace automation is high, their employees would be aware of the employment risk. Therefore, they tend to seek protection from labor unions. In line with this rationale, a higher likelihood of AI replacement for a company should be associated with a higher percentage of workers seeking union protection. Our empirical evidence in Panel A, Table 6, supports our conjecture. When we regress the proxies of union protection on AI replacement and a set of control variables, we can see that the coefficients of *AI replacement* are positive and statistically significant at the l% level.

Next, we examine the role of union protection on the negative link between AI replacement and employees' benefits. Intuitively, when a firm can substitute human workers with robots, they dare to pay less attention to their employees' welfare if they lack protection from a labor union. To empirically examine this hypothesis, we construct an interaction term between the proxy of union protection and AI replacement and include the interaction term in the regression analysis. Our variables of interest are the interaction terms of *Cov_percent* × *AI replacement* and *Mem_percent* × *AI replacement*. The empirical results in Panel B, Table 6 suggest the negative link between AI replacement and employees' well-being is attenuated by the existence of union protection. On the other hand, it indicates that the decreased employee welfare due to AI replacement is further exacerbated without union protection[4].

[Please Insert Table 6 here]

[4] For brevity, we only report the impact of AI replacement on overall employee-related CSR performance for the mechanism analysis. The results using workplace safety measures are qualitatively similar and are available upon request.

6.3. Labor-induced operating leverage

One salient benefit of substituting human workers with automated capital is that firms subjected to a considerable level of fixed labor cost can significantly reduce the part of the operating leverage induced by labor. Indeed, Bates, Du, and Wang (2020) document that firms with a high likelihood of utilizing automated capital to replace their human labor tend to adopt aggressive financial policies. They argue that substituting humans with robots enables such firms to reduce their operating leverage stemming from high levels of fixed labor cost and, in turn, to increase their financial flexibility. Therefore, such firms choose to use aggressive monetary policies. In line with their argument, we expect that a firm's desire to replace human workers with automated capital is relatively high if it incurs a high fixed cost of labor. However, if such firms are aware that they can eventually get rid of the high labor-indued operating leverage, they may think it is not necessary to focus on human employees' benefits. Thus, we expect the performance of employee-related CSR to be even worse when a firm has a high level of labor-induced operating leverage.

To test this channel, following Bates, Du, and Wang (2020), we measure labor-induced operating leverage by Selling, General & Administrative expenses (SG&A) scaled by total assets. We split the sample of firm-year observations into two firms: firms with below-median labor-induced operating leverage and firms with above-median labor-induced operating leverage. We report the empirical results in Table 7. In Table 7, column (1) are the firms with below-median SG&A ratio, while column (2) is the firms with above-median SG&A ratios. We can see clearly that *Employee CSR* is only statistically significant within the above-median group, suggesting that firms susceptible to workplace automation tend to downplay employees' when they face a significantly high level of labor-induced operating leverage.

[Please Insert Table 7 here]

6.4. Cash holdings

Cohn and Wardlaw (2016) show that when companies face financial constraints (e.g., low cash balance), they are inclined to compromise on the safety requirements at the workplace. In another vein, Bates, Du, and Wang (2020) indicate that firms choose to hold less cash if they can substitute human workers with automated capital. The rationale is that workplace automation enables firms to relax their operating leverage and adopt less conservative financial policies. Thus, such firms do not need to hold the precautionary cash to deal with adverse cash flow shocks when their operating leverage is high. Therefore, such firms may invest aggressively while paying less attention to their employees' well-being, given that they are also aware that they can eventually replace human workers with AI. In line with the above arguments, we postulate that firms have an even poorer record of employee CSR when their cash ratio is lower. Table 8 displays the empirical results. Our focus is the interaction term between AI replacement and cash ratio. Consistent with our hypothesis, the coefficients of the interaction term are positive and statistically significant at a 1% level, indicating that more cash holdings can mitigate the negative influence of AI replacement on employees' welfare, while fewer cash holdings can worsen the such negative impact.

[Please Insert Table 8 here]

6.5. Labor cost

So far, we have understood that a significant concern that motivates companies to choose automated capital is the sticky fixed cost of hiring human laborers. Substituting such labor with robotic capital can substantially alleviate human labor-related fixed costs, and firms, in turn can achieve improved financial flexibility. However, such a concern can become more pronounced if the labor cost is high. To empirically investigate this channel, we rely on the data on state-level minimum wage that is available from the U.S. Department of Labor website. To make our results more convincing, we use both nominal and real minimum wages (adjusted for inflation). As a higher state-level minimum wage implies higher labor costs for the companies in a given state, companies are more willing to replace the expensive human workers with robots if they can. Thus, they could care even less about human workers' well-being. To test this channel, we construct an interaction term between minimum wage and AI replacement. Specifically, we interact AI replacement with either nominal or real minimum wage. The results in Table 9 reveal that the negative impact of AI replacement becomes more pronounced when firms face higher labor costs, regardless of using nominal or real minimum wage.

[Please Insert Table 9 here]

6.6. Job market condition

When the job market is of a bearish form with a high unemployment rate, it can be tricky for human workers to seek alternative job opportunities if they are not satisfied with the current working conditions and salary. Under this scenario, workers have weak bargaining power for better treatment from their employers. The moral hazard problem, in this case, can further exacerbate as employers know their workers do not have available alternatives; employers can somehow take advantage of a bearish job market and choose not to make an effort on employee-related CSR. Thus, we conjecture that the impact of AI replacement on employee CSR tends to be more negative when the unemployment rate is high. To empirically test this mechanism, we add an interaction term between the state-level unemployment rate and AI replacement to our baseline model and focus on the interaction term's coefficient. The state-level unemployment data is obtained from the U.S. Department of Labor website. The empirical results about the job market condition channel are exhibited in Table 10. We can see that the coefficients of *Unemployment* × *AI replacement* are negative and statistically significant at conventional levels. These results reveal that a high unemployment rate can further exacerbate the negative influence of workplace automation on firms' employee-related CSR performance.

[Please Insert Table 10 here]

7. Actual costs of poor employee CSR practice

In this section, we would like to investigate further whether or not there exist any real consequences for the firms capable of utilizing the AI replacement while paying less attention to employees' well-being.

7.1. Employer reputation

Firstly, we want to explore whether there is a reputation cost for firms with a higher likelihood of workforce automation if they overlook their employees' well-being. To proxy an employer's reputation, we use Fortune magazine's ranking of best employers (top 100 employers) published annually between 2005 and 2016. We then construct an indicator variable as the dependent variable that takes the value of one if a firm is ranked as one of the top 100 employers by *Fortune* and zero otherwise. Since the dependent variable is binary, we estimate a probit regression in which *AI replacement* is the explanatory variable. The results are displayed in Table 11. In Table 11, the coefficient of *AI replacement* is negative and statistically significant at 5%. About the marginal effect, a one-standard-deviation increase in AI replacement decreases the probability of being recognized as one of the best employers by 11.5%. We interpret the finding as that firms with a high likelihood of utilizing automated capital can suffer significant reputation damage, presumably because such firms tend to overlook their employees' well-being.

[Please Insert Table 11 here]

7.2. Firm performance

Next, we investigate whether such unethical behavior can trigger any negative financial implications for firms susceptible to AI replacement. Firstly, we examine the firm value using Tobin's q and ROA. We then use the interaction term between *AI replacement* and *Employee CSR* to investigate the research question. Surprisingly, from the first two columns of Table 12, we can see that the interaction term coefficients are negative and statistically significant at either 1% or 5% level, suggesting that firms rationally choose to decrease employee CSR if AI can easily replace employees. Otherwise, the firm's value will fall. Economically, a one-standard-deviation increase in the interaction term reduces Tobin's q by 0.129. Overall, the above evidence indicates that firms' unethical behavior is rewarded by the higher firm performance.

[Please Insert Table 12 here]

7.3. Workplace safety

Given that an emerging but growing string of studies has shed light on the link between various corporate finance activities and workplace safety (see, e.g., Cohn and Wardlaw, 2016; Caskey and Ozel, 2017; Bradley, Mao, and Zhang, 2021), we want to further investigate the negative impact of workplace automation on corporate employees' well-being by examining the workplace safety among the firms with a high likelihood of substituting human workforce with automated capital. Specifically, we intend to determine whether such firms are associated with higher social costs of poorer workplace safety records (higher workplace injury rates) than their counterparts. We obtain the data about workplace safety from the OSHA Data Initiative Program under the Occupational Safety Health Administration (OSHA)[5]. OSHA consists of survey data from private sector establishments on reported injuries and illnesses stemming from work-related activities between 2002 and 2011. Our sample stops at 2011 because the OSHA data is unavailable beyond 2011. Also, we start from 2002 because, in 2002, OSHA updated its recoding criteria for workplace injuries, illnesses, and casualties, making it incomparable with the data before 2002. Therefore, the empirical analysis of workplace safety is based on our subsample from 2002 to 2011. All establishments under OSHA jurisdiction are required to make records of the injuries and illnesses as well as casualties attributed to work-

[5] It is worth noting that OSHA does not collect workplace safety information regulated by other federal agencies. For instance, the Mine Safety and Health Administration and the Federal Aviation Administration.

related activities that are available for OSHA inspection if 11 or more injuries or illnesses, or losses incur among the employees unless an establishment belongs to the industry that is on the OSHA exemption list[6]. We then manually merge each establishment available in OSHA to firms included in *Compustat* by the company name on an annual basis (see, e.g., Cohn and Wardlaw, 2016; Caskey and Ozel, 2017; Bradley, Mao, and Zhang, 2021).

Following the abovementioned literature, we create three proxies to capture a firm's workplace safety. The first one is the Total Case Rate (*TCR*), which is constructed as the sum of all injuries and illnesses as well as casualties that lead to days away from work or with work restriction or work transfer, and other cases required to record scaled by the aggregate number of hours worked by all employees in a given establishment year, and then multiplied by 200,000. The second one is called the injury rates with days away, restricted, or transferred (*DART*), which is computed as the number of injuries and illnesses that lead to days away from work or with work restriction or transfer, scaled by the total number of hours worked by all employees in a given establishment year, and then multiplied by 200,000. The third one is called injury rates with days away from work (*DAFW*), which equals the number of injuries and illnesses that lead to days away from work, divided by the aggregate number of hours worked by total employees in a given establishment year, and then multiplied by 200,000. Finally, our subsample includes 35,742 firms and 23,438 establishments that belong to 1,985 unique firms, and at last, 3,151 firm-year observations.

We apply the same empirical strategy adopted to investigate the impact of AI replacement on employee-related CSR to examine how AI replacement may affect the injuries and casualties in the corporate workplace. The model is specified as follows in Equation (4):

$$Workplace\ safety_{jt} = \beta_1 \cdot AI\ replacement_{jt} + \beta_2 \cdot Controls_{jt} + Year\ FE + Industry\ FE + \varepsilon_{jt} \qquad (4)$$

6 OSHA exempts an industry if it has a history of low accident rates.

Table 13 reports the empirical result of AI replacement's influence on workplace safety. Where the dependent variables are the three proxies of workplace safety, TCR, DART, and DAFW, our interest variable remains AI replacement. We can see clearly that the coefficients of all three measures are negative and statistically significant at conventional levels, indicating a substantial and adverse link between AI replacement and workplace safety. Economically speaking, a one-standard-deviation increase in firms' capability of replacing human workers with robots triggers 0.626 (*TCR*) more injuries and casualties in the corporate workplace. Thus, we argue that the empirical analysis strongly underpins our hypothesis that firms' susceptibility to utilizing automated capital to replace human laborers can deteriorate workplace safety, echoing our main finding on employee-related CSR.

[Please Insert Table 13 here]

8. Robustness test

Even though endogeneity was previously discussed, other factors might impact the outcomes. To ensure the article's overall robustness, we will further incorporate these potential influencing aspects into the analytical framework in this part.

8.1.The replacement by industrial robots

The independent variable *AI replacement* refers the possibility of AI replacing people. Nevertheless, it is essential to note that industrial robots have been extensively utilized in various industries, and our research should consider this reality. To empirically test the influence of robots on *Employee CSR,* we add two variables - *Robots installed* and *Robots stock* - to our baseline model. The results shown in Table 14 provide compelling evidence that the inclusion of the impact of robots on the coefficient of *AI replacement* yields a negative, statistically significant outcome - one in line with our benchmark results. Interestingly, the coefficients for *Robots install* and *stock* are both significantly positive. This could be attributed to the fact that the replacement of workers by robots tends to be a progressive process, meaning that the costs of replacing labor with robots rise as the complexity of the required tasks increases. In environments where robots work, employees represent a scarce resource, thus commanding a higher value for employers who consequently pay more attention to *Employee CSR*.

[Please Insert Table 14 here]

8.2.The impact of routine-task intensity (RTI)

To evaluate the extent to which our findings concerning *AI replacement* and *Employee CSR* are attributable to the substitutability of routine-task labor versus non-routine-task labor, we include Autor and Dorn's (2013) industry-by-year measure of routine-task intensity (RTI) into our models.

Table 15 presents the empirical results. The first and second columns of results show that the coefficients of variables RTI are insignificant regardless of the *AI replacement* inclusion or exclusion, indicating that the substitutability of routine-task labor versus non-routine-task labor does not influence *Employee CSR*. Furthermore, column (2) shows that, even when RTI is incorporated into the analysis framework, the coefficient of the main explanatory variables remains negative and is statistically significant at the 5% level. This indicates that a firm's potential to replace routine-task laborers with automation does not affect our main regression results, which are therefore considered relatively robust.

[Please Insert Table 15 here]

8.3. ESG data

Research objects of ours are equally concerned with the relevant welfare of employees. Currently, research into ESG is experiencing vigorous growth and, to some extent, is an evolution and development from CSR research. We explore the influence of *AI replacement* on ESG by replacing the explanatory variables and further examine the robustness of our results.

Table 16 reports the impact of *AI replacement* on employees' welfare using the Thomson/Refinitiv ESG database data. The dependent variable is the *Workforce ESG,* the total workforce-related ESG performance score. Workforce-related ESG includes diversity and opportunity, employment quality, health and safety, and training and development. Our results are economically significant; for instance, a one-standard-deviation increase in *AI replacement* can reduce *Workforce ESG* by 0.015. Moreover, this further attests to the robustness of our results.

[Please Insert Table 16 here]

9. Conclusion

In this research, we focus on the dark side of adopting breakthrough technologies into business operations in modern corporations and seek empirical evidence for the potential shortcomings of workplace automation in public firms. The capability of utilizing AI and automation to replace human workers could trigger a moral hazard. Specifically, if corporate managers anticipate that they are highly likely to substitute the current labor with automated capital, they tend to devote fewer resources to employees' well-being. Therefore, such firms tend to have higher workplace injuries and overall employee-related CSR performance. Our empirical analysis well bolsters our conjecture. We find significant empirical evidence that firms more susceptible to workplace automation have more dangerous workplaces associated with more injuries and casualties and lower employee CSR performance than their counterparts. However, our empirical evidence suggests that firms' capability of workplace automation only affects employee-related CSR while it does not affect CSR performance in other categories. To attenuate the endogeneity concern and further strengthen our causal interpretation, we use instrumental variable analysis and rely on a plausibly exogenous shock to perform the difference-in-differences estimation. Again, the results are consistent with our baseline results. To further identify the potential mechanisms through which workplace automation discourages employees’ well-being, by decomposing the employee-related CSR into different categories, we find that firms that are more susceptible to workplace automation significantly decrease their commitment to employee involvement as well as human capital development, conversely, the concerns on their relationship with the union, employees’ health and safety and retirement benefits have been significantly exacerbated.

Additionally, our analyses indicate that the negative link between workplace automation and employees’ well-being is more substantial in firms adopting more aggressive financial policies and having higher labor-induced operating leverage. Also, the effect of automation becomes more negative on employees’ welfare when the workers have less union protection. Moreover, we find some plausible mechanisms at the state and country level. Our analysis suggests that firms with the ability to adopt automated capital are inclined to ignore employees’ benefits when the labor costs are higher, when the labor market is of a bearish market, and when the firms locate in less religious counties. Finally, overlooking employees’ well-being incurs reputation damage and is financially costly. Firms with more capability of replacing human labor with robots while neglecting employees’ welfare are less likely to be ranked as the best U.S. employers by *Fortune*. Financially speaking, such firms are surprisingly rewarded with higher accounting performance and firm value, indicating an ethical conundrum for corporate managers.

There is no clear clue when machine learning can reach fully adaptative intelligence like humans and when full automation can be achieved. Such a research question is worth investigating. Even though AI has been replacing more and more non-routine jobs requiring skilled workers, at the current stage, human beings are still far more adaptive than artificial intelligence. Humans are still far more able to adapt to change than artificial intelligence (AI). Thus, in the foreseeable future, human workers' crucial role will still be in partially automated workplaces. However, we show that overlooking human workers can trigger substantial financial costs of poor performance and the social cost of high injury rates in the workplace. Our research can assist policymakers in better understanding the importance of humans' and robots' coexistence and making decisions on how to ameliorate the coexistence. Besides the firm-level benefits of AI applications, a critical help at the aggregate level is that AI can contribute to the sustainable environment by reducing the corporate carbon footprint (*Forbes*, 2020). Hence, policymakers should encourage AI applications healthily and ethically.

References

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review* 108(6), 1488-1542.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy* 128(6), 2188-2244.

Attig, N., El Ghoul, S., Guedhami, O., & Suh, J. (2013). Corporate social responsibility and credit ratings. *Journal of Business Ethics*, 117(4), 679-694.

Autor, D., & Salomons, A. (2018). Is automation labor-displacing? Productivity growth, employment, and labor share. *NBER Working paper*.

Bates, T. W., Du, F., & Wang, J. J. (2020). Workplace automation and corporate financial policy. Available at: https://papers.ssrn.com/3556935. *Unpublished Working Paper*.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* 119(1), 249-275.

Bradley, D., Mao, C. X., & Zhang, C. (2021). Does Analyst Coverage Affect Workplace Safety? *Management Science*, forthcoming.

Brammer, S., Williams, G., & Zinkin, J. (2007). Religion and attitudes to corporate social responsibility in a large cross-country sample. *Journal of business ethics* 71(3), 229-243.

Büchel, B., & Floreano, D. (2021). *Tesla's problem: overestimating automation, underestimating humans.* Available at: https://www.imd.org/research-knowledge/articles/teslas-problem-overestimating-automation-underestimating-humans/ (accessed on December 1, 2021).

Cahan, S. F., Chen, C., Chen, L., & Nguyen, N. H. (2015). Corporate social responsibility and media coverage. *Journal of Banking & Finance* 59, 409-422.

Caskey, J., & Ozel, N. B. (2017). Earnings expectations and employee safety. *Journal of Accounting and Economics* 63(1), 121-141.

Cheng, X., Lyandres, E., Zhou, K., & Zhou, T. (2021). Industrial Robots and Finance. Available at: https://papers.ssrn.com/3736000. *Unpublished Working Paper*.

Cohn, J. B., & Wardlaw, M. I. (2016). Financing constraints and workplace safety. *The Journal of Finance* 71(5), 2017-2058.

Cohn, J., Nestoriak, N., & Wardlaw, M. (2021). Private equity buyouts and workplace safety. *The Review of Financial Studies* 34(10), 4832-4875.

David, H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review* 103(5), 1553-97.

Deng, X., Kang, J. K., & Low, B. S. (2013). Corporate social responsibility and stakeholder value maximization: Evidence from mergers. *Journal of Financial Economics* 110(1), 87-109.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review* 86(1), 59-100.

Donangelo, A. (2014). Labor mobility: Implications for asset pricing. *The Journal of Finance* 69(3), 1321-1346.

Du, X., Jian, W., Zeng, Q., & Du, Y. (2014). Corporate environmental responsibility in polluting industries: Does religion matter? *Journal of Business Ethics* 124(3), 485-507.

El Ghoul, S., Guedhami, O., Kwok, C. C., & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance* 35(9), 2388-2406.

Electrek (2021, August 4). *Tesla Teases Future Products Using Artificial Intelligence.* Available at: https://electrek.co/2021/08/04/tesla-teases-future-products-using-artificial-intelligence/amp/ (accessed on December 1, 2021).

Financier Worldwide (2021, December 1). *Clarity and Comfort: The Importance of Quality Auditing.* Available at: https://www.financierworldwide.com/clarity-and-comfort-the-importance-of-quality-auditing (accessed on December 7, 2021).

Forbes (2019, March 1). *Inside Tesla's Model 3 Factory, Where Safety Violations Keep Rising.* Available at: https://www.forbes.com/sites/alanohnsman/2019/03/01/tesla-safety-violations-dwarf-big-us-auto-plants-in-aftermath-of-musks-model-3-push/?sh=7f49f54554ce (accessed on December 1, 2021).

Forbes (2020, November 16). *The Role of AI In Carbon Reduction and Increased Efficiency for Energy.* Available at: https://www.forbes.com/sites/forbestechcouncil/2020/11/16/the-role-of-ai-in-carbon-reduction-and-increased-efficiency-for-energy/?sh=6f096a045c58 (accessed on December 1, 2021).

Fortune (2018, October 22). *The Four Waves of A.I.* Available at: https://fortune.com/2018/10/22/artificial-intelligence-ai-deep-learning-kai-fu-lee/ (accessed on December 1, 2021).

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerization? *Technological forecasting and social change* 114, 254-280.

Furman, J. L. (2013). The America COMPETES Acts: The Future of US Physical Science and Engineering Research? *Innovation policy and the economy* 13(1), 101-149.

Gonzalez, H. B. (2015, May). The America COMPETES Acts: An Overview. Library of Congress, Congressional Research Service.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review* 104(8), 2509-26.

Graetz, G., & Michaels, G. 2018. Robots at work. *Review of Economics and Statistics* 100,753-68.

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation* 18(7), 1527-1554.

Hirsch, B. T., & Macpherson, D. A. (2003). Union membership and coverage database from the current population survey: Note. *ILR Review* 56(2), 349-354.

Jiao, Y. (2010). Stakeholder welfare and firm value. *Journal of Banking & Finance* 34(10), 2549-2561.

Jiraporn, P., Jiraporn, N., Boeprasert, A., & Chang, K. (2014). Does corporate social responsibility (CSR) improve credit ratings? Evidence from geographic identification. *Financial Management* 43(3), 505-531.

Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance* 43, 1-13.

Lee, D. (2017). Corporate social responsibility and management forecast accuracy. *Journal of Business Ethics* 140(2), 353-367.

Leigh, J. P. (2011). Economic burden of occupational injury and illness in the United States. *The Milbank Quarterly* 89(4), 728-772.

Moore, S., & Simon, J. L. (1999). The greatest century that ever was: 25 miraculous trends of the past 100 years. *Cato Institute*.

Ramasamy, B., Yeung, M. C., & Au, A. K. (2010). Consumer support for corporate social responsibility (CSR): The role of religion and values. *Journal of Business Ethics* 91(1), 61-72.

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)* 45(2), 212-218.

Webb, M. (2020). The impact of artificial intelligence on the labor market. Available at: https://papers.ssrn.com/3482150. *Unpublished Working Paper*.

Zhang, M. B. (2019). Labor - technology substitution: Implications for asset pricing. *The Journal of Finance* 74(4), 1793-1839.

Table 1: Summary Statistics

This table reports the summary statistics (i.e., the mean, median, standard deviation, 25% and 75% quantiles, and the number of data points) for main variables. Variable definitions are provided in Appendix A1.

Variable	N	Mean	Std. Dev.	Median	P25	P75
Employee CSR	18388	-0.032	0.164	0.000	-0.083	0.000
Other CSR	18388	-0.209	0.493	-0.217	-0.533	0.000
AI replacement	18388	0.468	0.145	0.469	0.349	0.594
Leverage	18388	0.184	0.201	0.133	0.001	0.281
Tobin's q	18388	4.214	5.548	2.230	1.106	4.855
HHI	18388	653.155	613.579	436.117	327.842	744.471
ROA	18388	0.102	0.179	0.125	0.074	0.179
Cash ratio	18388	0.137	0.147	0.097	0.037	0.197
Firm size	18388	6.963	1.619	6.836	5.787	8.034
Firm age	18388	21.888	15.950	16.000	9.000	32.000
R&D	18388	0.051	0.101	0.008	0.000	0.066

Table 2: Correlation Matrix

This table reports the correlation matrix for the main sample. Variable definitions are provided in Appendix A1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) *Employee CSR*	1										
(2) *Other CSR*	0.466	1									
(3) *AI replacement*	-0.108	-0.007	1								
(4) *Leverage*	-0.053	-0.042	0.096	1							
(5) *Tobin's q*	0.047	0.025	-0.162	-0.196	1						
(6) *HHI*	-0.085	-0.080	0.393	0.070	-0.106	1					
(7) *ROA*	0.059	0.069	0.205	0.018	-0.070	0.128	1				
(8) *Cash ratio*	0.016	-0.012	-0.284	-0.281	0.340	-0.185	-0.326	1			
(9) *Firm size*	0.061	0.122	0.178	0.292	-0.188	0.117	0.320	-0.397	1		
(10) *Firm age*	0.023	0.108	0.243	0.040	-0.224	0.032	0.175	-0.248	0.467	1	
(11) *R&D*	0.035	0.007	-0.359	-0.149	0.312	-0.253	-0.574	0.470	-0.360	-0.214	1

Table 3. The impact of AI replacement on employee CSR

This table reports the impact of AI replacement on employee CSR among U.S. public firms between 1999 and 2016. The dependent variable is employee CSR. All variables are defined in Appendix A1. Year and industry fixed effects are included in columns (1), and year and firm fixed effects are included in columns (2). We report coefficient estimates with *p-values* in parentheses below. P-values are calculated using industry-level clustered standard errors according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**
	Employee CSR	***Employee CSR***
AI replacement	-0.085**	-0.136***
	(0.034)	(0.007)
Leverage	-0.039***	-0.022*
	(0.001)	(0.054)
Tobin's q	0.000	-0.000
	(0.580)	(0.272)
HHI	0.001	0.003
	(0.195)	(0.353)
ROA	-0.019	-0.017
	(0.168)	(0.166)
Cash ratio	0.008	0.005
	(0.339)	(0.541)
Firm size	0.000	0.013
	(0.805)	(0.134)
Firm age	-0.015	-0.027
	(0.609)	(0.228)
R&D	-0.085**	-0.136***
	(0.034)	(0.007)
Constant	-0.199	-0.482
	(0.614)	(0.283)
Industry FE	Yes	No
Firm FE	No	Yes
Year FE	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.134	0.453

Table 4. Quasi-natural experiment 1 (PSM-DID analysis)

This table reports the impact of AI replacement on employee CSR using a difference-in-differences analysis following the approach of Bate, Du, and Wang (2022). The dependent variable is employee CSR. *Flooding* is a dummy variable that takes the value of 1 in years 2011 and 2012; 0 in other years. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included in columns (1), and year and firm fixed effects are included in columns (2). We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) ***Employee CSR***	**(2)** ***Employee CSR***
AI replacement × Flooding	0.074**	0.050**
	(0.012)	(0.015)
AI replacement	-0.148***	-0.129**
	(0.000)	(0.024)
Controls	Yes	Yes
Industry FE	Yes	No
Firm FE	No	Yes
Year FE	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.122	0.464

Table 5. Decomposition of the employee CSR

This table reports the impact of AI replacement on different components of the employee CSR. he dependent variables in Column (1)-(6) represent each of the six components of the employee CSR: *Union density*, *Union relationship concerns*, *Employee involvement*, *Human capital development*, *Health and safety concerns*, and *Retirement benefits concerns*, respectively. All other independent variables are defined in the Appendix A1. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Union density	***Union relationship concerns***	***Employee involvement***	***Human capital development***	***Health and safety concerns***	***Retirement benefits concerns***
AI replacement	0.063***	0.049*	-0.124**	-0.204***	0.167**	0.184*
	(0.005)	(0.073)	(0.034)	(0.000)	(0.028)	(0.057)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13077	15568	14115	2996	18388	12017
Adj. R-sq	0.165	0.091	0.117	0.032	0.037	0.152

Table 6. Union coverage and membership

This table reports the impact of union coverage and membership on the link between AI replacement and employee CSR. Panel A reports how AI replacement affects employees' union participation. In column (1) , the dependent variable is union coverage in percentage, while in column (2), the dependent variable is union membership percentage. Panel B shows how employees' union participation influence the link between AI replacement and employee CSR. The interaction factor between union participation and AI replacement is included. All other independent variables are defined in the Appendix A1. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A

	(1) ***Cov_percent***	**(2)** ***Mem_percent***
AI replacement	0.092***	0.091***
	(0.000)	(0.000)
Controls	Yes	Yes
Industry	Yes	Yes
Year	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.504	0.505

Panel B

	(1) ***Employee CSR***	**(2)** ***Employee CSR***
AI replacement	-0.100***	-0.100***
	(0.000)	(0.000)
Cov_percent	-0.024	
	(0.222)	
AI replacement × Cov_percent	0.487***	
	(0.000)	
Mem_percent		-0.018
		(0.366)
AI replacement × Mem_percent		0.469***
		(0.000)
Controls	Yes	Yes
Industry	Yes	Yes
Year	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.160	0.160

Table 7. Labor-induced operating leverage

This table reports the influence of labor-induced operating leverage on the link between AI replacement and employees' well-being. Labor-induced fixed cost is proxied by SG&A expenses scaled by total assets. The dependent variable is employee CSR. In column (1), the sample includes the observations with below-median SG&A ratio, while in column (2), the sample includes the observations with above-median SG&A ratio. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**
	SG&A ***Below median***	***SG&A*** ***Above median***
Employee CSR	-0.107	-0.106***
	(0.171)	(0.009)
Controls	Yes	Yes
Industry	Yes	Yes
Year	Yes	Yes
Observations	8501	8259
Adj. R-sq	0.145	0.152

Table 8. Cash holdings

This table reports the influence of cash holdings on the link between AI replacement and employees' well-being. The dependent variable is employee CSR. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) ***Employee CSR***
AI replacement	-0.419**
	(0.021)
AI replacement × *Cash ratio*	1.434***
	(0.008)
Cash ratio	0.044
	(0.585)
Controls	Yes
Industry	Yes
Year	Yes
Observations	18388
Adj. R-sq	0.170

Table 9. Labor cost

This table reports the influence of labor cost on the link between AI replacement and employees' well-being. Labor cost is measured with state-level nominal and real minimum wage. The dependent variable is employee CSR. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**
	Employee CSR	***Employee CSR***
AI replacement	-0.030	-0.030
	(0.115)	(0.108)
AI replacement × Real_wage		-0.023***
		(0.008)
Real_wage		-0.002
		(0.217)
AI replacement × Wage	-0.024***	
	(0.002)	
Wage	-0.001	
	(0.485)	
Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.173	0.175

Table 10. Job market condition

This table reports the influence of labor cost on the link between AI replacement and employees' well-being. Job market condition is measured by state-level unemployment rate. The dependent variable is employee CSR All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)
	Employee CSR
AI replacement	-0.027
	(0.403)
AI replacement × Unemploy	-0.009*
	(0.058)
Unemploy	-0.000
	(0.772)
Controls	Yes
Industry FE	Yes
Year FE	Yes
Observations	18388
Adj. R-sq	0.170

Table 11. Employer reputation

This table reports the impact of AI replacement on the reputation of the U.S. employers, proxied by the ranking of best employers among U.S. public firms between 2005 and 2016. The dependent variable, *Best employer*, is a dummy variable takes the value of one a firm is ranked as one of the top 100 best employers by *Fortune* in a given year, and zero otherwise. Column (1) shows the results from the probit regression. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)
	Best employer
AI replacement	-0.708**
	(0.036)
Controls	Yes
Industry FE	Yes
Year FE	Yes
Observations	15568
Pseudo R-sq	0.165

Table 12. Firm value and accounting performance

This table reports the implications on firm value as well as accounting performance for firms that are highly capable of replacing human workers by automated capital while they downplay the employees' well-being. In column (1), the dependent variable is Tobin's q, while in column (2), the dependent variable is ROA. All other independent variables are defined in the Appendix A1. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**
	Tobin's q	***ROA***
AI replacement	-0.893	0.027
	(0.456)	(0.212)
Employee CSR	0.828***	0.049***
	(0.003)	(0.000)
AI replacement × Employee CSR	-4.970**	-0.121***
	(0.017)	(0.003)
Controls	Yes	Yes
Industry	Yes	Yes
Year	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.207	0.389

Table 13. The impact of AI replacement on workplace safety

This table reports the impact of AI replacement on workplace safety among U.S. public firms between 2002 and 2011. The dependent variables are *TCR*, *DART*, and *DAFW* in column (1), column (2), and column (3) respectively. All variables are defined in the Appendix A1. All independent variables are as of the end of the prior year. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**	**(3)**
	TCR	***DART***	***DAFW***
AI replacement	6.798**	3.472**	1.238***
	(0.012)	(0.024)	(0.009)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	25033	25033	25033
Adj. R-sq	0.203	0.203	0.284

Table 14. The impact of actual workforce replacement by industrial robots on employee CSR

This table reports the impact of actual workforce replacement by industrial robots on employee CSR among U.S. public firms between 1999 and 2016. The dependent variable is employee CSR. All variables are defined in the Appendix A1. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) *Employee CSR*	(2) *Employee CSR*	(3) *Employee CSR*	(4) *Employee CSR*
Robots install	2.553		2.814*	
	(0.116)		(0.085)	
Robots stock		3.990**		3.937**
		(0.024)		(0.025)
AI replacement			-0.211**	-0.200*
			(0.042)	(0.054)
Controls	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Observations	9399	9399	9399	9399
Adj. R-sq	0.118	0.118	0.173	0.175

Table 15. The impact of routine-task intensity (RTI) on employee CSR

This table reports the impact of RTI on employee CSR among U.S. public firms between 1999 and 2016. The dependent variable is employee CSR. All variables are defined in the Appendix A1. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1) ***Employee CSR***	**(2)** ***Employee CSR***
RTI	-0.008	0.021
	(0.760)	(0.398)
AI replacement		-0.095**
		(0.019)
Controls	Yes	Yes
Industry	Yes	Yes
Year	Yes	Yes
Observations	18385	18385
Adj. R-sq	0.126	0.128

Table 16. The impact of AI replacement on employee-oriented welfare (ESG data)

This table reports the impact of AI replacement on employees' welfare by using the data from Thomson/Refinitiv ESG database. The dependent variable is the workforce total score. All variables are defined in the Appendix A1. All independent variables are as of the end of the prior year. Year and industry fixed effects are included. We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)
	Workforce ESG
AI replacement	-0.106*
	(0.054)
Controls	Yes
Industry	Yes
Year	Yes
Observations	11074
Adj. R-sq	0.173

Appendix A1.

This table provides the variable definitions used in this research.

Variables	Definition	Source
Main variables		
Employee CSR	The difference between adjusted total strength score and adjusted total concern score in the employee dimension	KLD database
Other CSR	The sum of the scores in the other five dimensions	KLD database
AI replacement	The weighted average capability of a firm to substitute human capital with automated capital across all occupations in various industries the company operate in	Frey and Osborne (2017) and Bates, Du, and Wang (2020)
Other variables		
Flooding	An indicator variable that equals one for the years of 2011 and 2012, capturing the duration of hard drive crisis in Thailand triggered by the Thailand flooding occurred in 2011.	Bates, Du, and Wang (2022)
Cash ratio	Cash balance over total assets in a fiscal year	Compustat
Firm size	Natural logarithm value of total assets in a fiscal year	Compustat
Leverage	Total long-term debts over total assets in a fiscal year	Compustat
R&D	Research and development expenses over total assets in a fiscal year	Compustat
ROA	Operating income over total assets in a fiscal year	Compustat
HHI	Sum of the squared market share of the total sales of each firm in a 3-digit standard industrial classification (SIC) industry of a fiscal year	Compustat
Tobin's q	Market value of assets divided by the book value of assets. Market value of assets is calculated as: total assets – book value of equity + market value of equity. Market value of equity is calculated by the number of common shares outstanding multiplies the share price	Compustat
Firm age	Logarithm of number of months since the first appearance of a firm in CRSP	CRSP
Cov_percent	The annual percentage of employees in one industry who are covered by the union.	unionstats.gsu.edu
Mem_percent	The annual percentage of employees in one industry who are member of the union.	unionstats.gsu.edu
SG&A ratio	Selling, general and administrative expenses scaled by total assets	Compustat
Wage	Minimum wages under state law	U.S. Department of Labor
Real wage	Inflation-adjusted minimum wages under state law deflated	U.S. Department of Labor
Unemploy	The annual unemployment rate of different states	U.S. Bureau of Labor Statistics

Best Employer	An indicator variable that equals one if a firm is ranked as one of the top 100 best employers by *Fortune* in a given year, and zero otherwise.	Fortune Magazine
TCR	The sum of all injuries and illnesses as well as casualties that lead to days away from work or with work restriction or work transfer, and other cases required to record scaled by the aggregate number of hours worked by all employees in a given establishment year. This variable is the firm level average of all establishments of a firm	OSHA Data Initiative Program
DART	The number of injuries and illnesses that lead to days away from work or with work restriction or transfer, scaled by the total number of hours worked by all employees in a given establishment year. This variable is the firm level average of all establishments of a firm	OSHA Data Initiative Program
DAFW	The number of injuries and illnesses that lead to days away from work, divided by the aggregate number of hours worked by total employees in a given establishment year. This variable is the firm level average of all establishments of a firm	OSHA Data Initiative Program
Robots install	The number of new installed industrial robots (in thousands) in a 3 digit SIC (before 2002) or 4-digit NAICS (after 2002) industry in a fiscal year.	International Federation of Robotics (IFR)
Robots stock	The number of new robots in operation stock (in thousands) in a 3 digit SIC (before 2002) or 4-digit NAICS (after 2002) industry in a fiscal year.	IFR
RTI	The the weighted average (weighted by wage) routine-task intensity across all occupational employment in a 3 digit SIC (before 2002) or 4-digit NAICS (after 2002) in a fiscal year. The variable captures a firm's potential to replace routine-task labors by automation.	Autor and Dorn (2013), Knesl (2022), OES
Workforce ESG	The total score in workforce related ESG performance. Workforce related ESG includes diversity and opportunity, employment quality, health and safety, and training and development.	Thomson/Refinitiv ESG database

Appendix A2. Placebo test: the impact of AI replacement on other CSR

This table reports the impact of AI replacement on other CSR among U.S. public firms between 1999 and 2016. The dependent variable is other CSR (adjusted CSR score from the other six dimensions except employee CSR). All variables are defined in the Appendix A1. Year and industry fixed effects are included in columns (1), and year and firm fixed effects are included in columns (2). We report coefficient estimates with *p-values* in parentheses below. *p-values* are calculated using clustered standard errors at industry level according to 2-digit SIC code. *, **, and *** refer to statistical significance at 10%, 5%, and 1% levels, respectively.

	(1)	**(2)**
	Other CSR	***Other CSR***
AI replacement	-0.005	0.024
	(0.952)	(0.882)
Leverage	-0.099**	-0.001
	(0.010)	(0.981)
Tobin's q	-0.000	-0.000***
	(0.360)	(0.002)
HHI	0.008***	0.059*
	(0.006)	(0.094)
ROA	-0.004	0.038
	(0.941)	(0.298)
Cash ratio	0.062**	0.167***
	(0.014)	(0.000)
Firm size	0.010**	0.018**
	(0.039)	(0.013)
Firm age	0.067	-0.209***
	(0.221)	(0.002)
R&D	-0.005	0.024
	(0.952)	(0.882)
Constant	-2.883**	-14.270***
	(0.046)	(0.000)
Industry FE	Yes	No
Firm FE	No	Yes
Year FE	Yes	Yes
Observations	18388	18388
Adj. R-sq	0.037	0.152