The Elusive Returns to AI Skills: Evidence from a Field Experiment

Teo Firpo,[a] Lukas Niemann,[b] Anastasia Danilov [a]

[a] *School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Strasse 1, 10178, Berlin, Germany.*
[b] *Now at Tanso Technologies, Munich, Germany.*

As firms increasingly adopt Artificial Intelligence (AI) technologies, how they adjust hiring practices for skilled workers remains unclear. This paper investigates whether AI-related skills are rewarded in talent recruitment by conducting a large-scale correspondence study in the United Kingdom. We submit 1,185 résumés to vacancies across a range of occupations, randomly assigning the presence or absence of advanced AI-related qualifications. These AI qualifications are added to résumés as voluntary signals and not explicitly requested in the job postings. We find no statistically significant effect of listing AI qualifications in résumés on interview callback rates. However, a heterogeneity analysis reveals some positive and significant effects for positions in Engineering and Marketing. These results are robust to controlling for the total number of skills listed in job ads, the degree of match between résumés and job descriptions, and the level of expertise required. In an exploratory analysis, we find stronger employer responses to AI-related skills in industries with lower exposure to AI technologies. These findings suggest that the labor market valuation of AI-related qualifications is context-dependent and shaped by sectoral innovation dynamics.

Keywords: return to skills, technological change, labor market, hiring, signaling, human capital, field experiment, AI-related skills

JEL classification: O33, J23, J24, I26

1. Introduction

Over the past decade, firms across sectors have increasingly integrated Artificial Intelligence (AI) technologies[1] into their operations (Babina et al., 2024). In parallel, recent empirical research documents rising labor market demand for AI-related skills in different job functions, often accompanied by wage premiums for workers who possess them (Alekseeva et al., 2021; Squicciarini & Nachtigall, 2021; Stapleton et al., 2021; Zhang et al., 2021). In response to this evolving demand, AI-related training and education programs have proliferated rapidly. For example, the Stanford Institute for Human-Centered AI reports that the number of AI-focused tertiary education programs tripled between 2017 and 2024 (Maslej et al., 2024). Similarly, Coursera – the leading online learning platform – notes that, in 2024 alone, four new learners enrolled in AI-related courses every minute (Coursera, 2025).

These developments suggest that even workers in non-technical roles who acquire complementary AI-related skills may benefit from enhanced employment prospects. Indeed, from a theoretical perspective, there are several reasons to expect employers to value such skills in hiring decisions.

First, a growing body of research links AI adoption to significant productivity gains at both the firm and individual levels. Czarnitzki et al. (2023) show that firms integrating AI technology experience measurable improvements in productivity. AI use is also associated with higher innovation outcomes (Rammer et al., 2022). At the individual level, AI tools have been found to improve worker performance across various roles, from consulting (Dell'Acqua et al., 2023) to customer service (Brynjolfsson et al., 2025). These findings suggest that employers

[1] We follow Acemoglu and Restrepo (2019) and Agrawal et al. (2019) in defining AI as a set of machine learning-based abilities that lower the cost of prediction and pattern recognition, thereby enabling or even automating decision-making in concrete tasks. This definition extends beyond large language models and generative AI – the focus of much recent debate – to include tasks such as forecasting, logistics, and computer vision.

may come to view AI-related qualifications as valuable, not only for improving firm performance but also for enhancing worker productivity.

A second reason why employers may favor candidates with AI-related qualifications lies in the general nature of AI-related human capital. In the framework of personnel economics, general skills are defined as those that enhance a worker's productivity across multiple firms, whereas firm-specific skills generate value only within a particular firm (Becker, 1964). Because general skills are transferable to other firms, employers have limited incentives to invest in them. The training costs for such skills are typically borne by the workers, either through formal education or other pre-employment investments. If AI-related skills are indeed general in nature – as is plausible given their wide applicability across tasks and sectors (Brynjolfsson et al., 2019; Goldfarb et al., 2023) – employers might prefer candidates who already acquired them outside of the organization.

Additionally, several studies suggest a positive effect of multiskilling on firms. Carmichael and MacLeod (1993) argue that multiskilled workers are particularly valuable in the context of technological change, as they have stronger incentives to support labor-saving innovations. Compared to single-skilled workers, they face better prospects of reassignment to other roles within the firm, reducing their resistance to technological displacement. While Carmichael and MacLeod's model focuses on internal labor markets, our study provides empirical evidence on whether such multiskilling signals are rewarded in external hiring decisions. Indeed, empirical evidence supports the productivity benefits of multiskilling. Kim and Park (2003) show that multiskilled workforces are positively associated with firm-level labor productivity growth, particularly in environments characterized by uncertainty. Similarly, Farnham and Hutchinson (2011) document a positive association of multiskilling on establishment-level labor productivity and other firm outcomes. In this context, AI-related qualifications may function as a signal of multiskilling – demonstrating workforce adaptability and broader task flexibility in technologically dynamic workplaces.

Taken together, this evidence suggests that employers should value candidates with AI-related qualifications, as they signal skills likely to enhance individual productivity, firm performance, and technological transformation. However, whether employers reward complementary AI skills in hiring decisions remains an open question, particularly when these skills are not yet explicitly required. Our study addresses this gap by providing causal evidence on how employers respond to AI-related qualifications in early-career recruitment for roles not primarily focused on AI. In a large-scale correspondence study in the UK, we submit 1,185 job applications to entry-level vacancies across a range of occupations, randomly varying whether the résumés include AI-related qualifications.

Our main finding reveals that, on average, including AI-related qualifications in résumés has no statistically significant effect on the probability of receiving a positive response from recruiters. However, this aggregate result masks considerable heterogeneity across job functions. We observe positive and statistically significant effects for applications to Engineering and Marketing roles, while the estimated effects are null in Finance and Accounting, Human Resources (HR), Information Technology (IT), and Logistics and Supply Chain positions. The overall null results are robust to different statistical specifications. To better understand these patterns, we explore several potential explanations for the limited overall returns to AI-related qualifications.

First, we examine whether our null finding might be attributable to features of our experimental design – particularly the specific AI-related qualifications included in the résumés. To assess this, we conduct a survey with 771 professionals with hiring experience, presenting them with the résumés and asking structured questions about the content therein.

Respondents judge the AI-related qualifications to be both noticeable and relevant. Moreover, they perceive them as a signal of higher digital competence, though also associated with somewhat lower social skills. Furthermore, we find empirical evidence that participants expect candidates with AI-related qualifications to receive a higher number of interview invitations. Taken together, the survey's findings suggest that the absence of a statistically significant effect in our correspondence experiment is unlikely to be an artifact of the experimental design. Instead, it likely reflects a genuinely weak employer preference for AI-related qualifications in the studied hiring context.

A second possible explanation for our null finding relates to heterogeneity in the skill requirements specified in the job postings to which we submit applications. Such variation may confound treatment effects, as the marginal value of AI-related qualifications could depend on the breadth, type, and specificity of skills demanded by employers. To account for this, we analyze the textual content of the targeted job postings and map the extracted skill requirements to the ESCO classification system. This allows us to construct measures of (i) the total *number of skills* requested, (ii) the degree of *match* between job requirements and our résumé content, and (iii) the level of *expertise* required. Controlling for these metrics does not alter our results.

Finally, we examine the interaction with the industry-level exposure to AI automation. To do so, we incorporate the industry-level *AI exposure* measure developed by Felten et al. (2021) into our analysis. We find that listing AI-related qualifications increases the likelihood of receiving a callback by approximately 4.4 percentage points, but only in industries that are not highly exposed to AI automation. At the same time, we observe a negative and statistically significant interaction between the treatment indicator and industry-level AI exposure. This pattern indicates that employers in more AI-exposed sectors are less likely to respond favorably to AI-related qualifications.

By providing causal evidence on how employers value AI-related qualifications in entry-level recruitment, our paper contributes to two strands of the literature. Our first contribution is to the research on returns to AI-related skills and to skills more broadly. A substantial body of research examines the returns to skills, with a particular focus on their acquisition in education and schooling (for a review, see Acemoglu and Autor (2011)). Several seminal studies highlight the labor market value of specific skillsets in hiring and compensation, including cognitive ability (Deming & Kahn, 2018; Koedel & Tyhurst, 2012), social skills (Deming, 2017; Ham et al., 2025; Piopiunik et al., 2020), leadership skills (Kuhn & Weinberger, 2005) and other noncognitive skills (Bassi & Nansamba, 2022), past experiences such as entrepreneurship (Kacperczyk & Younkin, 2022), and personality traits such as conscientiousness and cooperation (Heinz & Schumacher, 2017; Wehner et al., 2022). An emerging strand of this literature focuses on digital skills, beginning with competencies related to information and communication technologies (ICT). Falck et al. (2021) estimate the returns to ICT skills using international survey data, employing the staggered rollout of broadband as an instrumental variable. They find that workers with ICT skills earn higher wages, primarily due to their selection into better paid occupations with a greater share of abstract tasks.

In the context of AI, Pouliakas et al. (2025) investigate the wage premium for AI-related skills among programmers in Europe using observational workforce data. They find a substantial return to AI-related skills that remains unexplained by standard human capital measures such as years of education and work experience. Moreover, two notable studies focus on AI-related skills in online labor platforms. Duch-Brown et al. (2022) document a 3.0%-3.2% wage premium for workers on AI projects, attributing it to strong demand for, and limited supply of, freelance AI experts. Stephany and Teutloff (2024) report a 21% wage premium for online freelancers with AI-related skills, which they attribute to a high degree of complementarity with other competencies.

Yet despite growing interest in AI-related skills, existing evidence remains largely observational and limited to technical domains of the labor market or platform-based freelance work. Our study contributes by providing causal evidence from a field experiment on how employers respond to complementary AI-related qualifications in entry-level hiring contexts where such skills are not explicitly required. To our knowledge, the only related experimental study is Blanco and López Bóo (2020), who examine the effect of listing ICT skills on interview callbacks, finding small positive effects of 1.1 to 3.6 percentage points. We extend this work by focusing specifically on complementary AI-related skills and testing their causal impact in a broad set of entry-level occupations that do not explicitly require such competencies. In doing so, we provide novel evidence on how emerging digital skills are evaluated by employers in general, rather than within AI-specialized or freelance occupations.

Our second contribution is to the literature on the role of AI in shaping evolving skill demand. Building on the seminal work of Autor et al. (2003), which conceptualize jobs as "bundles of tasks," recent research has explored how AI and automation influence the skills required of workers. These studies emphasize that new technologies, including AI, are more likely to reshape job roles by eliminating certain tasks and creating new ones rather than fully automating entire occupations (Acemoglu & Restrepo, 2018; Brynjolfsson et al., 2018).

A central research question to this literature is whether AI will complement or substitute existing worker skills, and how these dynamics will affect workers with different skill profiles. Frey and Osborne (2017), for instance, argue that AI and computerization will primarily substitute for low-skill, low-wage jobs, making low-skilled workers the most vulnerable to technological displacement. Some empirical studies support this view, showing that AI adoption can increase labor inequality by reducing demand for low-skilled workers and enhancing the position of high-skilled workers (Balsmeier & Woerter, 2019; Holm & Lorenz, 2022; Xie et al., 2021). In contrast, Webb (2020), using a task-based exposure measure, concludes that AI may primarily affect high-skilled, specialized occupations, marking a departure from earlier waves of automation driven by software and robotics. In the context of online labor platforms, recent evidence shows that freelancers in occupations highly exposed to AI-driven automation have experienced significant declines in both employment and earnings (Hui et al., 2024).

Recent studies also highlight how AI exposure influences hiring and skill demand at the firm level. Acemoglu et al. (2022) employ measures of "exposure to AI" at the firm level[2] to show that firms exposed to AI tend to reduce hiring for non-AI roles and change skill requirements even for roles not directly requiring AI qualifications. These findings suggest that AI technologies not only reshape the tasks within existing jobs but also alter the broader skill requirements of firms, even for roles not directly related to AI.

Building on this literature, we ask whether acquiring AI skills – when not central to one's core occupation – offers a labor market advantage. We explore how the returns to AI skills differ depending on the industry-level exposure to AI automation. We find that acquiring AI skills may be beneficial for job seekers in industries with lower exposure to AI, but these advantages tend to disappear in highly exposed sectors. This aligns with the broader view that the impact of new technologies depends in part on the match between workforce skills and

[2] This includes a measure at the occupation-level developed by Felten et al. (2018). We employ the updated version of this measure described in Felten et al. (2021). The latter measures AI exposure at the industry level and is derived from the occupation-level measure.

technological needs (Acemoglu & Restrepo, 2019), and underscores the importance of aligning retraining efforts with sector-specific patterns of technological adoption.

The rest of the paper is structured as follows. In Section 2, we describe the research methodology. Section 3 outlines the results. Section 4 discusses our findings and Section 5 concludes.

2. Research methodology

To understand the effect of AI-related qualifications on the success of job seekers, we conduct a large-scale field experiment employing the correspondence study methodology.[3] This approach, which involves sending fictitious but realistic résumés to real job postings and tracking employers' responses, is well established in empirical organization science (Levine et al., 2023). While originally developed to detect labor market discrimination (Bartoš et al., 2016; Bertrand & Mullainathan, 2004; Kline et al., 2022), it has also been applied to study how specific applicant attributes such as formal qualifications (Deming et al., 2016; Verhaest et al., 2018) or entrepreneurial experience (Kacperczyk & Younkin, 2022) influence hiring outcomes. This approach allows us to isolate the marginal effect of AI-related qualifications on interview callback rates, holding all other résumé attributes constant.

Between March and November 2021, we sent 1,185 applications to open vacancies identified on major UK online job platforms,[4] including Indeed.co.uk, Monster.co.uk, and Reed.co.uk. We restrict applications to entry-level positions requiring at most one year of professional experience, and exclude postings that demand rare or highly specialized skills

[3] The experiment is pre-registered in the AEA Registry with reference number <*The registration number will be added after the review*>. We obtained ethical approval from the German Association for Experimental Economic Research e.V. (<*The registration number will be added after the review*>).
[4] Unlike in other countries (e.g., Germany), in the UK labor market during the time when our study was conducted it was common to send job applications consisting only of résumés, without supplementary documents, such as transcripts, references, or certifications.

(e.g., proficiency in Mandarin). We further limit the sample to job functions that do not involve direct participation in AI development, such as software engineering roles. This restriction serves two purposes. First, it aligns with our objective of estimating the complementary value of AI-related skills in non-AI-focused occupations. Second, AI development roles typically demand verifiable evidence of technical capability such as links to GitHub repositories or published research, which we cannot credibly or ethically simulate in a field experiment. To avoid cross-contamination of treatment effects or detection risk, we submit only one application per company. This ensures a clean between-subjects experimental design, in which each employer is exposed to only one treatment condition.

Each identified job posting is randomly assigned to one of two experimental conditions: a "treatment group," which receives a résumé that includes additional AI-related qualifications and a "control group," which receives an otherwise identical résumé without mentioning such qualifications. Job postings are randomized within job function strata to ensure balanced treatment assignment across occupational categories. In total, we send 591 applications to the control group, and 594 to the treatment group. Detailed information on the application procedures is provided in the Online Appendix.

The applications are sent to roles across six distinct job functions, representing a broad cross-section of typical organizational activities, including Engineering; Finance and Accounting; HR; IT; Logistics and Supply Chain; and Marketing.[5] These job functions are

[5] The pre-registered estimated total sample size was 600 observations, assuming a baseline callback rate of 17%, an effect size of 9.4 percentage points, a power of 0.8, and a significance level of 0.05, and implied n = 200 applications for one job function. The baseline callback rate was determined through a small pilot. Initially, we pre-registered only three functions: Finance and Accounting, HR, and Marketing. The other three job functions (IT, Logistics and Supply Chain, and Engineering) were added subsequently, after we observed no significant overall treatment effect across specifications. The main purpose of collecting observations for the other job functions is to increase the representativeness of occupations as well as the sample size and to be able to detect smaller effects. While some of the heterogeneous effects discussed in sections 3.3 and 3.4 yield marginally significant results on the reduced (pre-registered) sample for some specifications, these results are not robust across models and outcomes. We report the analysis on the reduced sample in the Online Appendix.

selected based on recent research evidence indicating significant potential for the adoption and integration of AI technologies across a wide range of organizational activities (Alekseeva et al., 2021; Squicciarini & Nachtigall, 2021; Stapleton et al., 2021).

For each job function, we construct two résumés that are identical in all respects except for the inclusion of signals of AI-related qualifications. These signals are selected based on case studies and empirical data on the most frequently listed AI skills, as reported in Alekseeva et al. (2021), and complemented with use cases identified in industry reports (for a detailed description see Table 1 in the Online Appendix). Consistent with our definition of AI as a prediction-oriented technology (Agrawal et al., 2019), AI-related skills cover a broad range of tasks from classification, anomaly detection, risk forecasting and recommendations. In the treatment condition, signals of AI-related skills are added into five résumé sections, such as "[u]sing Deep Learning techniques for image based part classification" in Engineering or "IBM Watson" in HR. These signals of AI competencies are tailored to each job function. The control résumés are identical but do not mention these AI-related skills. In total, we create 12 unique résumés: one treatment and one control version for each of the six job functions. The distribution of applications by job function is reported in Table 1.

Table 1 – Sample size by job function

Job function	Total	Treatment (AI)	Control
Engineering	200	100	100
Finance	189	95	94
HR	196	99	97
IT	200	100	100
Logistics	200	100	100
Marketing	200	100	100
Total	1,185	594	591

Note: The table reports the number of applications submitted by job function and treatment status. "HR" refers to Human Resources, "IT" to Information Technology, "Finance" to Finance and Accounting, "Logistics" for Logistics and Supply Chain.

We focus on two measures of callback rates and use them as dependent variables in our analysis.[6] The first, *strict callback*, includes only interview invitations, whether via email or phone.[7] The second, *broad callback*, captures all positive employer responses, including interview invitations, follow-up requests, or invitations to apply for alternative positions. We track responses for 40 days following each application. To minimize any inconvenience for employers, we respond promptly to each callback with a brief and polite rejection.

In addition to employer callbacks, our dataset includes detailed information on the job postings and firm-level characteristics. For each vacancy to which we apply, we extract the full job posting text and systematically code the skills requested by the employer. We complement this with data on the company advertising the position, collected from company websites and official financial records available through the UK Companies House registry. These firm-level variables include industry sector, job location, number of employees, annual revenues, and year of establishment. Full documentation on these data sources and coding procedures is provided in the Online Appendix.

3. Results

3.1 Causal impact of AI-related qualifications on callback rates

Table 2 presents descriptive statistics on callback rates by treatment status and job function. Panel A reports statistics of strict callback rates, which include only explicit interview invitations, while Panel B presents statistics of broad callback rates, which additionally capture follow-up requests and other expressions of employer interest. Across all job functions, strict callback rates range from 3.17% to 11.50%, and broad callback rates range between 6.35% and

[6] These measures follow from the literature on correspondence experiments (S. Baert et al., 2015; Verhaest et al., 2018).
[7] Overall, 91.1% of callbacks were via email.

19.50%. Both measures are consistently lower for Finance compared to other job functions (*p*-values < 0.088), based on two-sided tests of proportions reported in Table 3 of the Online Appendix.

We test the average treatment effect of AI-related qualifications on callback rates across the full sample and find no statistically significant difference between treatment and control groups for either outcome measure: For strict callback rates the *p*-value is 0.408 (*test statistic* = -9), and for broad callback rates the *p*-value is 0.203 (*test statistic* = -17), based on a two-sided Fisher-Pitman permutation test for two independent samples with 200,000 replications, henceforth "FPP2S test."[8]

However, the lack of statistically significant overall effects masks heterogeneity across job functions. In both Marketing and Engineering, résumés listing AI-related qualifications receive higher callback rates compared to those in the control group. In Marketing, strict callback rates are 16.00% for AI résumés compared to 7.00% for the control group (*p*-value = 0.075, *test statistic* = -9, FPP2S test), while broad callback rates are 24.00% versus 12.00% (*p*-value = 0.043, *test statistic* = -12, FPP2S test). In Engineering, strict callback rates are 10.00% for AI résumés compared to 4.00% for the control group (*p*-value = 0.163, *test statistic* = -6, FPP2S test), while broad callback rates are 20.00% versus 8.00% (*p*-value = 0.024, *test statistic* = -12, FPP2S test).

[8] Our pre-registered analysis plan proposed a simple test of proportions, which is less conservative than the FPP2S test but delivers virtually the same results (reported in Table 4 in the Online Appendix).

Table 2 – Descriptives of callback rates by job function and treatment

Panel (A): Strict callback rates						
Job function	Overall	Treatment (AI)	Control	Relative frequency (AI/control)	p-values	Test statistic
Engineering	7.00%	10.00%	4.00%	2.50	0.163	-6
Finance	3.17%	3.16%	3.19%	0.99	1.000	0
HR	7.65%	9.09%	6.19%	1.47	0.622	-3
IT	7.50%	5.00%	10.00%	0.50	0.283	5
Logistics	11.00%	9.00%	13.00%	0.69	0.496	4
Marketing	11.50%	16.00%	7.00%	2.29	0.075*	-9
All functions	8.02%	8.75%	7.28%	1.20	0.408	-9
Panel (B): Broad callback rates						
Job function	Overall	Treatment (AI)	Control	Relative frequency (AI/control)	p-values	Test statistic
Engineering	14.00%	20.00%	8.00%	2.50	0.024**	-12
Finance	6.35%	6.32%	6.38%	0.99	1.000	0
HR	13.78%	12.12%	15.46%	0.78	0.637	3
IT	16.50%	13.00%	20.00%	0.65	0.254	7
Logistics	19.50%	21.00%	18.00%	1.17	0.725	-3
Marketing	18.00%	24.00%	12.00%	2.00	0.043**	-12
All functions	14.77%	16.16%	13.37%	1.21	0.203	-17

Note: *Strict callback rates* include only interview invitations; *Broad callback rates* include all positive responses. p-values and test statistics from two-sided Fisher-Pitman permutation tests for two independent samples with 200,000 replications comparing rates between the treatment and control groups. The total number of observations is 1,185. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Tables 3 and 4 present the estimation results for strict and broad callbacks respectively for the following model:

$$Y_i = \alpha + \beta T_i + \mathbf{\Gamma}'\mathbf{\Phi}_i + \boldsymbol{\lambda}'(T_i \cdot \mathbf{\Phi}_i) + \mathbf{\Theta}'\mathbf{X}_i + u_i \qquad (1)$$

Where Y_i is a binary outcome equal to one if application i receives a callback and zero otherwise; T_i is a treatment indicator equal to one if the application i included the AI résumé, and zero otherwise; Φ_i is a 5×1 vector of five job function dummies; X_i is a vector of employer-level controls, including industry sector, establishment age, job location (coded as a binary variable equal to one if the job was based in London or the South-East, and zero otherwise), and firm size (measured by the number of employees).[9] The term u_i denotes the idiosyncratic error.

The coefficient β captures the estimated average treatment effect of AI-related qualifications across all job functions. The 1×5 vector of interaction coefficients λ' captures heterogeneity in the treatment effect by job function. These coefficients – β and λ' – constitute the primary focus of our analysis. We estimate equation (1) both with and without the inclusion of employer controls. As indicated in our pre-registration, we employ both a linear probability model (LPM) and a Probit model (reporting average marginal effects) with robust standard errors.

Across all model specifications, the estimated average treatment effect of listing AI-related qualifications is not statistically significant, with p-values ranging from 0.346 to 0.990. However, the interaction between the treatment indicator and the Marketing job function dummy is positive and marginally significant at the 10% level in all LPM specifications: For strict callbacks, the coefficient size is around 0.09 and the p-values are 0.082 and 0.072 (Table 3, Models (5) without employer-level controls and (7) with employer-level controls, respectively). For broad callbacks the coefficient size is around 0.12 and the p-values are 0.064 in Model (5) (without employer-level controls), and 0.087 in Model (7) (with employer-level controls). These estimates suggest that including AI-related qualifications increases the

[9] While the number of employees was not explicitly specified as a control variable in our pre-analysis plan, excluding it does not materially alter the direction or magnitude of our results. The pre-analysis plan did include 'wage offered' as a proposed control; however, this variable is missing for over 37% of our sample. Including it as a control thus substantially reduces the effective sample size. Nonetheless, when wage is included, the estimated effects remain qualitatively similar and our core conclusions are unaffected. Complete results are available upon request.

probability of receiving an interview invitation for marketing roles by approximately 9 percentage points and a broader callback by 12 percentage points. Similarly, the interaction between the treatment dummy and the Engineering job function dummy in the LPM models is positive and statistically significant, but only for broad callbacks. AI-related qualifications increase the probability of a broad callback by at least 11 percentage points, with a *p*-value of 0.047 (Model (5) without employer-level controls), and 0.081 (Model (7) with controls).

The sample size of 1,185 observations provides us with 80% power to detect a treatment effect size of at least 4.8 percentage points for the strict callback rate and 6.0 percentage points for the broad callback rates, with a two-sided test of proportions and means of 7.28% and 13.37% (i.e., the control group values) for the strict and broad callbacks respectively.[10] While our estimates do not reveal statistically significant average treatment effects, we cannot rule out the existence of smaller positive effects that fall below these minimum detectable effect sizes.

[10] We conduct the achieved power analysis using Stata's *power* command (see full details in the Online Appendix).

Table 3 – Treatment effect on strict callbacks

	(1) LPM	(2) Probit	(3) LPM	(4) Probit	(5) LPM	(6) Probit	(7) LPM	(8) Probit
Treatment (AI)	0.015	0.015	0.015	0.015	-0.000	-0.001	-0.000	0.002
	(0.016)	(0.016)	(0.016)	(0.016)	(0.026)	(0.052)	(0.027)	(0.052)
Treatment (AI) × Engineering					0.060	0.068	0.059	0.063
					(0.044)	(0.066)	(0.047)	(0.067)
Treatment (AI) × HR					0.029	0.030	0.005	0.003
					(0.046)	(0.064)	(0.047)	(0.066)
Treatment (AI) × IT					-0.050	-0.052	-0.044	-0.049
					(0.045)	(0.065)	(0.048)	(0.066)
Treatment (AI) × Logistics					-0.040	-0.030	-0.040	-0.030
					(0.052)	(0.062)	(0.055)	(0.062)
Treatment (AI) × Marketing					0.090^{*}	0.070	0.095^{*}	0.077
					(0.052)	(0.062)	(0.053)	(0.062)
Constant	0.073^{***}	0.073^{***}	0.024	0.028^{**}	0.032^{*}	0.032^{*}	0.051^{*}	0.035
	(0.011)	(0.011)	(0.015)	(0.012)	(0.018)	(0.018)	(0.030)	(0.022)
Job function dummies	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Controls	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Observations	1185	1185	1185	1185	1185	1185	1141	1121

Note: Linear probability model (LPM) and Probit model results (reporting average marginal effects and constants for the predicted probability of the outcome when all exogenous variables are equal to zero); robust standard errors are reported in parentheses. The dependent variable *Strict callback* is equal to one if the application received a callback for an interview, and zero otherwise. *Treatment (AI)* is equal to one if the résumé includes AI-related skills and zero otherwise. *Job function dummies* are dummies for each of the job functions (Engineering, HR, IT, Logistics, and Marketing), with Finance being the reference category. *Controls* include industry fixed-effects, firm size (number of employees), company age, and a location dummy if the job is based in London or the South-East of England. The pre-registration also indicated *wage offered* as a control but is excluded as it is missing for over 37% of the sample; the exclusion does not qualitatively affect the results. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4 – Treatment effect on broad callbacks

	(1) LPM	(2) Probit	(3) LPM	(4) Probit	(5) LPM	(6) Probit	(7) LPM	(8) Probit
Treatment (AI)	0.028	0.028	0.028	0.028	-0.001	-0.001	0.006	0.011
	(0.021)	(0.021)	(0.021)	(0.020)	(0.036)	(0.064)	(0.038)	(0.064)
Treatment (AI) × Engineering					0.121^{**}	0.127	0.111^{*}	0.108
					(0.061)	(0.082)	(0.064)	(0.083)
Treatment (AI) × HR					-0.033	-0.033	-0.072	-0.077
					(0.061)	(0.081)	(0.064)	(0.083)
Treatment (AI) × IT					-0.069	-0.062	-0.062	-0.061
					(0.064)	(0.080)	(0.068)	(0.081)
Treatment (AI) × Logistics					0.031	0.026	0.006	-0.003
					(0.067)	(0.078)	(0.072)	(0.079)
Treatment (AI) × Marketing					0.121^{*}	0.106	0.114^{*}	0.096
					(0.065)	(0.079)	(0.067)	(0.079)
Constant	0.134^{***}	0.134^{***}	0.049^{**}	0.056^{***}	0.064^{**}	0.064^{**}	0.138^{***}	0.105^{**}
	(0.014)	(0.014)	(0.021)	(0.017)	(0.025)	(0.025)	(0.044)	(0.045)
Job function dummies	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Controls	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Observations	1185	1185	1185	1185	1185	1185	1141	1126

Note: Linear probability model (LPM) and Probit model results (reporting average marginal effects and constants for the predicted probability of the outcome when all exogenous variables are equal to zero); robust standard errors are reported in parentheses. The dependent variable *Broad callback* is equal to one if the application received any positive response from the employer, and zero otherwise. *Treatment* is equal to one if the résumé includes AI-related skills and zero otherwise. *Job function dummies* are dummies for each of the job functions (Engineering, HR, IT, Logistics, and Marketing), with Finance being the reference category. *Controls* include industry fixed-effects, firm size (number of employees), company age, and a location dummy if the job is based in London or the South-East of England. The pre-registration also indicated *wage offered* as a control but is excluded as it is missing for over 37% of the sample; the exclusion does not qualitatively affect the results. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Overall, our findings offer a nuanced view on the labor market returns to AI-related qualifications. While entry-level job applications listing complementary AI-related qualifications receive significantly higher callback rates for vacancies in Marketing and Engineering, we detect no significant effects in other job functions or for the treatment on average. This pattern is somewhat surprising in light of existing empirical evidence documenting rising demand for AI-related skills across a broad range of occupations. Although Marketing and Engineering are widely recognized as functions with high potential for AI integration, similar potential has also been identified in functions such as Logistics and Supply Chain, as well as IT – with HR and Finance and Accounting not far behind (Chui et al., 2018). Given this widespread potential for applying AI technology, one might have expected a more uniform positive response to AI qualifications across job functions.

In the following section, we explore three potential explanations for our findings: (i) the design features of résumés used in our correspondence study, (ii) the characteristics of the job postings, and (iii) the industry level exposure to AI.

3.2 Manipulation check of résumés and AI-related qualifications

To assess whether our findings are driven by design-related artifacts, we systematically evaluate key features of the résumés used in the experiment.[11] In particular, we examine whether: (i) the signals of AI skills included in the treatment résumés are easily noticeable; (ii) these qualifications are perceived as relevant for the targeted job roles; (iii) the treatment résumés signal higher general competencies; and (iv) the treatment résumés are perceived as indicating overqualification.

To investigate these dimensions, we conduct an online survey with 771 professionals who report having hiring experience and domain expertise in at least one of the six job functions examined in our study. Participants are recruited via the online platform Prolific.com.[12]

[11] The analyses presented in this section (as well as Sections 3.3 and 3.4) were not part of the original pre-analysis plan. They were conducted after observing a null effect in the correspondence study, a finding that went against our expectations. They should therefore be interpreted as exploratory.
[12] The survey was conducted in 2022.

The survey comprises four parts. In Part I, participants are randomly assigned to view either the treatment or control résumé corresponding to their area of professional expertise (e.g. the Marketing résumés are presented to participants with experience in marketing). After viewing a résumé, participants assess the applicant's skills and the likelihood of receiving an interview invitation. The randomization of résumés in Part I enables a between-subjects comparison. In Part II, participants are shown the alternative version of the résumé and respond to the same set of questions, allowing for within-subject comparison. Part III elicits assessment of the perceived authenticity of both résumés and relevance of listed qualifications – both AI-related and non-AI – for an entry-level position. Part IV collects standard socio-demographic data from participants, including age, gender, and current employment characteristics. The full survey is available in the Online Appendix.

First, we test whether participants notice the presence of AI-related qualifications in the résumés. Attention to specific résumé characteristics might be limited, and if survey participants fail to notice the AI-related content, it is plausible that recruiters in the correspondence study also overlook it. This mechanism could help account for the null average treatment effect reported above.

To test this mechanism, we ask participants in Part III of the survey to recall which of the two résumés they viewed contained AI-related qualifications. Importantly, participants cannot revisit the résumés during Part III. The correct response is incentivized with a £0.10 bonus. The results indicate that 95.46% of respondents can correctly identify and recall the résumé containing the AI-related qualifications. This recall rate is significantly higher than the 50% benchmark expected under random guessing (p-value < 0.001, two-sided Binomial probability test). The recall rates are above 91.80% for all job functions (see Table 5 in the Online Appendix).

Second, we examine whether AI-related qualifications influence the perception of résumés across a broader set of skill dimensions. As predicted by signaling theory, workers may invest in additional education or training not solely for its direct productivity benefits, but because such investments serve as costly signals of otherwise unobserved ability, allowing employers to differentiate between high- and low-ability applicants (Arrow, 1973; Spence, 1973). If employers interpret AI-related qualifications as such a signal, we would expect them to prefer candidates who list such qualifications, even in jobs where AI skills are not explicitly required. At the same time, these qualifications may serve as a negative signal if employers associate AI-related qualifications with weaker social or interpersonal skills, potentially explaining why AI-related qualifications do not translate into positive effects on callback rates.

To test the signaling mechanism, participants are asked to evaluate each résumé on five distinct skill categories: *AI-related; advanced digital; basic digital; cognitive;* and *social*. For each category, participants respond to the question "*How would you assess the candidate's [category] skills?*" using a five-point Likert scale ranging from one ("Very low") to five ("Very high"). These evaluations are collected for both the treatment and control résumés in Parts I and II of the survey. A brief description of each skill category, as provided to participants, is listed in Table 6 of the Online Appendix.

Since the only difference between the treatment and control résumés lies in the inclusion of AI-related qualifications, any observed difference in skill assessment scores can be attributed to these qualifications. Using a non-parametric test, we find that respondents rated treatment résumés significantly higher across nearly all skill dimensions (ranging from 0.25 to 1.64 standard deviations, see Table 5). The sole exception is the *social skills* category, where control résumés receive 0.08 standard deviations higher average scores (mean = 3.808, SD = 0.747) compared to treatment résumés (mean = 3.744, SD = 0.780). While this difference is statistically significant (p-value = 0.009, *test statistic* = -49, two-sided Fisher-Pitman permutation test for paired replicates, henceforth FPPR test), the absolute magnitude is small.

Although we randomize the résumé order to mitigate potential order effects in our survey, the within-subjects survey design may still introduce systematic biases in how participants evaluate résumés. To account for such potential biases, we estimate a regression model in which the dependent variable is the standardized score assigned to a given skill category. The independent variables include: (i) a dummy equal to one if the résumé contains the AI-related qualifications and zero otherwise (*AI résumé*); (ii) a dummy equal to one if the résumé was presented first and zero if it was seen second (*First résumé*); and (iii) an interaction between these two indicators. Table 6 reports the results.

The estimated regression coefficients largely align with the findings from the non-parametric analysis. The coefficients for the *AI résumé* dummy are positive and statistically significant for all skill categories except *social skills*, indicating that résumés listing AI-related qualifications are rated more highly than control résumés when presented second (by 1.2, 0.8, and 0.5 standard deviations respectively). Moreover, the sum of the *AI résumé* coefficient and the interaction term is also positive and statistically significant for *AI-related*, *advanced digital* and *basic digital* skills, suggesting that the treatment résumé receives higher ratings (by 1.34, 0.44, 0.24 standard deviations, respectively) even when presented first. However, the statistical significance of the *First résumé* dummy and its interaction with *AI résumé* dummy in some specifications indicates that presentation order may still influence skill assessment for certain categories. Nevertheless, the résumés with AI-related qualifications receive significantly higher scores for skills in the *AI-related*, *advanced digital* and *basic digital skills* categories regardless of their presentation order. The inclusion of AI-related qualifications results in significantly higher ratings for *cognitive skills* only when the participants view them in Part II

(0.39-0.41 standard deviations). Finally, when presented first, the treatment résumés are rated by 0.14-0.15 standard deviations lower on the *social skills* dimension. This effect is not statistically significant when respondents view the treatment résumés in Part II. Overall, we interpret these results as confirming that participants perceive the treatment résumés as having higher AI-related, basic digital and advanced digital skills, with suggestive evidence of lower social skills and modest improvements in cognitive skills that appear sensitive to presentation order.

Table 5 – Perceived candidates' skillfulness (expert survey)

Skills categories	Treatment (AI)	Control	*p*-value	Test statistic
AI-related	4.125 (0.822)	2.523 (1.106)	0.000***	1235
Advanced digital	4.014 (0.767)	3.482 (0.928)	0.000***	401
Basic digital	4.454 (0.664)	4.187 (0.740)	0.000***	206
Cognitive	4.014 (0.663)	3.844 (0.726)	0.000***	131
Social	3.744 (0.780)	3.808 (0.747)	0.009***	-49

Note: The table reports the mean response to the question "*How would you assess the candidate's [category] skills?*" for the corresponding résumé × type skills on a scale from one "Very low" to five "Very high." Standard deviations are reported in parentheses. Each respondent (N = 771) rates two résumés. *p*-values and test statistics from a two-sided Fisher-Pitman permutation test of paired replicates with 200,000 runs testing the difference in ratings of the treatment and control résumés. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Third, we examine whether participants perceive the AI-related qualifications as relevant to the targeted job roles. A lack of perceived relevance could explain why résumés containing AI skills do not receive higher callback rates: Even if recruiters noticed the AI skills, they may have judged them as unrelated to the core responsibilities of the position, and therefore uninformative in assessing candidate suitability.

To test this mechanism, Part III of the survey asks participants to rate their agreement with the statement: "The AI skills (This included skills such as "…" and a bullet point stating the candidate had experience "…")[13] would significantly improve a candidate's ability to execute the [*job function]* assistant role". Responses are recorded on a five-point Likert scale ranging from one ("Strongly disagree") to five ("Strongly agree"). Overall, 57.07% of respondents selected "Agree" or "Strongly agree," while only 19.33% selected "Disagree" or "Strongly disagree." The difference is statistically significant (p-value < 0.001, *z-score* = 26.542, two-sided one-sample test of proportions). The pattern holds consistently across all job functions, with all p-values below 0.001 (two-sided one sample tests of proportions; see Table 7 in the Online Appendix).

Fourth, we examine whether candidates with résumés listing AI-related qualifications are perceived as overqualified for entry-level positions, even when the skills themselves are viewed as relevant to the job. While firms are generally assumed to prefer more qualified candidates, recent research highlights frictions that may arise when candidates exceed job requirements. Employers may form adverse assessments in such cases, anticipating higher wage demands, lower job satisfaction, or an increased likelihood of early turnover (Korpi & Tåhlin, 2009; McGuinness & Wooden, 2009). A recent set of empirical studies, however, show that perceived overqualification either does not affect, or even has a positive effect on the probability of being invited for an interview (Deming et al., 2016; Van Beek et al., 1997; Verhaest et al., 2018). The only exception is Humburg and Van Der Velden (2015), who find that candidates holding a PhD were less likely to be invited to interview for a junior position compared to those with bachelor's or master's degrees. These findings suggest that in practice,

[13] Instead of the ellipsis participants are shown the specific AI-related qualifications and experiences listed in the résumé that they saw earlier. In the survey on Marketing résumés, for example, these are replaced with entries such as "IBM Watson" and "Using AI / Machine Learning algorithms for personalised customer recommendations (...)". For the full list of qualifications by job function please see Table 1 in the Online Appendix.

firms' preferences for qualifications may be increasing but not strictly monotonic – they may plateau or even reverse when additional qualifications are expected to introduce new costs.

To test this mechanism, we ask participants *"In your opinion, how well do the candidate's qualification and experience fit an entry-level [job function] position?*". Responses are recorded on a five-point Likert scale ranging from one ("Underqualified") to five ("Overqualified"). As in earlier parts of the survey, participants answer this question for each résumé.

The comparison of mean evaluations provides some evidence that résumés listing AI-related qualifications are perceived as more overqualified (mean = 3.838, SD = 0.763) compared to control résumés (mean = 3.790, SD = 0.749). This difference is marginally statistically significant (p-value = 0.096, *test statistic* = 37.0, two-sided FPPR test), but economically rather small in magnitude, corresponding to a change of 0.04 standard deviations. Table 8 in the Online Appendix reports the differences in comparison of means by job function, showing no statistically significant difference at the job function level.

Table 7 reports the regression analysis testing the robustness of perceived qualification effect to résumé presentation order. The findings indicate that presentation order significantly affects the perception of AI-related overqualification. When shown in Part II of the survey, the AI résumés are significantly more likely to be rated as overqualified than control résumés, with effect size ranging from 0.58 to 0.60 standard deviations (p-values < 0.001). In contrast, when presented first, AI résumés are perceived as 0.05 to 0.47 standard deviations less overqualified than control résumés (p-values < 0.001, two-sided Wald tests). These results suggest that the perceived overqualification associated with AI credentials is sensitive to whether the résumé is evaluated first or second.

Table 6 – AI qualifications and skill perception (expert survey)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		AI-related skills			Advanced digital skills			Basic digital skills	
AI résumé	1.194***	1.201***	1.197***	0.753***	0.757***	0.758***	0.509***	0.505***	0.489***
	(0.058)	(0.057)	(0.058)	(0.067)	(0.066)	(0.068)	(0.073)	(0.073)	(0.075)
First résumé	-0.079	-0.072	-0.076	0.176**	0.180**	0.181**	0.262***	0.258***	0.242***
	(0.063)	(0.062)	(0.063)	(0.075)	(0.072)	(0.074)	(0.074)	(0.073)	(0.074)
AI résumé × First résumé	0.153*	0.138*	0.146*	-0.313***	-0.321***	-0.325***	-0.271**	-0.262**	-0.231*
	(0.083)	(0.081)	(0.083)	(0.113)	(0.110)	(0.113)	(0.122)	(0.121)	(0.125)
Constant	-0.595***	-0.516***	-0.831***	-0.386***	-0.164**	-0.213	-0.318***	-0.354***	-1.548***
	(0.047)	(0.061)	(0.260)	(0.052)	(0.080)	(0.344)	(0.055)	(0.091)	(0.392)
Job function dummies	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Socio-demographic controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Treatment (AI) + AI × First = 0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
First + Treatment (AI) × First = 0	0.116	0.160	0.140	0.027	0.023	0.022	0.894	0.946	0.881
Observations	1,542	1,542	1,538	1,542	1,542	1,538	1,542	1,542	1,538

Note: Linear regression (OLS) with robust standard errors clustered at the participant level and reported in parentheses. The dependent variables are: Standardized assessment scores for each of the five skill dimensions (AI-related, advanced digital, basic digital, cognitive, social). *AI résumé* is a dummy equal to one if the résumé assessed contained AI qualifications. *First résumé* is a dummy equal to one if the résumé being assessed was seen first, and zero otherwise. In the line *Treatment (AI) + AI × First = 0* we report the *p*-values for a two-sided Wald test. The line *First + Treatment (AI) × First* = 0 presents the *p*-values for a two-sided Wald test. *Job function dummies* are dummies for each of the six occupations (Engineering, HR, IT, Logistics and Marketing, with Finance as the reference category). *Controls* include participant's gender, age, highest educational degree achieved, employment status, employment type, industry of employment, years of work experience, hiring experience, general knowledge of AI and general work experience with AI. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6 (continued) – AI qualifications and skill perception (expert survey)

	(10)	(11)	(12)	(13)	(14)	(15)
	Cognitive skills			Social skills		
AI résumé	0.411***	0.412***	0.385***	0.064	0.067	0.023
	(0.071)	(0.070)	(0.071)	(0.072)	(0.071)	(0.073)
First résumé	0.207***	0.208***	0.181**	0.162**	0.165**	0.122*
	(0.074)	(0.072)	(0.074)	(0.070)	(0.070)	(0.070)
AI résumé × First résumé	-0.337***	-0.340***	-0.292**	-0.294**	-0.300**	-0.206
	(0.123)	(0.120)	(0.123)	(0.129)	(0.128)	(0.130)
Constant	-0.225***	-0.088	-1.962***	-0.039	-0.019	-1.707***
	(0.053)	(0.084)	(0.368)	(0.053)	(0.091)	(0.387)
Job function dummies	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Socio-demographic controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Treatment (AI) + Treatment (AI) × First = 0	0.303	0.302	0.193	0.001	0.001	0.011
First + Treatment (AI) × First = 0	0.057	0.053	0.112	0.074	0.067	0.257
Observations	1,542	1,542	1,538	1,542	1,542	1,538

Note: Linear regression (OLS) with robust standard errors clustered at the participant level and reported in parentheses. The dependent variables are: Standardized assessment scores for each of the five skill dimensions (AI-related, advanced digital, basic digital, cognitive, social). *AI résumé* is a dummy equal to one if the résumé assessed contained AI qualifications. *First résumé* is a dummy equal to one if the résumé being assessed was seen first, and zero otherwise. In the line *Treatment (AI) + Treatment (AI) × First = 0* we report the *p*-values for a two-sided Wald test. The line *First + Treatment (AI) × First* = 0 presents the *p*-values for a two-sided Wald test. *Job function dummies* are dummies for each of the six occupations (Engineering, HR, IT, Logistics and Marketing, with Finance as the reference category). *Controls* include participant's gender, age, highest educational degree achieved, employment status, employment type, industry of employment, years of work experience, hiring experience, general knowledge of AI and general work experience with AI. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7 – AI qualifications and perceptions of overqualification and employment chances (expert survey)

	(1)	(2)	(3)	(4)	(5)	(6)
	Overqualification			*Ln(*expected callbacks)		
AI résumé	0.596***	0.592***	0.583***	0.254***	0.253***	0.244***
	(0.075)	(0.074)	(0.075)	(0.045)	(0.044)	(0.044)
First résumé	0.733***	0.729***	0.720***	0.118**	0.118**	0.108**
	(0.066)	(0.066)	(0.068)	(0.047)	(0.047)	(0.047)
AI résumé × First résumé	-1.065***	-1.058***	-1.028***	-0.292***	-0.292***	-0.271***
	(0.118)	(0.116)	(0.118)	(0.087)	(0.087)	(0.086)
Constant	-0.398***	-0.606***	0.359	3.139***	3.144***	1.849***
	(0.053)	(0.083)	(0.340)	(0.034)	(0.061)	(0.272)
Job function dummies	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Socio-demographic controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Treatment (AI) + Treatment (AI) × First = 0	0.000	0.000	0.000	0.426	0.424	0.575
First + Treatment (AI) × First = 0	0.000	0.000	0.000	0.000	0.000	0.000
Observations	1,542	1,542	1,538	1,536	1,536	1,532

Note: Linear probability model (LPM) clustered at the participant level. Standard errors reported in parentheses. The dependent variables are: *Overqualification* (whether the participant estimated the candidate to be under- or over-qualified), standardized; and *Ln(expected callbacks)* as the natural log of how many times out of 50 the participant estimated that the résumé was invited for an interview in the correspondence study; *AI résumé* is a dummy equal to one if the résumé assessed contained AI qualifications. *First résumé* is a dummy equal to one if the résumé being assessed was seen first, and zero otherwise. The line *Treatment* (*AI)* + *Treatment (AI)* × *First* = 0 presents the *p*-values for a Wald test. The line *First* + *Treatment (AI)* × *First* presents the *p*-values for a two-sided Wald test. *Job function dummies* are dummies for each of the six occupations (Engineering, HR, IT, Logistics and Marketing, with Finance as the reference category). *Controls* include participant's gender, age, highest degree achieved, employment status, employment type, industry of employment, years of work experience, hiring experience, general knowledge of AI and general work experience with AI. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fifth, we examine whether participants expect candidates with AI-related qualifications to receive more interview invitations than those without such qualifications. To test this, participants in Parts I and II of the survey are informed that the résumé they just viewed was submitted for an application to a real job posting. They are then asked to estimate how many times, out of 50 submissions, the application received an interview invitation. Responses that deviate by fewer than three points from the correct answer are incentivized with a £0.10 bonus.

On average, respondents estimate that résumés listing AI-related qualifications receive 2.76 more interview invitations than control résumés. This difference is statistically significant (p-value = 0.000, *test statistic* = 2124, two-sided FPPR test). The pattern holds within each job function: respondents consistently estimate more callbacks for AI résumés, with all differences statistically significant (all p-values < 0.074, two-sided FPPR test; see Table 9 in the Online Appendix for more details).

While respondents, on average, expect AI résumés to receive more interview invitations, this perceived advantage may be influenced by the order in which résumés are presented. To test for such order effects, we estimate the regression models reported in Table 7 (Models (4) – (6)). The results indicate that the higher expected callbacks rates for AI résumés are sensitive to presentation order. Although the coefficients of the *AI résumé* dummy suggest a 27.6% to 28.9% increase in expected callbacks, the interaction term with the *First résumé* dummy is negative, of similar magnitude, and statistically significant. As a result, the combined effect, capturing the effect of the AI-related qualifications in résumés presented Part I of the survey, is not statistically different from zero (all p-values > 0.424, two-sided Wald tests that the combined effect is zero). These findings suggest that, in Part I of the survey, respondents do not expect AI résumés to receive more interview invitations than control résumés. However, in Part II of the survey, control résumés presented after AI résumés are expected to receive lower callbacks than AI résumés presented after control résumés.

To sum up, we explore five potential mechanisms that could explain the absence of effects of AI-related qualifications on callbacks observed in our correspondence study. First, we verify that the treatment manipulation is salient: nearly all participants correctly identify which résumé contains AI-related skills, ruling out that AI qualifications are not visible to evaluators as a likely explanation. Second, we examine skill inferences and find that AI résumés are rated significantly higher in technical skill domains but slightly lower in social skills, particularly when shown first – indicating that AI credentials may signal a trade-off in perceived competencies. Third, we show that a majority of participants perceive the AI qualifications as relevant to the entry-level job roles, suggesting that a lack of perceived applicability does not drive the null result. Fourth, we test whether AI résumés are perceived as overqualified. While they are rated as more overqualified when shown second, this effect does not exist when they are presented first, revealing a strong sensitivity to presentation order. Fifth, we find that participants expect AI résumés to receive more interview invitations, but again only when these résumés are viewed second. Together, these findings highlight the importance of order effects and suggest that perceptions of AI qualifications are context-dependent, shaped not only by résumé content but also by the sequence in which it is evaluated.

3.3 Heterogeneous treatment effects by vacancy and job-posting type

The objective of this section is to examine how employer responses to AI-related qualifications correlate with features of job postings. This analysis tests whether the overall null effect masks meaningful heterogeneity in callback rates conditional on vacancy characteristics. Specifically, we assess whether the likelihood of receiving a callback in response to AI qualifications varies with attributes of the job postings to which applications were submitted. To this end, we analyze the variation in the textual content of the job advertisements in our sample.

We begin by extracting the skills required in each job posting using the Lightcast (formerly Burning Glass Technologies) open API.[14] This process generates a structured list of required skills for each vacancy, based on a detailed and widely validated taxonomy (Acemoglu et al., 2022; Hershbein & Kahn, 2018). For each job posting, we calculate the total

[14] These data could not be collected for 10 advertisements overall (less than 1% of the sample). The methodology to obtain the skills is described in detail in Burning Glass Technologies (2019).

number of skills and standardize this measure across the sample. The median job posting in our sample requires 18 skills with a standard deviation of 10.614 (see Figures 1 and 2 in the Online Appendix for the overall and job-function-specific distributions of this variable, respectively). We incorporate the resulting variable as a control in our main regression specification and interact it with the treatment indicator. The results, reported in Panel A in Table 8, indicate that neither the treatment effect nor its interaction with the number of required skills is statistically significant (all p-values > 0.176). This suggests that the likelihood of receiving a callback in response to AI-related qualifications does not vary with the number of skills required in the job posting.

Next, we map the list of extracted skills from each job posting to the European Skills, Competences, Qualifications and Occupations (ESCO) taxonomy using an open API.[15] The same procedure is then undertaken with the skills extracted from our résumés. This allows us to construct a measure of alignment between the qualifications listed in the résumés and those required by the job postings. Specifically, we define a match index for each vacancy i:

$$Match_i = \frac{Number\ of\ skills\ overlapping\ in\ the\ résumé\ and\ the\ job\ ad}{Number\ of\ skills\ mentioned\ in\ the\ job\ ad}$$

This results in a variable that varies between zero and one; the median job posting has a match value of 0.059 with a standard deviation of 0.071 (see Figures 1 and 3 of the Online Appendix for the overall and job-function-specific distributions, respectively). We include the standardized match index and its interaction with the treatment indicator in the main regression specifications. The results, reported in Panel B of Table 8, indicate that neither the main treatment effect nor its interaction with the match index is statistically significant (all p-values

[15] Nesta's Skills Extractor Library (2022), available at https://nestauk.github.io/ojd_daps_skills/build/html/about.html

> 0.131). This suggests that the null effect of AI-related qualifications on callback rates is not driven by variations in the degree of alignment between résumé skills and those specified in the job posting.

Finally, we examine whether the level of expertise implied by the job postings can explain our results. Recent literature has explored whether AI technologies are more likely to automate expert tasks rather than routine ones (Acemoglu et al., 2022; Acemoglu & Restrepo, 2018; Brynjolfsson et al., 2018). To test whether the treatment effect varies with the level of required expertise, we construct a text-based measure of expertise following Autor (2024), who argues that tasks demanding higher expertise tend to be described using less common vocabulary. Specifically, we apply the Dale-Chall Complexity Index, which captures the proportion of words in a text that are not included in a list of 3,000 commonly understood words – originally compiled by Dale and Chall (1948) to reflect vocabulary easily understood by fourth-grade children. The measure is calculated as follows:

$$DCC \equiv 1 - \frac{N_{dc}}{N_{words}}$$

where N_{dc} is the number of words mentioned in the text that appear on the Dale-Chall list and N_{words} is the total number of words in the text. Values closer to one denote higher required expertise, while values closer to zero indicate lower required expertise.

We compute this measure using the official ESCO skills descriptions of the skills mentioned in each job posting, rather than the job descriptions themselves. This approach allows us to capture the *normalized required expertise,* while avoiding firm-level stylistic differences in job advertisement language. The resulting variable has a median value of 0.698 with a standard deviation of 0.075 (the overall and job-function-specific distributions are presented in Figures 1 and 4 on the Online Appendix, respectively).

We incorporate the standardized expertise measure into our main regression specification, along with an interaction term with the treatment indicator. The results, presented in Panel C of Table 8, show that neither the treatment effect nor its interaction with the expertise index is statistically significant in any specification (all *p*-values > 0.155). This finding suggests that the level of expertise required for a vacancy does not systematically affect the likelihood of receiving a callback in response to AI-related qualifications.

Finally, we manually verify whether employers request AI-related skills by systematically searching each job posting.[16] We find only two such job postings (in Logistics and IT, both in the control condition). Including a dummy to control for these two job postings does not change any of our results.

In sum, the analysis presented in this section provides no evidence that the null effect of AI-related qualifications on callback rates varies with job posting characteristics. Across a range of vacancy-level attributes – including the number of required skills, the degree of skill match between the résumé and the job posting, and the level of required expertise – AI-related qualifications have no effect on the likelihood of receiving a callback. These findings suggest that the limited employer response to AI qualifications observed in our correspondence study is not confined to specific types of vacancies but rather reflects a broader pattern.

[16] We search the job postings for the words 'AI', 'Artificial Intelligence', 'ML', 'Machine Learning', and all the skills included in our treatment résumés (see Online Appendix for details). Results available upon request.

Table 8 – Treatment effect and job posting characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Number of skills required						
	Strict callback			Broad callback		
Treatment (AI)	0.015	0.015	0.014	0.028	0.028	0.023
	(0.016)	(0.016)	(0.017)	(0.021)	(0.021)	(0.022)
Number of skills required	0.005	0.010	0.007	0.018	0.031*	0.028
	(0.008)	(0.012)	(0.012)	(0.011)	(0.017)	(0.018)
Treatment (AI) × Number of skills required		-0.009	-0.014		-0.023	-0.028
		(0.017)	(0.018)		(0.023)	(0.024)
Constant	0.073***	0.073***	0.103***	0.135***	0.135***	0.242***
	(0.011)	(0.011)	(0.033)	(0.014)	(0.014)	(0.046)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	1,175	1,175	1,132	1,175	1,175	1,132

	(7)	(8)	(9)	(10)	(11)	(12)
Panel B: Match in skills required						
	Strict callback			Broad callback		
Treatment (AI)	0.016	0.016	0.014	0.032	0.031	0.027
	(0.016)	(0.016)	(0.017)	(0.021)	(0.021)	(0.022)
Match	-0.005	0.005	0.002	-0.015*	-0.008	-0.013
	(0.007)	(0.013)	(0.015)	(0.009)	(0.016)	(0.019)
Treatment (AI) × Match		-0.015	-0.013		-0.012	-0.003
		(0.016)	(0.018)		(0.019)	(0.022)
Constant	0.073***	0.074***	0.102***	0.134***	0.135***	0.232***
	(0.011)	(0.011)	(0.033)	(0.014)	(0.014)	(0.045)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	1,168	1,168	1,126	1,168	1,168	1,126

	(13)	(14)	(15)	(16)	(17)	(18)
Panel C: Expertise						
	Strict callback			Broad callback		
Treatment (AI)	0.016	0.016	0.014	0.030	0.030	0.025
	(0.016)	(0.016)	(0.017)	(0.021)	(0.021)	(0.022)
Expertise	-0.012	-0.006	-0.005	-0.023**	-0.007	-0.005
	(0.007)	(0.010)	(0.011)	(0.010)	(0.012)	(0.013)
Treatment (AI) × Expertise		-0.012	-0.011		-0.031	-0.025
		(0.015)	(0.015)		(0.020)	(0.021)
Constant	0.073***	0.073***	0.102***	0.134***	0.134***	0.234***
	(0.011)	(0.011)	(0.033)	(0.014)	(0.014)	(0.045)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	1,175	1,175	1,132	1,175	1,175	1,132

Note: Linear probability model (LPM) with robust standard errors reported in parentheses. The dependent variable *Strict (Broad) callback* is equal to one if the application received a callback, and zero otherwise. *Strict callback* includes only interview invitations; *Broad callback* includes all positive responses. *Treatment* is equal to one if the submitted résumé includes AI skills and zero otherwise. *Number of skills required* is the standardized number of skills mentioned in the job advertisement. *Match* is the standardized proportion of skills mentioned in the job advertisement that are also present in the candidate's résumé. *Expertise* is the standardized Dale-Chall index score based on the skills requested in the job advertisement. *Controls* include dummies for the job functions, a dummy equal to one if the job is based in London or the South-East of England and zero otherwise, and controls for companies' characteristics (number of employees, company age, industry dummies). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3.4 Exploring the role of exposure to AI automation

We next investigate whether industry-level exposure to AI-driven automation moderates the treatment effect. To this end, we employ the AI automation exposure index developed by Felten et al. (2021), which measures the share of job tasks within each industry deemed suitable for AI automation. Higher values of this index indicate greater exposure – that is, industries in which a larger proportion of tasks are susceptible to AI-driven task automation.

We assign each job vacancy an AI exposure score based on the industry classification of the posting firm.[17] The original measure is standardized with mean zero and standard deviation equal to one. The median job posting in our sample has a value of 0.242 and a standard deviation of 1.173. The full and job-function-specific distributions are presented in Figures 1 and 5 of the Online Appendix, respectively.

We incorporate the measure of AI industry exposure into our main regression specification and interact it with the treatment variable. The results of these estimations are presented in Table 9. The coefficient for the treatment indicator on broad callbacks (Panel B) is positive and marginally statistically significant in Models (9), (10), and (12), suggesting that listing AI-related qualifications increases the likelihood of receiving a positive callback by approximately 3.9 to 4.7 percentage points, but only in industries not highly exposed to AI automation. The interaction term is negative and statistically significant for the broad callback rate in Models (3)-(6), with *p*-value ranging between 0.027 and 0.046. This indicates that the positive effect of AI qualifications on callback rates diminishes significantly as the AI exposure measure increases. Moreover, Wald tests of that the treatment effect plus its interaction with the AI exposure indicator are equal to zero confirm that the total effect of listing AI

[17] The original AI exposure index from Felten et al. (2021) is based on US NAICS codes, which we convert to UK SIC codes at the five-digit level. For more details see the Online Appendix.

qualifications in industries with AI exposure around one standard deviation above mean is not significantly different from zero (all p-values > 0.727).

Taken together, these results suggest that AI-related qualifications may increase positive callback rates only in industries that are less exposed to AI automation. In contrast, in more AI-exposed industries, the same qualifications do not yield any advantage. One possible interpretation is that in highly AI-exposed industries, firms may already employ a sufficiently AI-competent workforce or view AI qualifications as less differentiating. This pattern may reflect a saturation effect, where the marginal return to signaling AI-related qualifications declines once such technologies become widely embedded in standard work processes.

Table 9 – Treatment effect and exposure to AI

	Panel A: Strict callback					
	(1)	(2)	(3)	(4)	(5)	(6)
	LPM	Probit	LPM	Probit	LPM	Probit
Treatment (AI)	0.014	0.014	0.021	0.023	0.018	0.021
	(0.016)	(0.016)	(0.016)	(0.018)	(0.017)	(0.018)
AI exposure	0.008	0.008	0.016	0.017	0.020*	0.021*
	(0.007)	(0.007)	(0.010)	(0.011)	(0.011)	(0.011)
Treatment (AI) × AI exposure			-0.016	-0.017	-0.017	-0.018
			(0.014)	(0.014)	(0.014)	(0.014)
Constant	0.070***	0.069***	0.067***	0.065***	0.072***	0.071***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.022)	(0.022)
Controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Treatment (AI) + AI × Exposure = 0			0.782	0.706	0.968	0.887
Observations	1,185	1,185	1,185	1,185	1,141	1,141
	Panel B: Broad callback					
	(7)	(8)	(9)	(10)	(11)	(12)
	LPM	Probit	LPM	Probit	LPM	Probit
Treatment (AI)	0.027	0.027	0.044**	0.047**	0.036	0.039*
	(0.021)	(0.021)	(0.022)	(0.023)	(0.022)	(0.023)
AI exposure	0.010	0.010	0.029**	0.031**	0.033**	0.034**
	(0.009)	(0.009)	(0.013)	(0.013)	(0.013)	(0.014)
Treatment (AI) × AI exposure			-0.038**	-0.039**	-0.036**	-0.036**
			(0.018)	(0.018)	(0.018)	(0.018)
Constant	0.130***	0.130***	0.122***	0.120***	0.146***	0.145***
	(0.014)	(0.014)	(0.014)	(0.015)	(0.028)	(0.029)
Controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Treatment (AI) + AI × Exposure = 0			0.805	0.727	0.986	0.889
Observations	1,185	1,185	1,185	1,185	1,141	1,141

Note: Linear probability model (LPM) and Probit model (reporting average marginal effects and constants for the predicted probability of the outcome when all exogenous variables are equal to zero) results; robust standard errors are reported in parentheses. The dependent variable *Strict (Broad) callback* is equal to one if the application received a callback, and zero otherwise. *Strict callback* includes only interview invitations; *Broad callback* includes all positive responses. *Treatment (AI)* is equal to one if the résumé includes AI skills and zero otherwise. *AI exposure* is the measure of exposure to AI at the industry level from Felten et al. (2021). *Controls* include dummies for the job functions, number of employees, company age, and a dummy if the job is based in London or the South-East of England. In the line *Treatment (AI) + AI × Exposure = 0* we report the p-values for a two-sided Wald test. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4. Discussion

This study provides causal evidence on the return to AI-related qualifications for jobseekers in entry-level positions. In particular, we investigate the impact of complementary AI qualifications on the likelihood of receiving invitations to interview. We conduct a large-scale correspondence study, submitting 1,185 fictitious job applications to real entry-level vacancies across six job functions in the UK. By randomly assigning each vacancy to receive either a résumé that includes AI-related qualifications (treatment group) or an identical application without such qualifications (control group), we are able to isolate the effect of AI-related skills on the employers' recruitment interest. Our primary outcome variable of interest is whether an application receives a callback or not.

Contrary to widespread expectations and prevailing labor market narratives, we find no statistically significant difference in callback rates for résumés with and without AI-related qualifications. This finding is surprising given the documented surge in employer demand for AI expertise, the proliferation of AI-related job postings, and the expanding relevance of AI tools across occupational contexts (Alekseeva et al., 2021; Chui et al., 2018; Squicciarini & Nachtigall, 2021; Stapleton et al., 2021).

To understand our findings, it is essential to assess whether our study had sufficient power to detect economically significant impacts of AI-related qualifications. Our sample of 1,185 observations offers 80% statistical power to detect effects of at least 4.8 percentage points for strict callbacks and 6.0 percentage points for broad callbacks (based on the empirical baseline callback of 7.28 and 13.37 percentage points respectively). These effects are larger than the coefficients in our main regressions (Tables 3 and 4 above) which are as large as 1.5 percentage points for strict callbacks and 2.8 percentage points for broad callbacks. To understand if a sample size with sufficient power would have yielded statistically significant results, we perform Monte Carlo simulations. To achieve sufficient power of 80% for estimates of 1.5 (2.8) percentage points for the strict (broad) callbacks respectively, we simulate 1,000 stratified samples of 10,212 (5,072) observations and estimate our main specifications. In the sample of 5,072, the average *p*-value for the treatment indicator for the broad callback is below 0.05, but the average *p*-value for the strict callback is above 0.1. In the sample of 10,212, the average estimated treatment coefficient for both strict and broad callbacks is at least statistically significant at the 5% level and of size 1.5 and 2.8 respectively. If we expand our specifications to include interaction terms with the occupation dummies, all treatment coefficients remain insignificant. This indicates that if there are any effects on callbacks from including AI-related qualifications in résumés, they are likely below 1.5 and 2.8 percentage points for strict and broad callbacks respectively. Empirically larger sample sizes would be required to detect such effects at a statistically significant level. Future research should take into account the requirement of larger sample sizes in a between-subject design for correspondence studies.

To contextualize this effect size, it is useful to compare it to the results of recent correspondence studies. Our suggested effect sizes translate to a 21% increase in the probability of receiving a callback. This is comparable to the 18% increase from listing a master's (vs a bachelor's) degree found by Verhaest et al. (2018) and the 24% decrease resulting from the inclusion of a for-profit (as opposed to non-profit) degree for business-related jobs in the study by Deming et al. (2016). Conversely, much larger effects are observed in the correspondence studies by Lennon (2021) (88% decrease from listing an online vs in-person degree) and Kacperczyk and Younkin (2022) (46% decrease resulting from experience founding a business). Discrete-choice experiments, where HR managers are asked to choose between two résumés, also find much larger effects resulting from candidate characteristics such as social skills or personality traits (Piopiunik et al., 2020; Wehner et al., 2022).

Furthermore, an increase of 21% in the probability of receiving a callback would indicate a larger effect than the 12% increase from listing an internship found by Baert et al. (2021), and, most relevant to our case, the 8% increase identified by Blanco and López Bóo (2020) for advanced ICT skills. These effects are more in line with the expectations of HR professionals in our expert survey, who predicted that including AI-related qualifications would increase the probability of being invited for an interview by 9.6%. Overall, these comparisons highlight that, even if there were a return to AI-related skills that we cannot observe due to our sample size, it is likely small and of the magnitude of gaining internship experience or enhancing one's general ICT skills.

More broadly, the expert survey confirms that our findings are not driven by our experimental design. Importantly, survey respondents confirm that the AI qualifications included in our résumés are both salient and relevant to the entry-level job roles. Additionally, they are associated with higher AI-related competencies, as well as in basic and advanced digital technology. However, the presence of AI-related qualifications also signals weaker social skills and is slightly considered as overqualification. These perceptions highlight the complex interplay of skill signaling and employer interpretation.

As an exploratory analysis, we conduct a series of robustness checks to assess whether job and firm characteristics moderate the treatment effect. Specifically, we control for the number of required skills, the skill match between résumé and posting, and the level of required expertise. None of these factors change the null average effect of AI qualifications on callback rates. However, we find heterogeneity by industry-level exposure to AI automation. Résumés listing AI qualifications receive more callbacks in industries with lower exposure to AI technologies, whereas no such effect is observed in industries with higher exposure. This suggests that the signaling value of AI qualifications may depend on the broader technological context of the hiring industry. Indeed, the industry-specific heterogeneity may also reflect structural differences in how AI technologies interact with human labor. In sectors with high exposure to AI automation, AI is more likely to substitute for routine or codifiable tasks, particularly within specialized teams that already employ dedicated AI professionals (Acemoglu et al., 2022). In such contexts, general AI qualifications may offer limited marginal value for non-specialist roles, since core technical functions are already covered. Conversely, in industries with lower AI exposure, AI-related skills may serve as complements. Here, candidates with AI-related qualifications could be viewed as valuable for enabling or supporting digital transformation processes.

It is also important to note the methodological limitations of our study. The correspondence design we employ captures employer responses only up to the interview invitation stage. Although previous correspondence studies have shown that callbacks are a reliable proxy for hiring probability (Bertrand & Mullainathan, 2004), our data do not allow us to observe actual hiring decisions, wage offers, or career progression. It remains possible that AI-qualified candidates, despite receiving similar callback rates, may be favored later in the hiring funnel or command higher compensation or steeper career progression once employed. Moreover, our analysis focuses on entry-level positions. These roles typically emphasize general skills and standardized application materials, making them well-suited to an experimental study. More senior roles often require non-transferable credentials or detailed work histories that are difficult to standardize experimentally. While our survey results confirm the relevance of AI qualifications for the jobs we study, we cannot exclude the possibility that such qualifications yield greater labor market returns in more senior or specialized positions.

Thus, the potential for stronger signaling effects in more advanced labor market segments remains unexplored. Future research could investigate how this dynamic evolves over time as AI adoption continues to reshape the demand for digital competencies across occupations and industries. Moreover, examining whether similar patterns hold in mid-career or AI-specific job markets, and to what extent these dynamics are shaped by social fit and skill complementarities, may provide deeper insights into how AI qualifications are evaluated in real-world hiring contexts.

It is also important to contextualize our findings temporally, given the rapidly changing nature of AI. Our correspondence experiment was conducted in 2021, with data collection ending in early 2022, before the launch of ChatGPT and the widespread adoption of LLMs. This may have important implications, as it is likely the average recruiter had limited exposure to AI during the timeline of our experiment. From this perspective, our study should be interpreted as providing a reference point for the returns to AI skills at the early stage of technological development. Nevertheless, our contribution is to demonstrate the returns to skills related to an emerging technology. Our timing – when AI was already in the process of being adopted by firms (Alekseeva et al., 2021; Babina et al., 2024) but before it obtained the status of a 'hyped' technology (Floridi, 2024) – may be more suitable for this purpose. Future research could explore whether, as the technology matures, recruiters respond differently to signals of AI qualifications.

From a policy perspective, our findings caution against a one-size-fits-all approach to digital upskilling. Universal AI training programs may not add uniform value across all industries. Targeted upskilling initiatives focused on industries or occupations where AI adoption is emerging may be more effective in enhancing employability. Additionally, our findings call attention to the question of whether workers should bear the costs of AI-related human capital investments, especially when labor market rewards remain uncertain.

5. Conclusion

In sum, our findings challenge the assumption that AI-related qualifications unambiguously enhance employability in early-career recruitment. While such skills might be valued in abstract or strategic terms, they do not automatically translate into interview opportunities, at least not in the entry-level labor market in job functions such as HR, Finance, Marketing, Engineering, IT and Logistics. Even when highly visible and positively perceived by hiring professionals, AI-related qualifications did not consistently improve early-stage hiring outcomes across a broad range of entry-level occupations. However, these skills do appear to generate positive effects in industry sectors where AI has not yet become fully embedded. This divergence between experts' expectations and empirical outcomes suggests that the labor market returns to AI skills are context-dependent.

Acknowledgements: The authors are grateful to Charlotte Knickrehm, Judith Künnecke, Max Thon, Anja Schöttner, Heiner Schumacher, and Dirk Sliwka for their comments; to Johannes Beckers, David Alexander Behrens, Haoyu Cao, Alexis Hall, and Müge Özdemir for their research assistance, and to participants at the 29th Annual Conference of the Society for Institutional and Organizational Economics (SIOE), Annual Meeting of the Committee for Organizational Economics in 2025 (OrgEcon, VfS), XXIV. Symposium zur ökonomischen Analyse der Unternehmung, German Economic Association of Business Administration (GEABA), Digital Economcy Workshop in 2025 (DIW), PSE-CEPR Policy Forum in 2025, Annual meeting of the German Association for Experimental Economic Research Wirtschaftsforschung in 2024 (GfeW), Connect AI Berlin: Advancing AI Research and Societal Impact in 2023 at Max Planck Institute for Human Behavior; as well as participants of the research seminars at RWTH Aachen (2025) and RMIT University Melbourne for their feedback.

Funding: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 280092119 - TRR 190 Rationality and Competition and the Einstein Center Digital Future.

References

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics* (Vol. 4, pp. 1043–1171). Elsevier. https://doi.org/10.1016/S0169-7218(11)02410-5

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*(S1), S293–S340. https://doi.org/10.1086/718327

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. https://doi.org/10.1257/aer.20160696

Acemoglu, D., & Restrepo, P. (2019). Artificial Intelligence, Automation, and Work. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda* (pp. 197–236). The University of Chicago Press.

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. *Journal of Economic Perspectives*, *33*(2), 31–50. https://doi.org/10.1257/jep.33.2.31

Alekseeva, L., Azar, J., Giné, M., Samila, S., & Taska, B. (2021). The Demand for AI Skills in the Labor Market. *Labour Economics*, *71*(102002). https://doi.org/10.1016/j.labeco.2021.102002

Arrow, K. J. (1973). The Theory of Discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets*.

Autor, D. (2024). *Does Automation Replace Experts or Complement Expertise? The Answer is Yes.* [Joseph R. Schumpeter lecture]. European Economic Association, Rotterdam.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333.

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151,* 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Baert, B. S., Neyt, B., Siedler, T., Tobback, I., & Verhaest, D. (2021). Student internships and employment opportunities after graduation: A field experiment. *Economics of Education Review*, *83*, 102141. https://doi.org/10.1016/j.econedurev.2021.102141

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is There Less Discrimination in Occupations Where Recruitment Is Difficult? *ILR Review*, *68*(3), 467–500. https://doi.org/10.1177/0019793915570873

Balsmeier, B., & Woerter, M. (2019). Is this time different? How digitalization influences job creation and destruction. *Research Policy*, *48*(8), 103765. https://doi.org/10.1016/j.respol.2019.03.010

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition. *American Economic Review*, *106*(6), 1437–1475. https://doi.org/10.1257/aer.20140571

Bassi, V., & Nansamba, A. (2022). Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda. *The Economic Journal*, *132*(642), 471–511. https://doi.org/10.1093/ej/ueab071

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition*. NBER. https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-first-edition

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, *94*(4), 991–1013. https://doi.org/10.1257/0002828042002561

Blanco, M., & López Bóo, F. (2020). The Labor Market Returns to ICT Skills: A Field Experiment. In *Information Technologies and Economic Development in Latin America* (pp. 141–160). Anthem Press.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work*. *The Quarterly Journal of Economics*, *140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn, and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings*, *108*, 43–47. https://doi.org/10.1257/pandp.20181019

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. In *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Burning Glass Technologies. (2019). *Mapping the Genome of Jobs: The Burning Glass Skills Taxonomy*.

Carmichael, H. L., & MacLeod, W. B. (1993). Multiskilling, Technical Change and the Japanese Firm. *The Economic Journal*, *103*(416), 142–160. https://doi.org/10.2307/2234341

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., & Malhotra, S. (2018). *Notes from the AI frontier: AI adoption advances, but foundational barriers remain* (Discussion Paper).

Coursera. (2025). *Global Skills Report: Trusted skills insights for a rapidly changing world*. Coursera.

Czarnitzki, D., Fernández, G. P., & Rammer, C. (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior & Organization*, *211*, 188–205. https://doi.org/10.1016/j.jebo.2023.05.008

Dale, E., & Chall, J. (1948). A Formula for Predicting Readability. *Educational Research Bulletin*, *27*(1–20), 37–54.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. https://doi.org/10.2139/ssrn.4573321

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10.1093/qje/qjx022

Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., & Katz, L. F. (2016). The Value of Postsecondary Credentials in the Labor Market: An Experimental Study. *American Economic Review*, *106*(3), 778–806. https://doi.org/10.1257/aer.20141757

Deming, D., & Kahn, L. B. (2018). Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals. *Journal of Labor Economics*, *36*(S1).

Duch-Brown, N., Gomez-Herrera, E., Mueller-Langer, F., & Tolan, S. (2022). Market power and artificial intelligence work on online labour markets. *Research Policy*, *51*(3), 104446. https://doi.org/10.1016/j.respol.2021.104446

Falck, O., Heimisch-Roecker, A., & Wiederhold, S. (2021). Returns to ICT skills. *Research Policy*, *50*(7), 104064. https://doi.org/10.1016/j.respol.2020.104064

Farnham, M., & Hutchinson, E. (2011). The Effect of Multiskilling on Labor Productivity, Product Quality, and Financial Performance. In *Advances in the Economic Analysis of Participatory & Labor-Managed Firms* (Vol. 12, pp. 35–62). Emerald Group Publishing. https://doi.org/10.1108/S0885-3339(2011)0000012006

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217. https://doi.org/10.1002/smj.3286

Felten, E. W., Raj, M., & Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, *108*, 54–57. https://doi.org/10.1257/pandp.20181021

Floridi, L. (2024). Why the AI Hype is Another Tech Bubble. *Philosophy & Technology*, *37*(4), 128. https://doi.org/10.1007/s13347-024-00817-w

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, *52*(1), 104653. https://doi.org/10.1016/j.respol.2022.104653

Ham, C., Hann, R. N., Rabier, M., & Wang, W. (2025). Auditor Skill Demands and Audit Quality: Evidence from Job Postings. *Management Science*, *71*(7), 5805–5829. https://doi.org/10.1287/mnsc.2022.01640

Heinz, M., & Schumacher, H. (2017). Signaling cooperation. *European Economic Review*, *98*, 199–216. https://doi.org/10.1016/j.euroecorev.2017.06.017

Hershbein, B., & Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review*, *108*(7), 1737–1772. https://doi.org/10.1257/aer.20161570

Holm, J. R., & Lorenz, E. (2022). The impact of artificial intelligence on skills at work in Denmark. *New Technology, Work and Employment*, *37*(1), 79–101. https://doi.org/10.1111/ntwe.12215

Hui, X., Reshef, O., & Zhou, L. (2024). The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market. *Organization Science*, *35*(6), 1977–1989. https://doi.org/10.1287/orsc.2023.18441

Humburg, M., & Van Der Velden, R. (2015). Skills and the graduate recruitment process: Evidence from two discrete choice experiments. *Economics of Education Review*, *49*, 24–41. https://doi.org/10.1016/j.econedurev.2015.07.001

Kacperczyk, O., & Younkin, P. (2022). A Founding Penalty: Evidence from an Audit Study on Gender, Entrepreneurship, and Future Employment. *Organization Science*, *33*(2), 716–745. https://doi.org/10.1287/orsc.2021.1456

Kim, Y.-M., & Park, K. S. (2003). Multiskilling and Firm Performance. *Seoul Journal of Economics*, *16*(4), 387–422.

Kline, P., Rose, E. K., & Walters, C. R. (2022). Systemic Discrimination Among Large U.S. Employers. *The Quarterly Journal of Economics*, *137*(4), 1963–2036. https://doi.org/10.1093/qje/qjac024

Koedel, C., & Tyhurst, E. (2012). Math skills and labor-market outcomes: Evidence from a resume-based field experiment. *Economics of Education Review*, *31*(1), 131–140. https://doi.org/10.1016/j.econedurev.2011.09.006

Korpi, T., & Tåhlin, M. (2009). Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Economics*, *16*(2), 183–193. https://doi.org/10.1016/j.labeco.2008.08.004

Kuhn, P., & Weinberger, C. (2005). Leadership Skills and Wages. *Journal of Labor Economics*, *23*(3).

Lennon, C. (2021). How Do Online Degrees Affect Labor Market Prospects? Evidence from a Correspondence Audit Study. *ILR Review*, *74*(4), 920–947. https://doi.org/10.1177/0019793919899943

Levine, S. S., Schilke, O., Kacperczyk, O., & Zucker, L. G. (2023). Primer for Experimental Methods in Organization Theory. *Organization Science*, *34*(6), 1997–2025. https://doi.org/10.1287/orsc.2023.18093

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). *The AI Index 2024 Annual Report* (AI Index Steering Committee). Institute for Human-Centered AI.

McGuinness, S., & Wooden, M. (2009). Overskilling, Job Insecurity, and Career Mobility. *Industrial Relations: A Journal of Economy and Society*, *48*(2), 265–286. https://doi.org/10.1111/j.1468-232X.2009.00557.x

Piopiunik, M., Schwerdt, G., Simon, L., & Woessmann, L. (2020). Skills, signals, and employability: An experimental investigation. *European Economic Review*, *123*, 103374. https://doi.org/10.1016/j.euroecorev.2020.103374

Pouliakas, K., Santangelo, G., & Dupire, P. (2025). Are Artificial Intelligence (AI) Skills a Reward or a Gamble? Deconstructing the AI Wage Premium in Europe. *IZA Discussion Papers*, *17607*. https://doi.org/10.2139/ssrn.5099090

Rammer, C., Fernández, G. P., & Czarnitzki, D. (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy*, *51*(7), 104555. https://doi.org/10.1016/j.respol.2022.104555

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Squicciarini, M., & Nachtigall, H. (2021). *Demand for AI skills in jobs: Evidence from online job postings* (No. 03; Science, Technology and Industry Working Papers). OECD. https://dx.doi.org/10.1787/3ed32d94-en

Stapleton, K., Copestake, A., & Pople, A. (2021). AI, firms and wages: Evidence from India. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3957858

Stephany, F., & Teutloff, O. (2024). What is the price of a skill? The value of complementarity. *Research Policy*, *53*(1), 104898. https://doi.org/10.1016/j.respol.2023.104898

Van Beek, K. W. H., Koopmans, C. C., & Van Praag, B. M. S. (1997). Shopping at the labour market: A real tale of fiction. *European Economic Review*, *41*(2), 295–317. https://doi.org/10.1016/S0014-2921(96)00037-2

Verhaest, D., Bogaert, E., Dereymaeker, J., Mestdagh, L., & Baert, S. (2018). Do Employers Prefer Overqualified Graduates? A Field Experiment. *Industrial Relations: A Journal of Economy and Society*, *57*(3), 361–388. https://doi.org/10.1111/irel.12212

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *Working Paper*.

Wehner, C., De Grip, A., & Pfeifer, H. (2022). Do recruiters select workers with different personality traits for different tasks? A discrete choice experiment. *Labour Economics*, *78*, 102186. https://doi.org/10.1016/j.labeco.2022.102186

Xie, M., Ding, L., Xia, Y., Guo, J., Pan, J., & Wang, H. (2021). Does artificial intelligence affect the pattern of skill demand? Evidence from Chinese manufacturing firms. *Economic Modelling*, *96*, 295–309. https://doi.org/10.1016/j.econmod.2021.01.009

Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., Lyons, T., Manyika, J., Niebles, J. C., Sellitto, M., Shoham, Y., Clark, J., & Perrault, R. (2021). *The AI Index 2021 Annual Report* (No. arXiv:2103.06312). arXiv. https://doi.org/10.48550/arXiv.2103.06312

Online appendix

The Elusive Returns to AI Skills: Evidence from a Field Experiment

1. Correspondence study

This section contains additional information on the design and data collection of the correspondence experiment.

1.1. Additional information on the résumés

For each job function, we create two identical résumés that only differ in their inclusion of AI-related skills. The résumés have the following sections:

- Personal statement (a brief summary of the candidate's abilities)
- Education
- Skills (short bullets)
- Languages
- Hobbies
- Work experience
- Additional experience (describing online courses taken and volunteering activities)
- References (a statement indicating a reference can be provided)

The skills and details of the professional experience in the résumés vary by job function, but all retain the same experience level and a few characteristics. The candidate is male and has a common name,[1] and comes from Sheffield, England. He has an undergraduate and a master's degrees from Nottingham Trent University, both with a 2:1 grade (which corresponds to a high, but not extraordinary score). The skills section always includes relatively advanced IT skills that would be a pre-requisite for any application of AI, and an additional language. The candidate also has some work experience: a summer internship during his undergraduate studies, and a placement year during the master's degree.[2] The résumés also list a number of extracurricular activities (sport hobbies, volunteering) that signal social and other non-cognitive skills as well as an ability to cooperate, which have been found to be beneficial for

[1] The candidate's first and last names were selected based on top 50 entries in the Office for National Statistics and National Records of Scotland, respectively. A manual check confirmed that there are over a dozen profiles with the name on LinkedIn, which is helpful in avoiding detection by employers who might search for the candidate's name online.

[2] In the UK, master's degrees are usually only one year long. We select a two-year program to allow our candidate some work experience, while maintaining the same structure across job functions.

jobseekers (Heinz & Schumacher, 2017) and complementary with advanced IT skills (Håkanson et al., 2021). In sum, these characteristics allow us to apply for a range of roles that require either no or little work experience.

The AI skills presented in the treatment résumés are also tailored to the respective job function. They always appear in the same sections: a mention in the "Personal statement" section; an additional AI-related competency in the "Skills" section; an extra bullet point highlighting AI skills in both the "Education" and "Work Experience" sections; and a record of an AI online course listed under "Additional Experience." The main keywords were identified from the list of most requested AI skills in Alekseeva et al. (2021), and included: machine learning, artificial intelligence, deep learning, neural networks, and IBM Watson.[3] For each applicant, these AI skills are embedded in résumés through use cases in the respective occupational field. The use cases are based on trends and surveys from industry reports (e.g. from Gartner, Forrester Research), as well as interviews with sectoral experts. All references to AI in the résumés and the sources for the use cases are detailed in Table 1 below.

The résumés are available in the 'Résumés' folder of this Online Appendix. They are named after the respective job function (e.g. Marketing) and treatment group (e.g. AI).

1.2. Randomization procedure

Before applying for jobs, we generate a random allocation to decide if an application should use the treatment or control résumé. We stratify by job type. Because we cannot control how many job postings for each job type will be available on any given week, and we aim to ensure numerical balance between the treatment and control groups over the timeline of the experiment, we randomize job postings in blocks of four (e.g. Treatment-Control-Control-Treatment). The randomization was conducted in Stata (do files with replicable seeds are available in the Data Repository).

[3] Additionally, the IT résumé included the keyword 'Azure AI' which was directly mention by an industry expert as a key skill for the profile.

Table 1 – AI skills and use cases included in treatment résumés

Job function	*Mentions*	*Sources*
Engineering	• AI and ML for engineering • Using Deep Learning techniques for image based part classification • Developed a machine learning (ML) approach to anomaly detection in the production process • Applied AI methods to improve product evaluation with large scale experimental databases in Python • Machine Learning Engineering for Production MLOps (on Coursera.org)	McKinsey Global Institute (2019); Forrester Inc (2021)
Finance	• IBM Watson • Using AI/Machine Learning algorithms for fraud detection, riskmodeling, and financial forecasts and planning • Implementing and using IBM's Watson AI for finance • Using Machine Learning algorithms and tools for prescriptive financial forecasting and finance process automation with RPA • Neural Networks and Deep Learning (on Coursera.org)	Bachinskiy (2019); Balakrishnan et al. (2020); Bryan (2018)
HR	• IBM Watson • Using AI/Machine Learning algorithms for automated workforce planning, talent acquisition, and performance management • Implementing and using IBM's Watson AI for HR • Using Machine Learning algorithms and tools for AI-supported virtual HR assistants and employee skills management • Neural Networks and Deep Learning (on Coursera.org)	Pemberton (2018); Wang (2018)
IT	• Azure AI • Using AI/Machine Learning algorithms to improve or automate business processes and practices • Implementing AI knowledge mining solutions with Azure AI • Using tools for AI conversational solutions with firm clients • Neural Networks and Deep Learning (on Coursera.org)	Costello (2020); Forrester Inc (2019); Forrester Inc (2021)
Logistics	• Machine Learning • Using AI tools to improve demand and supply planning • Implementing a Machine Learning algorithm to improve demand forecasting • Using AI tools to develop a dynamic supply planning to optimise the supply chain flow • Machine Learning for Supply Chains Specialization (on Coursera.org)	Gartner (2019); McKinsey & Company (2021)
Marketing	• IBM Watson • Using AI/Machine Learning algorithms for personalised customer recommendations, prescriptive actions, and real-time analytics • Implementing and using IBM's Watson AI for marketing • Using Machine Learning algorithms and tools for marketing orchestration and real-time personalisation • Neural Networks and Deep Learning (on Coursera.org)	Baker (2020); Gartner, Inc. (2019)

Note: The table presents the mentions of AI skills in treatment résumés. The first column represents the job function; the second column the AI skills added to the résumés, using the exact wording from the résumés; the third column the sources of the use case for the corresponding job function.

1.3. Applying for jobs

We search for open vacancies using the top online job boards in the UK (the complete list includes: "CV Library," "Indeed," "Milkround," "Monster," "Reed," "Totaljobs"). We restrict our search to job offers that required either no or little professional experience (up to one year).[4] To do so, we make use of the following keywords: "assistant," "entry-level," "junior," followed by the specific job function keyword ("finance"/"accounting," "HR"/"human resources," "marketing," "IT," "engineer," "logistics"/"supply chain"). We then select suitable job offers based on the available search results. We avoid job openings that ask for specific, uncommon skills (e.g., speaking Mandarin). We apply to each vacancy once, following the pre-defined random allocation to decide whether to send the résumé that includes AI-skills (treatment) or not (control). We always complete a predefined randomization block in one sitting, to ensure numerical balance over time. To avoid detection, we do not apply to more than one job per employer. All applications were completed between March and November 2021. In total, we were able to apply to 1,185 job openings.

1.4. Collecting data from the job postings

We begin the application process with an Excel spreadsheet that contains a list of IDs and the randomization sequence described in section 1.2. above. For each row, we collect data directly from the job postings. This includes noting the company name, job posting title, location, job board on which it was advertised, company website (where available[5]), and wage offered (where available[6]). To store the job descriptions, we save the entire content of the website either as a Google Docs document or as an HTML file, placing each file inside a folder named after the ID of the application. We then employ a Python script to automatically extract the text from the job postings, alongside the application IDs, in a file named "job_descriptions.csv," which we incorporate into the main data during processing (under the variable 'raw_description"). Due to a manual error, these data were incorrectly stored for 10 advertisements and could therefore not be collected. All full job description files are available upon request.

1.5. Tracking callbacks

To track callbacks, we set up separate email addresses and UK phone numbers. Over the course of 40 days after each application, we monitor both the email inboxes and phone numbers to

[4] This could make the candidate – who holds a master's degree – slightly overqualified for certain entry-level jobs. However, recent evidence from another audit experiment shows this should give the candidate an advantage (Verhaest et al., 2018).
[5] No website was available for 101 of the job advertisements.
[6] This information was not available for 814 job advertisements.

keep track of callbacks. Upon receiving an email response, we categorize it as follows: invitation to interview; request for more information (such as willingness to relocate for the job); offer to apply for a separate job opening at the company; offer to interview for a separate job opening at the company; request to file a separate application form; rejection. All but the last two were categorized as "broad" callbacks. Only interview invitations are categorized as "strict" callbacks. Upon receiving a phone call, we follow the same categorization approach as for the emails, based on either (a) SMS messages (only if the company identifies itself); (b) voicemails (only if the company identifies itself[7]); (c) matching phone numbers to companies if they did not send an SMS message nor left a voicemail (in this case they are categorized as "broad" callback).

As soon as a company reaches out to the candidate, we reply via email politely rejecting the offer.

1.6. Matching job advertisements to company data

Using the company name as a starting point, we collected detailed data on the firms advertising the vacancies. We began by manually obtaining data on the postcode and region for the company location as advertised on the job posting. We then searched for the company in the UK's public registry for companies (Companies House), ensuring key details were correct by cross-referencing the location of incorporation, and searching the company's website for a Companies House number (the company website is usually available in the job posting). From Companies House we extract the industry code (SIC code),[8] the year of founding, the number of employees and revenues using the latest report (where available[9]). In cases where any of this information was missing (e.g. because not all companies are required to report revenues), we searched three alternative data sources: the companies' LinkedIn page (which tends to report the founding year of the company); Dun and Bradstreet, a commercial data analytics firm that provides some public information on companies; and the ORBIS database. We report the source of each data point in the variables "employees_source," "revenues_source," and "year_source."

Through this approach, we obtain a broad coverage of our main control variables of interest (see Table 2 below for a breakdown).

[7] In all cases companies sending an SMS or leaving a voice mail identified themselves clearly.
[8] In a few cases, this information was not available, as the company's sector was therefore coded as "Other."
[9] This information was not available for 50 companies (employees) and 406 companies (revenues).

Table 2 – Coverage of key control variables

Variable	*Missing*	*Available*
Founding year	0.76%	99.24%
Sector (SIC)	1.01%	98.98%
Employees	4.22%	95.78%
Location	0.08%	99.92%
Job description	0.84%	99.16%

1.7. Matching job advertisements to the measure of "exposure to AI" by Felten et al. (2021)

Felten et al. (2021) developed a measure of exposure to AI at the industry level based on the tasks carried out by occupations within that industry. Their measure provides a score at the 4-digit NAICS code (from the US Bureau of Labor Statistics). Since our job advertisements are categorized using the UK's Standard Industrial Classification (SIC) codes, we must first use a crosswalk to convert the scores to UK industries. Because no direct NAICS-to-UK SIC crosswalk exists, we first convert the NAICS codes into International Standard Industrial Classification of All Economic Activities (ISIC, the global classification developed by the UN) codes, and only then to UK SIC, following the procedures suggested by the UK Office for National Statistics (ONS).[10] We manually check the results. This approach, however, leaves us with around 600 unmatched observations. We employ a simple large language model (LLM) based approach, creating word embeddings from the NAICS descriptions and matching them to the UK SIC definitions.[11] We again manually check the results. Combining these two approaches, we are able to match all observations to a score for AI exposure.

[10] See https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/mapinguksic2007codestonaic for an explanation.

[11] The Python code is available upon request.

2. Prolific survey

We design a complementary online survey experiment to elicit perceptions of résumés that included or excluded AI-related skills. The survey was administered through Prolific, an online platform that allows for the recruitment of participants based on precise screening criteria. The study was conducted in December 2022. The full survey can be found in the folder ‘Prolific’ of this Online Appendix.

2.1 Participants and Screening

To ensure the relevance of the assessments, we pre-screen participants based on two key criteria: prior experience making hiring decisions and past or current experience working in one of the six occupational sectors used in our correspondence study (Engineering, Finance, HR, IT, Logistics, and Marketing). Participants are excluded if their responses during the screener were inconsistent with their Prolific pre-screening information. Moreover, we embed three attention checks within the survey—asking participants to select a specific option—and discard responses that fail three attention checks (as recommended by the Prolific platform).

2.2 Design and Implementation

We program the survey with Qualtrics. We use embedded variables to customize each résumé to the participant's background, ensuring high contextual relevance. AI skill descriptions and applications match those used in the field experiment, e.g., IBM Watson for HR and Finance roles, Azure AI for IT roles, etc., with job-appropriate use cases (e.g., fraud detection, workforce planning, marketing automation). Each participant is exposed to both treatment and control résumés in randomized order to control for order effects.

2.3 Data Anonymity and Ethics

Participants are provided informed consent, and are informed that participation is voluntary and anonymous. Participants ae compensated for their time (£1.80 base pay plus performance-based bonuses).

3. Additional results

3.1 Additional tables

Table 3 – Differences in callback rates by function

Panel (A): Strict callback rates

Job function	Engineering (7.00%)	Finance (3.17%)	HR (7.65%)	IT (7.50%)	Logistics (11.00%)	Marketing (11.50%)
	p-values					
Engineering	-	0.088*	0.803	0.847	0.162	0.120
Finance			0.053*	0.059*	0.003**	0.002**
HR				0.954	0.253	0.194
IT					0.227	0.173
Logistics						0.874

Panel (B): Broad callback rates

Job function	Engineering (14.00%)	Finance (6.35%)	HR (13.78%)	IT (16.50%)	Logistics (19.50%)	Marketing (18.00%)
	p-values					
Engineering	-	0.013**	0.949	0.487	0.141	0.275
Finance			0.016**	0.002**	0.000***	0.000***
HR				0.450	0.126	0.250
IT					0.435	0.691
Logistics						0.701

Note: The table reports *p*-values from two-sided tests of proportions comparing overall callback rates between the occupation in the corresponding row and column respectively. The columns report the occupation and the callback rates in parentheses. *Strict callbacks* include only interview invitations; *Broad callbacks* include all positive responses. The total number of observations is 1,185. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4 – Callback rates by job function and treatment

Panel (A): Strict callback rates

Job function	Overall	AI	Control	Ratio (AI/control)	*p*-values	Test statistic
Engineering	7.00%	10.00%	4.00%	2.50	0.096*	-1.663
Finance	3.17%	3.16%	3.19%	0.99	0.989	0.013
HR	7.65%	9.09%	6.19%	1.47	0.444	-0.765
IT	7.50%	5.00%	10.00%	0.50	0.179	1.342
Logistics	11.00%	9.00%	13.00%	0.69	0.366	0.904
Marketing	11.50%	16.00%	7.00%	2.29	0.046**	-1.995
All functions	8.02%	8.75%	7.28%	1.20	0.349	-0.937

Panel (B): Broad callback rates

Job function	Overall	AI	Control	Ratio (AI/control)	*p*-values	Test statistic
Engineering	14.00%	20.00%	8.00%	2.50	0.014**	-2.445
Finance	6.35%	6.32%	6.38%	0.99	0.985	0.019
HR	13.78%	12.12%	15.46%	0.78	0.497	0.679
IT	16.50%	13.00%	20.00%	0.65	0.182	1.334
Logistics	19.50%	21.00%	18.00%	1.17	0.592	-0.535
Marketing	18.00%	24.00%	12.00%	2.00	0.027**	-2.209
All functions	14.77%	16.16%	13.37%	1.21	0.175	-1.356

Note: *Strict callbacks* include only interview invitations; *Broad callbacks* include all positive responses. *p*-values and test statistics from two-sided tests of proportions comparing rates between AI and control groups. The total number of observations is 1,185. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5 – Percentage of participants correctly guessing which résumé contained AI skills (expert survey)

Job function	Percentage that guessed correctly	P-value	N
Engineering	93.80%	0.000***	129
Finance	97.67%	0.000***	129
HR	94.66%	0.000***	206
IT	99.24%	0.000***	132
Logistics	91.80%	0.000***	122
Marketing	96.23%	0.000***	53
All functions	95.46%	0.000***	771

Note: The table reports the percentage of Prolific survey participants correctly guessing which of the two résumés contained AI skills, after seeing them in random order. *p*-values from two-sided Binomial tests against a random guess of 50%. The total number of observations is 771. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6 – Descriptions of skill categories (expert survey)

Skills categories	*Description*
AI-related	AI-related digital skills are employed to apply the methods and concepts of artificial intelligence in business functions. Such skills often include advanced programming and algorithm development, and applying machine learning methods.
Advanced digital	Advanced digital skills are employed to apply theoretical and analytical knowledge when using technology. These skills often include data analysis, coding, web and app development, etc.
Basic digital	Basic digital skills are employed to complete simple tasks using rudimentary digital devices and applications. These skills often include digital communication, creating and managing documents/spreadsheets, installing software updates, etc.
Cognitive	Cognitive skills are employed to process incoming information. These skills often include analysis, research skills, problem solving, critical thinking, math and statistics.
Social	Social skills are employed in interpersonal settings. These skills often include skills such as communication, teamwork, collaboration, negotiation.

Note: The table reports the descriptions given to the Prolific study participants for each of the skills categories.

Table 7 – Perceived importance of AI skills for execution of the job role (expert survey)

Job function	Agree	Disagree	Neither … nor …	P-values	Z-scores
Engineering	64.34%	13.18%	22.48%	0.000***	17.179
Finance	54.26%	22.48%	23.26%	0.000***	8.647
HR	51.94%	22.33%	25.73%	0.000***	10.205
IT	52.27%	25.00%	22.73%	0.000***	7.236
Logistics	66.39%	11.48%	22.13%	0.000***	19.032
Marketing	56.60%	18.87%	24.53%	0.000***	7.022
All functions	57.07%	19.33%	23.61%	0.000***	26.542

Note: The table reports the responses to the question "The AI skills [*list of AI skills from the résumé*] would significantly improve a candidate's ability to execute the [*job function*] assistant role" on a scale from one "Strongly disagree" to five "Strongly agree." "Agree" and "Strongly agree" are reported together as "Agree;" "Disagree" and "Strongly disagree" are reported together as "Disagree;" "Neither / nor" stands for "Neither agree nor disagree." *p*-values and test statistics (z-scores) from a two-sided one sample test of proportions comparing the share of "Agree" and "Disagree" responses. The total number of observations is 771. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8 – Perceived under- or over-qualification of candidates by résumé (expert survey)

Job function	Treatment (AI)	Control	*p*-values	Test statistics
Engineering	3.605 (0.744)	3.690 (0.737)	0.228	-11
Finance	3.853 (0.697)	3.783 (0.637)	0.321	9
HR	3.791 (0.739)	3.748 (0.755)	0.489	9
IT	4.030 (0.761)	3.924 (0.768)	0.132	14
Logistics	3.951 (0.801)	3.861 (0.806)	0.267	11
Marketing	3.811 (0.833)	3.717 (0.794)	0.556	5
All functions	3.838 (0.763)	3.790 (0.749)	0.096*	37

Note: The table reports the mean response from one ("Underqualified") to five ("Overqualified") to the question: "In your opinion, how well do the candidate's qualification and experience fit an entry-level [*job function*] role?" *p*-values and test statistics control from two-sided Fisher-Pitman permutation tests for paired replicates with 200,000 runs testing the difference in perception of the treatment and control résumés. The total number of observations is 771. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9 – Mean expected (strict) callbacks by job function (expert survey)

Job function	Treatment (AI)	Control	Implied effect size in pp	*p*-value	Test statistic
Engineering	32.504 (13.437)	28.512 (13.448)	7.984	0.000***	515
Finance	33.341 (12.52)	30.202 (12.25)	6.279	0.000***	405
HR	31.617 (13.022)	29.000 (12.67)	5.233	0.001***	539
IT	29.697 (13.733)	28.098 (13.421)	3.197	0.074*	211
Logistics	31.557 (13.005)	29.385 (13.443)	4.344	0.006***	265
Marketing	27.679 (14.706)	24.113 (12.752)	7.132	0.033**	189
All functions	31.445 (13.296)	28.690 (13.029)	5.510	0.000***	2,124

Note: The table reports mean responses to the question: "Out of 50 companies, how many companies invited the candidate with this résumé for an interview?" Standard deviations in the parentheses. *Implied effect size in pp* is a variable that captures the expected difference in callbacks between treatment and control résumés (treatment effect) in percentage points. *p*-values and test statistics from two-sided Fisher-Pitman permutation tests for paired replicates with 200,000 runs comparing AI and control résumés. The total number of observations is 771. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10 – Balance in control variables by treatment group

Variable	Overall	Control	AI	p-value
Company's age	31.179 (36.974) [1176]	31.242 (35.046) [591]	31.115 (38.854) [585]	0.953
Location dummy for London or the South-East	0.295 (0.456) [1185]	0.296 (0.457) [594]	0.296 (0.457) [591]	0.905
Number of employees	1,183.336 (5,664.098) [1142]	1,185.724 (5,410.291) [579]	1,180.881 (5,918.603) [563]	0.988

Note: Descriptive statistics from the correspondence study for selected control variables by treatment group. Robust standard deviations are shown in parentheses. The number of observations is reported in square brackets. p-values are from two-sided tests of proportions comparing means between the AI and control applications. Comparisons by job function also yield statistically insignificant differences in means (all p-values > 0.138), with one exception: *Company age* in the Logistics function differs between control and treatment groups (36.194 vs. 46.980; p-value of 0.074).

Table 11 – Response time by treatment group

Job function	Overall	Control	AI	p-value
Engineering	6.938 (7.270) [81]	7.081	6.818	0.872
Finance	6.793 (8.048) [58]	7.387	6.111	0.552
HR	6.631 (7.676) [65]	6.118	7.194	0.577
IT	6.619 (8.599) [105]	6.772	6.438	0.844
Logistics	9.923 (10.893) [117]	11.547	7.962	0.076
Marketing	4.506 (5.542) [83]	4.122	4.881	0.536
All functions	7.106 (8.558) [509]	7.549	6.629	0.226

Note: Descriptive statistics from the correspondence study. *Response time* is the difference in days between the date of the application and the date of the recruiter's response (if any). Standard deviations are shown in parentheses. The number of observations is reported in square brackets. p-values are from two-sided t-tests of means between the AI and control applications.

Table 12 – Sample distribution and AI exposure by industry sectors

Industry sector (UK SIC)	Share of companies			*p*-value	Mean AI exposure
	Overall	AI	Control		
Accommodation and food services	0.020	0.024	0.017	0.417	-1.001
Administrative and support services	0.119	0.126	0.112	0.438	0.515
Agriculture, forestry and fishing	0.004	0.005	0.003	0.658	-1.591
Arts, entertainment and recreation	0.010	0.010	0.010	0.993	-0.180
Construction	0.034	0.025	0.042	0.105	-1.045
Education	0.023	0.020	0.025	0.551	1.444
Electricity, gas, steam and air conditioning supply	0.006	0.003	0.008	0.254	0.248
Financial and insurance activities	0.055	0.054	0.056	0.882	1.920
Human health and social work activities	0.024	0.012	0.036	0.007***	0.302
Information and communication	0.142	0.149	0.134	0.426	1.457
Manufacturing	0.204	0.205	0.203	0.920	-0.318
Mining and quarrying	0.002	0.003	0.000	0.157	-1.009
Other or missing	0.012	0.008	0.015	0.279	1.597
Non-trading company	0.011	0.008	0.014	0.398	1.699
Other service activities	0.024	0.025	0.024	0.862	-0.291
Professional, scientific and technical activities	0.123	0.141	0.105	0.056*	1.464
Public administration and defence	0.003	0.002	0.005	0.315	0.685
Real estate activities	0.016	0.013	0.019	0.481	0.962
Transportation and storage	0.028	0.029	0.027	0.872	-1.189
Water supply, sewerage, waste management activities	0.007	0.003	0.010	0.155	-0.478
Wholesale and retail trade	0.133	0.131	0.135	0.838	-0.052
Observations	1,185	594	591	-	1,185

Note: The table reports the share of job vacancies by industry sector, aggregated at the first level of the UK SIC classification (values do not add up exactly to one because of rounding). *Control* and *AI* report the shares of job vacancies assigned to control and treatment (AI) résumés, respectively. *p*-values are from two-sided tests of proportions comparing means between the control and AI shares. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *Mean AI exposure* is computed for each industry based on Felten et al. (2021), taking the mean of scores for each first level SIC category in our sample.

3.2 Achieved power

We calculate the achieved power in our correspondence study using the 'power' command in Stata. Specifically, we take the means of the strict and broad callbacks measures from the control group in our sample, and perform the following analyses:

Strict callbacks:

power twoproportions .072758, test(chi2) power(0.8) n(1185) n1(591)

Broad callbacks:

power twoproportions .1336717, test(chi2) power(0.8) n(1185) n1(591)

3.3 Results from first wave of applications (Finance, HR and Marketing)

We present below the results for the 'first wave' of applications that were originally pre-registered (i.e. only for the HR, Marketing and Finance job functions).

Table 13 – Treatment effect on strict callbacks (first wave only)

	(1) LPM	(2) Probit	(3) LPM	(4) Probit	(5) LPM	(6) Probit	(7) LPM	(8) Probit
Treatment (AI)	0.040*	0.041*	0.040*	0.039*	-0.000	-0.001	0.011	0.011
	(0.022)	(0.022)	(0.022)	(0.022)	(0.026)	(0.049)	(0.028)	(0.048)
Treatment (AI) × HR					0.029	0.028	-0.001	-0.002
					(0.046)	(0.061)	(0.049)	(0.062)
Treatment (AI) × Marketing					0.090*	0.066	0.081	0.062
					(0.052)	(0.059)	(0.053)	(0.058)
Job function dummies	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Controls	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Observations	585	585	585	585	585	585	575	533

Note: Linear probability model (LPM) and Probit model results (reporting average marginal effects); robust standard errors are reported in parentheses. The dependent variable *Strict callback* is equal to one if the application received a callback for an interview, and zero otherwise. *Treatment (AI)* is equal to one if the résumé includes AI-related skills and zero otherwise. *Job function dummies* are dummies for each of the job functions HR and Marketing with Finance being the reference category. *Controls* include industry fixed-effects, firm size (number of employees), company age, and a location dummy if the job is based in London or the South-East of England. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 14 – Treatment effect on broad callbacks (first wave only)

	(1) LPM	(2) Probit	(3) LPM	(4) Probit	(5) LPM	(6) Probit	(7) LPM	(8) Probit
Treatment (AI)	0.029	0.029	0.030	0.028	-0.001	-0.001	0.017	0.026
	(0.028)	(0.028)	(0.028)	(0.027)	(0.036)	(0.057)	(0.040)	(0.056)
Treatment (AI) × HR					-0.033	-0.030	-0.087	-0.089
					(0.061)	(0.073)	(0.066)	(0.074)
Treatment (AI) × Marketing					0.121*	0.095	0.110	0.075
					(0.065)	(0.071)	(0.067)	(0.071)
Job function dummies	*No*	*No*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Controls	*No*	*No*	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Observations	585	585	585	585	585	585	575	547

Note: Linear probability model (LPM) and Probit model results (reporting average marginal effects); robust standard errors are reported in parentheses. The dependent variable *Broad callback* is equal to one if the application received any positive response from the employer, and zero otherwise. *Treatment* is equal to one if the résumé includes AI-related skills and zero otherwise. *Job function dummies* are dummies for each of the job functions HR and Marketing with Finance being the reference category. *Controls* include industry fixed-effects, firm size (number of employees), company age, and a location dummy if the job is based in London or the South-East of England. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 15 – Treatment effect and job advertisement attributes (first wave only)

	Panel A: Number of skills required					
	(1)	(2)	(3)	(4)	(5)	(6)
	Strict callbacks			Broad callbacks		
Treatment (AI)	0.041*	0.043*	0.043*	0.030	0.032	0.033
	(0.022)	(0.022)	(0.024)	(0.028)	(0.028)	(0.030)
Number of skills required	-0.004	0.009	0.012	-0.001	0.018	0.019
	(0.011)	(0.016)	(0.017)	(0.013)	(0.020)	(0.022)
Treatment (AI) × Number of skills required		-0.023	-0.042*		-0.033	-0.054**
		(0.021)	(0.022)		(0.026)	(0.027)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	580	580	570	580	580	570
	Panel B: Match in skills required					
	(7)	(8)	(9)	(10)	(11)	(12)
	Strict callbacks			Broad callbacks		
Treatment (AI)	0.041*	0.041*	0.038	0.030	0.029	0.025
	(0.022)	(0.022)	(0.024)	(0.028)	(0.028)	(0.030)
Match	0.002	-0.000	-0.003	0.002	-0.007	-0.018
	(0.010)	(0.012)	(0.015)	(0.013)	(0.020)	(0.024)
Treatment (AI) × Match		0.003	0.009		0.012	0.029
		(0.019)	(0.023)		(0.027)	(0.033)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	576	576	566	576	576	566
	Panel C: Expertise					
	(13)	(14)	(15)	(16)	(17)	(18)
	Strict callbacks			Broad callbacks		
Treatment (AI)	0.041*	0.043*	0.042*	0.031	0.034	0.032
	(0.022)	(0.022)	(0.024)	(0.028)	(0.028)	(0.030)
Expertise	-0.013	0.006	0.003	-0.024*	0.002	-0.003
	(0.011)	(0.012)	(0.014)	(0.013)	(0.015)	(0.018)
Treatment (AI) × Expertise		-0.036*	-0.033		-0.050*	-0.039
		(0.021)	(0.023)		(0.026)	(0.029)
Controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	580	580	570	580	580	570

Note: Linear probability model (LPM) with robust standard errors reported in parentheses. The dependent variable *Strict (Broad) callback* is equal to one if the application received a callback, and zero otherwise. *Strict callback* includes only interview invitations; *Broad callback* includes all positive responses. *Treatment* is equal to one if the submitted résumé includes AI skills and zero otherwise. *Number of skills required* is the standardized number of skills mentioned in the job advertisement. *Match* is the standardized proportion of skills mentioned in the job advertisement that are also present in the candidate's résumé. *Expertise* is the standardized Dale-Chall index score based on the skills requested in the job advertisement. *Controls* include dummies for the job functions, a dummy equal to one if the job is based in London or the South-East of England and zero otherwise, and controls for companies' characteristics (number of employees, company age, industry dummies). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 16 – Treatment effect and exposure to AI (first wave only)

	Panel A: Strict callback					
	(1)	(2)	(3)	(4)	(5)	(6)
	LPM	Probit	LPM	Probit	LPM	Probit
Treatment (AI)	0.038*	0.039*	0.054**	0.063**	0.049**	0.057**
	(0.022)	(0.023)	(0.024)	(0.028)	(0.023)	(0.027)
AI exposure	0.009	0.010	0.022*	0.029*	0.021*	0.028*
	(0.009)	(0.009)	(0.012)	(0.016)	(0.012)	(0.016)
Treatment (AI) × AI exposure			-0.026	-0.032	-0.025	-0.032
			(0.018)	(0.020)	(0.018)	(0.020)
Controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Joint test: AI + AI × Exposure			0.242	0.164	0.331	0.256
Observations	585	585	585	585	575	575
	Panel B: Broad callback					
	(7)	(8)	(9)	(10)	(11)	(12)
	LPM	Probit	LPM	Probit	LPM	Probit
Treatment (AI)	0.026	0.027	0.039	0.043	0.032	0.034
	(0.028)	(0.028)	(0.030)	(0.032)	(0.030)	(0.032)
AI exposure	0.013	0.013	0.023	0.026	0.022	0.024
	(0.011)	(0.011)	(0.016)	(0.017)	(0.016)	(0.017)
Treatment (AI) × AI exposure			-0.021	-0.024	-0.021	-0.022
			(0.022)	(0.023)	(0.022)	(0.023)
Controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*
Joint test: AI + AI × Exposure			0.543	0.492	0.732	0.681
Observations	585	585	585	585	575	575

Note: Linear probability model (LPM) and Probit model (reporting average marginal effects) results; robust standard errors are reported in parentheses. The dependent variable *Strict (Broad) callback* is equal to one if the application received a callback, and zero otherwise. *Strict callback* includes only interview invitations; *Broad callback* includes all positive responses. *Treatment (AI)* is equal to one if the résumé includes AI skills and zero otherwise. *AI exposure* is the measure of exposure to AI at the industry level from Felten et al. (2021). *Controls* include dummies for the job functions, number of employees, company age, and a dummy if the job is based in London or the South-East of England. *Joint test: AI + AI × Exposure* presents the *p*-values for a Wald test of joint significance that the *Treatment (AI)* and the interaction *Treatment (AI) × AI exposure* are jointly equal to zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3.4 Additional figures from correspondence study

Figure 1 – Cumulative distribution functions (CDF) of explanatory variables

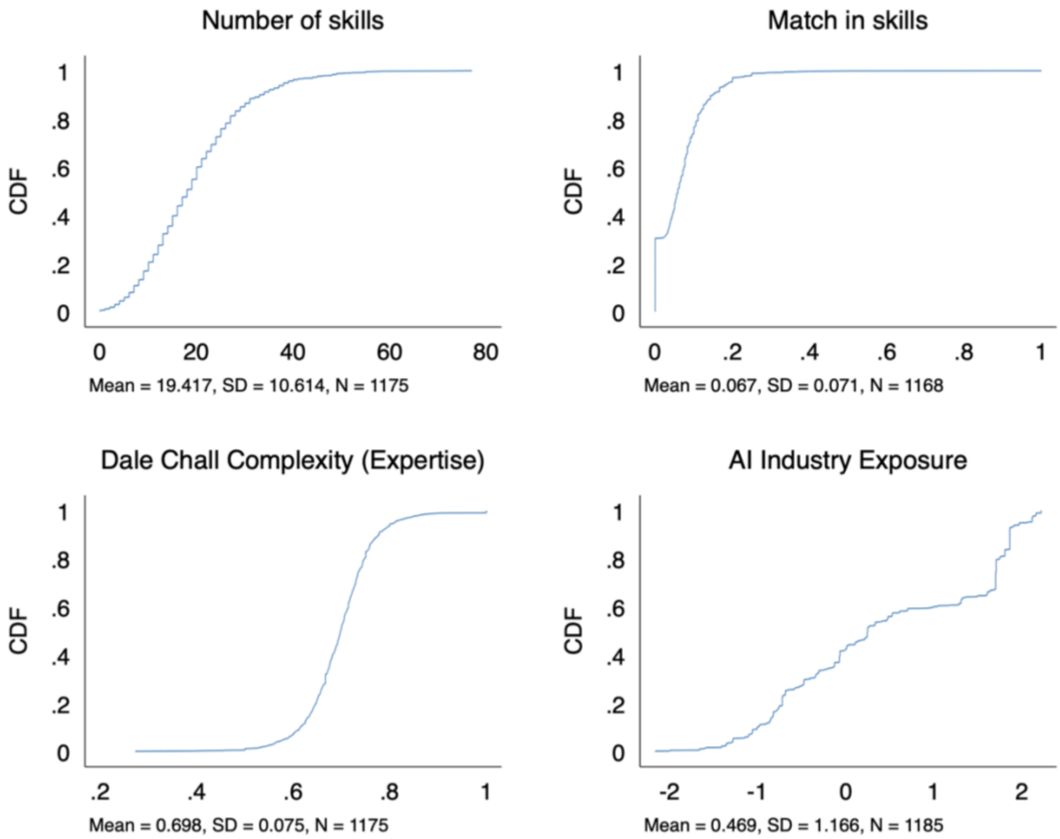


Figure 2 – Distribution of skills required in job postings by job function

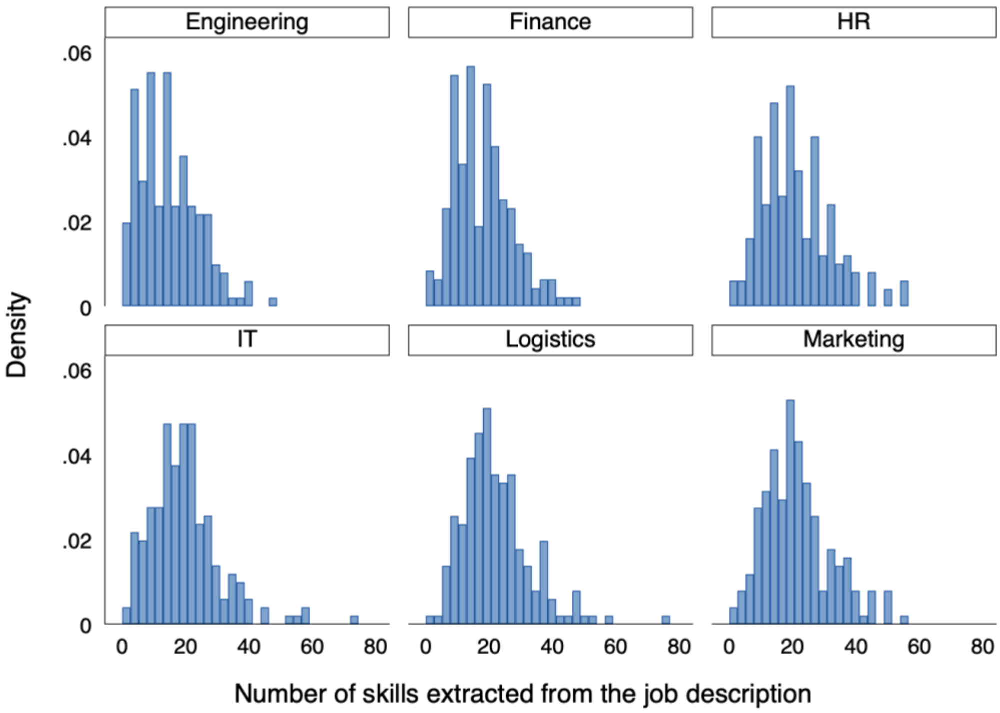


Figure 3 – Distribution of the match in skills required in job postings, and present in the candidate's résumé, by job function

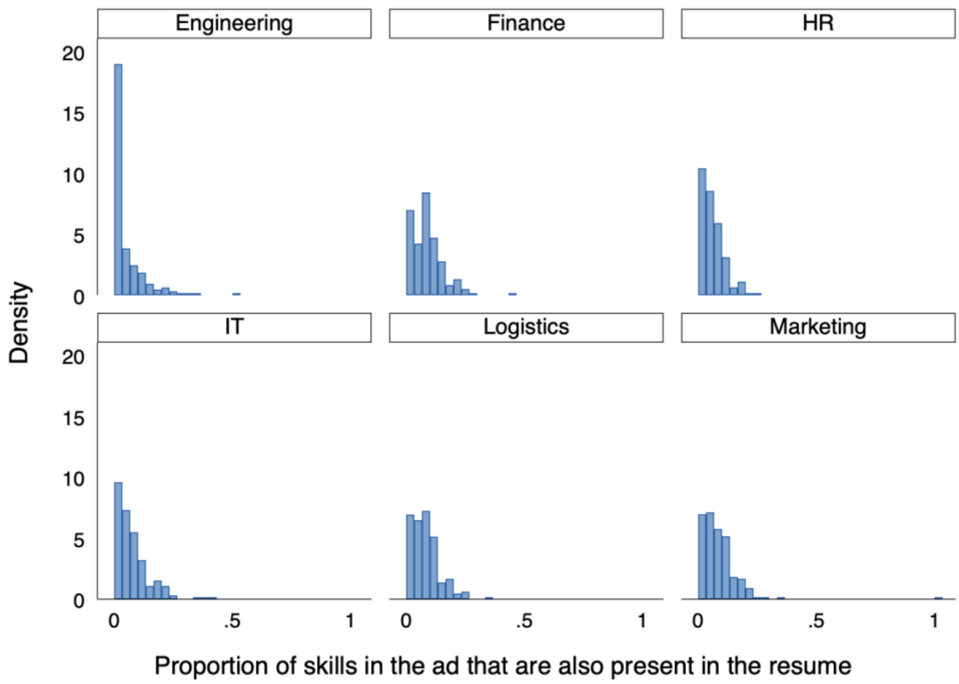


Figure 4 – Distribution of Dale-Chall Complexity measure (expertise) in job postings by job function

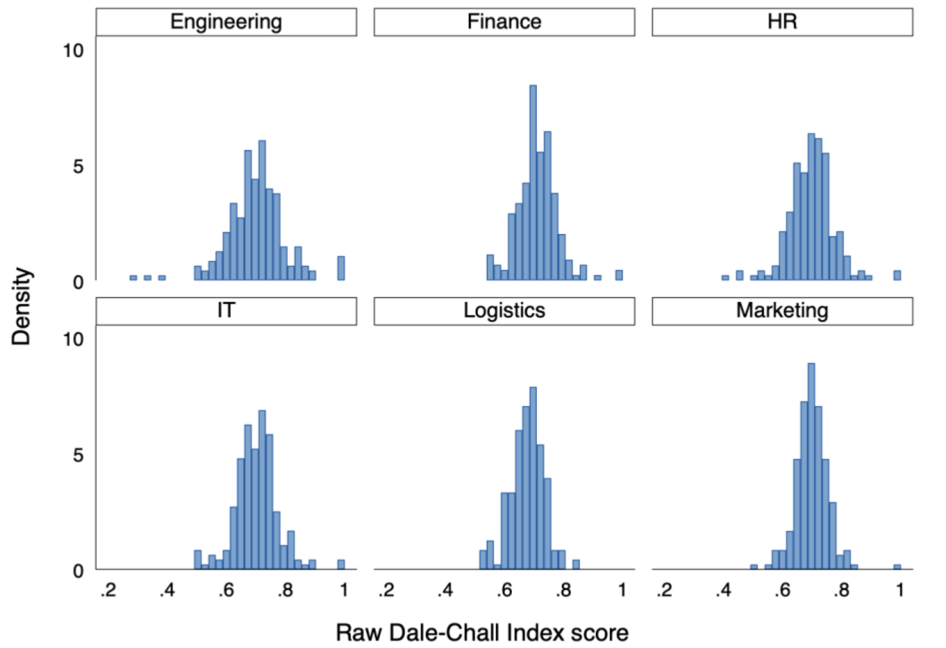


Figure 5 – Distribution of exposure to AI at the industry level (measure by Felten et al. (2021)) by job function

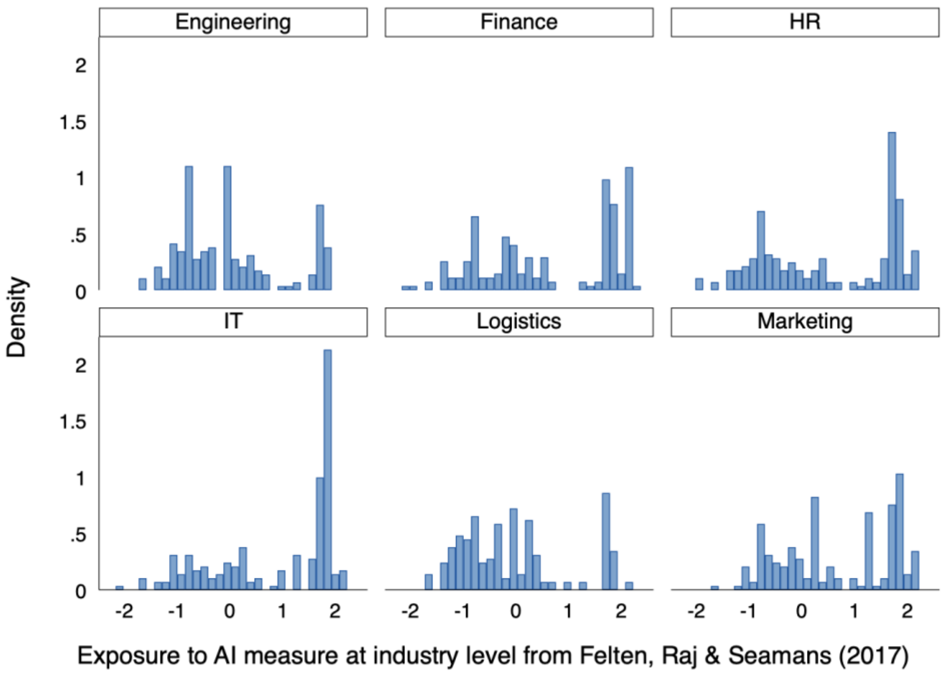
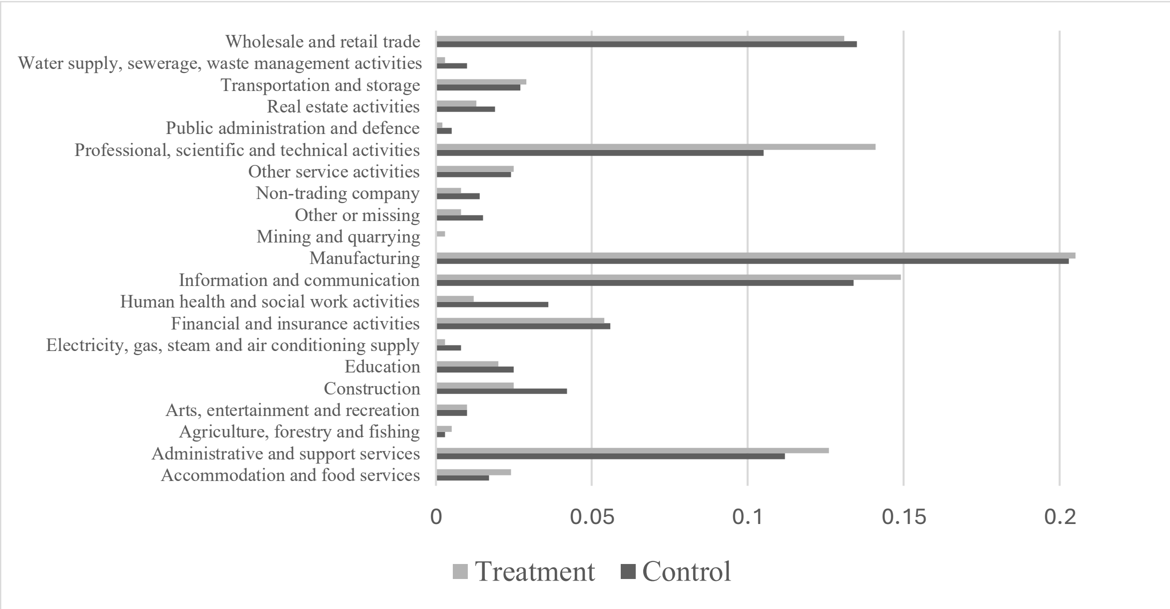


Figure 6 – Proportion of job advertisements by industry and treatment group



Note: Each bar represents the proportion of job roles in the sample advertised by companies in each sector. Sectors are aggregated at the highest level in the UK SIC classification.

3.5 Expert survey

AI Audit Experiment - Marketing

Start of Block: Consent

[Intro_text]

Welcome to our research study!

Study on human judgement and decision-making This research is being conducted by *[redacted for peer review]*. The intention of this study is to better understand human decision-making.

Compensation
The basic compensation fee for filling in our survey is £1.80. You can also earn additional compensation if you get certain responses right – you will be informed about this extra compensation during the survey.

Participation, Benefits and Risks
Participation in this research study is completely voluntary. You have the right to withdraw at any time or refuse to participate entirely. If you desire to withdraw, please simply close your internet browser. Risks are minimal for involvement in this study. It is very unlikely that answering these questions affects you emotionally or otherwise.

Confidentiality and Questions
All data obtained from participants will be kept anonymous. There will be no record that links the data collected from you with any personal data from which you could be identified (e.g., your name, address, email, etc.). Once anonymized, these data may be made available to researchers via accessible data repositories and possibly used for novel purposes. The data will be stored for at least 10 years. If you have any questions or comments or if you need support, please contact us by email: *[redacted for peer review]*.

Notes on data protection
All responses are collected anonymously. No personal data will be collected and analyzed that would allow you to be identified.
Your answers will be stored on servers within the EU and will only be analyzed in the context of this research project. No data will be forwarded to the questionnaire provider Qualtrics or other third parties.
Data transmission is encrypted and data security is certified with ISO27001. In this context, we would like to draw your attention to the following information from questionnaire provider Qualtrics:
Terms of Use https://www.qualtrics.com/terms-of-service/
Privacy policy https://www.qualtrics.com/privacy-statement/
Security declaration https://www.qualtrics.com/security-statement/
You also have the right to contact the official data protection officer at *[redacted for peer review]*. You may raise with her your questions or concerns regarding compliance with the

Privacy Policy, or complain about non-compliance. In this case, please contact the data protection officer of *[redacted for peer review]*.

Page Break

Q380 Before you proceed to the survey, please verify the captcha below.

Time_introtext Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q1.1 In order to proceed you need to confirm following questions:

	yes (1)	no (2)
I am 18 years or older. (1)	◯	◯
I have read and understood the information above. (2)	◯	◯
I want to take part in this research. (3)	◯	◯

Time_Q1.1 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Consent

Start of Block: Statement do not consent

JS

Q106 *As you do not wish to participate in this study, please return your submission on Prolific by selecting the 'Stop without completing' button.*

End of Block: Statement do not consent

Start of Block: Questions for screener validation

Q2.1 Do you have any experience in making hiring decisions (i.e. have you been responsible for hiring job candidates)?

◯ Yes (1)

◯ No (2)

Time_q2 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q471 In which of the following sectors are you primarily working now or did used to work in the past? (Select all that apply).

- ☐ Agriculture, Food and Natural Resources (1)
- ☐ Architecture and Construction (4)
- ☐ Arts (5)
- ☐ Business Management & Administration (6)
- ☐ Education & Training (7)
- ☐ Finance (8)
- ☐ Government & Public Administration (9)
- ☐ Medicine (10)
- ☐ Hospitality & Tourism (11)
- ☐ Information Technology (12)
- ☐ Legal (13)
- ☐ Policing (14)
- ☐ Military (15)
- ☐ Manufacturing (16)
- ☐ Marketing & Sales (17)
- ☐ Retail (18)
- ☐ Science, Technology, Engineering & Mathematics (19)
- ☐ Social Sciences (20)

☐ Transportation, Distribution & Logistics (21)

☐ Other (22)

Q472 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Questions for screener validation

Start of Block: Statement inconsistent screener response

JS

inconsistence_text *Thank you for answering these questions. You are ineligible for this study, as you have provided information which is inconsistent with your Prolific prescreening responses. This study is for people with hiring experience and with some experience in Marketing. Please return your submission on Prolific by selecting the 'Stop without completing' button.*

End of Block: Statement inconsistent screener response

Start of Block: Verification and Prolific ID

Q3.2 Please enter or confirm your Prolific ID here:

____

Time_Q3 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Verification and Prolific ID

Start of Block: Introduction

Q449 Introduction
This study consists of four parts.

In Part I and Part II, we will show you slightly different CVs (one in each part). These CVs were designed to apply to a job vacancy in the UK. We will ask you to assess the skills and employment chances of its (fictitious) owners.
In Part III, we will ask you general questions about both CVs.

To conclude, we will ask you a few questions about yourself in Part IV.

Q451 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Introduction

Start of Block: Intro Part I

part1_text Part I
Suppose you were recruiting a candidate for an entry-level ${e://Field/occupation} position.

Now, imagine you receive the following CV of a potential candidate who applies for an entry-level position. Please take as long as you need to familiarize yourself with this first CV. You will be asked to answer some questions about this CV.

*** Embedded variable *occupation* ∈ {marketing, finance, IT, engineering, supply chain, HR} ***

Q454 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Intro Part I

Start of Block: CV control display

Q419 We will first ask you about different types of skills from this potential candidate based on the information provided in this CV. Please, look and read it carefully first.

*** CV was presented here

Q453 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV control display

Start of Block: CV control skill evaluation

evaluation_control ${lm://Field/3} **How would you assess the candidate's ${lm://Field/1}?**

	Very low (42)	Somewhat low (43)	Neither low nor high (44)	Somewhat high (45)	Very high (46)
${lm://Field/2} (1)	○	○	○	○	○

	Field 1	Field 2	Field 3
1	social skills	Social skills	Social skills are employed in interpersonal settings. These skills often include skills such as communication, teamwork, collaboration, negotiation.
2	cognitive skills	Cognitive skills	Cognitive skills are employed to process incoming information. These skills often include analysis, research skills, problem solving, critical thinking, math and statistics.
3	basic digital skills	Basic digital skills	Basic digital skills are employed to complete simple tasks using rudimentary digital devices and applications. These skills often include digital communication, creating and managing documents/spreadsheets, installing software updates, etc.
5	advanced digital skills	Advanced digital skills	Advanced digital skills are employed for application of theoretical and analytical knowledge when using technology. These skills often include data analysis, coding, web and app development, etc.
6	AI-related digital skills	AI-related digital skills	AI-related digital skills are employed to apply the methods and concepts of artificial intelligence in business functions. Such skills often include advanced programming and algorithm development, and applying machine learning methods.
7	special skills, based on text above	Special skills	Special skills. For this question, simply click on the option 'very low' (this is an attention check).

Q421 Here is the same CV again for your reference.

*** CV was presented here

Q420

Q456 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV control skill evaluation

Start of Block: Transition Firm Decision Question 1

Q442 We will now ask you a question related to firms' decisions to provide a job interview for a candidate with this first CV you just saw.

Q457 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Transition Firm Decision Question 1

Start of Block: CV control hire beliefs

JS

Q414 **Job interview decision** In a real-life experiment conducted in 2021, we sent randomly selected companies the version of the CV that you just saw. All companies were looking to fill a vacancy for an entry-level ${e://Field/occupation} role.
Please estimate: Out of 50 companies, how many companies invited the candidate with this CV for an interview?
Think carefully, because you can earn additional bonus for your answer!
If your estimate deviates from the actual statistic from our study by less than 3 integers, you will earn a bonus compensation of £0.10, on top of your participation fee as announced on Prolific. This additional payment will be sent to you within 72 hours of your submission being approved. (If you encounter any issues with this payment, feel free to email us at *[redacted for peer review]*). If you do not answer correctly, you will not receive the additional bonus.

*** Embedded variable *occupation* ∊ {marketing, finance, IT, engineering, supply chain, HR} ***

---

hire_control How many interview invites did the candidate with this CV receive for 50 applications sent (from 0 to 50)?

____

Page Break

asset_control In your opinion, how well do the candidate's qualification and experience fit an entry-level ${e://Field/occupation} position? Indicate your answer on the scale from underqualified to overqualified.

	Underqualified (1)	Somewhat underqualified (2)	Neither under- or over-qualified (3)	Somewhat qualified (4)	Overqualified (5)
How would you evaluate this candidate's qualifications for the entry-level ${e://Field/occupation} position? (44)	○	○	○	○	○

*** Embedded variable *occupation* ∊ {marketing, finance, IT, engineering, supply chain, HR} ***

Q458 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV control hire beliefs

Start of Block: Start 2nd Part

Q381 Part IIWe will now show you the second, slightly different and fictitious CV and ask you the same questions as in Part I. Please answer these questions about the CV that you will see on the next page. Suppose you were recruiting a candidate for an entry-level ${e://Field/occupation} position.   Now, imagine you receive the following CV of a potential candidate who applies for an entry-level position. Please take as long as you need to familiarize yourself with this CV. You will be asked to answer some questions about this CV.

*** Embedded variable *occupation* ∊ {marketing, finance, IT, engineering, supply chain, HR} ***

Q459 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Start 2nd Part

Start of Block: CV ai display

Q422 We will first ask you about different types of skills from this potential candidate based on the information provided in this CV. Please, look and read it carefully first.

Q392

Q460 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV ai display

Start of Block: CV ai skills evaluation

evaluation_AI ${lm://Field/3} **How would you assess the candidate's ${lm://Field/1}?**

	Very low (42)	Somewhat low (43)	Neither low nor high (44)	Somewhat high (45)	Very high (46)
${lm://Field/2} (1)	o	o	o	o	o

	Field 1	Field 2	Field 3
1	social skills	Social skills	Social skills are employed in interpersonal settings. These skills often include skills such as communication, teamwork, collaboration, negotiation.
2	cognitive skills	Cognitive skills	Cognitive skills are employed to process incoming information. These skills often include analysis, research skills, problem solving, critical thinking, math and statistics.
3	basic digital skills	Basic digital skills	Basic digital skills are employed to complete simple tasks using rudimentary digital devices and applications. These skills often include digital communication, creating and managing documents/spreadsheets, installing software updates, etc.
5	advanced digital skills	Advanced digital skills	Advanced digital skills are employed to apply theoretical and analytical knowledge when using technology. These skills often include data analysis, coding, web and app development, etc.
6	AI-related digital skills	AI-related digital skills	AI-related digital skills are employed to apply the methods and concepts of artificial intelligence in business functions. Such skills often include advanced programming and algorithm development, and applying machine learning methods.
7	special skills, based on text above	Special skills	Special skills. For this question, simply click on the option 'very low' (this is an attention check).

Randomize loop order

Q424 Here is the same CV again for your reference.

*** CV was presented here

Q423

Q461 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV ai skills evaluation

Start of Block: Transition Firm Decision Question 2

Q444 We will now ask you a question related to firms' decision to provide a job interview for a candidate with this second CV that you just saw.

Q462 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Transition Firm Decision Question 2

Start of Block: CV ai hire beliefs

JS

Q404 **Job interview decision** In a real-life experiment conducted in 2021, we sent randomly selected companies this version of the CV that you just saw. All companies were looking to fill a vacancy for an entry-level ${e://Field/occupation} role. Please estimate: Out of 50 companies, how many companies invited the candidate with this CV for an interview? **Think carefully, because you can earn additional bonus for your answer!** If your estimate deviates from the actual statistic from our study by less than 3 integers, you will earn a bonus compensation of £0.10, on top of your participating fee as announced on Prolific. This additional payment will be sent to you within 72 hours of your submission being approved. (If you encounter any issues with this payment, feel free to email us at *[redacted for peer review]*). If you do not answer correctly, you will not receive the additional bonus.

*** Embedded variable *occupation* ∈ {marketing, finance, IT, engineering, supply chain, HR} ***

---

*

hire_AI How many interview invites did the candidate with this CV receive for 50 applications sent (from 0 to 50)?

____

Page Break

asset_AI In your opinion, how well do the candidate's qualification and experience fit an entry-level ${e://Field/occupation} position? Indicate your answer on the scale from underqualified to overqualified.

	Underqualified (1)	Somewhat underqualified (2)	Neither under- or over-qualified (3)	Somewhat overqualified (4)	Overqualified (5)
How would you evaluate this candidate's qualifications for the entry-level ${e://Field/occupation} position? (44)	○	○	○	○	○

*** Embedded variable *occupation* ∈ {marketing, finance, IT, engineering, supply chain, HR} ***

Q463 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: CV ai hire beliefs

Start of Block: Part III Both CVs

part2_text Part III In this part, we will ask you a few more questions about the two CVs you saw in Part I and Part II. The first question also includes a bonus payment for correct answers.

Page Break

recall_CV The only difference between the two CVs you saw was that one of them included skills and experiences related to Artificial Intelligence. ${e://Field/AI_skills} Which CV included this additional information? Bonus compensation: If you give the correct answer, you will earn an additional £0.10 on top of your participation fee.

◯ The first CV included the additional information about Artificial Intelligence (1)

◯ The second CV included the additional information about Artificial Intelligence (2)

*** Values of the variable *AI_skills* conditional on *occupation*

marketing	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using AI / Machine Learning algorithms for personalised customer recommendations (...)".
engineering	This included experience such as the candidate having "[d]eveloped a machine learning (ML) approach to anomaly detection in the production process".
finance	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using AI / Machine Learning algorithms for fraud detection, risk modeling, and financial forecasts and planning (...)".
IT	This included skills such as "Azure AI" and a bullet point stating the candidate had experience "Implementing AI knowledge mining solutions with Azure AI".
HR	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using machine learning algorithms and tools for AI-supported virtual HR assistants (...)".
Supply Chain	This included skills such as "Machine Learning" and a bullet point stating the candidate had experience "Using AI tools to develop dynamic supply planning (...)".

Page Break

Q213 In the next questions, we will ask you about the authenticity of the CVs and about the applicability of AI skills for an entry-level ${e://Field/occupation} job and for the workforce, in general.

*** Embedded variable *occupation* ∈ {marketing, finance, IT, engineering, supply chain, HR} ***

evaluation Please rate your agreement with the following statements on a scale from "strongly disagree" to "strongly agree":

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
The CVs seemed authentic (4)	○	○	○	○	○
The AI skills (${e://Field/AI_skills}) would significantly improve a candidate's ability to execute the ${e://Field/occupation} assistant role (5)	○	○	○	○	○
Young professionals and graduates should acquire AI skills (13)	○	○	○	○	○
For this question, simply check the middle option (this is an attention check) (6)	○	○	○	○	○
The other skills in the CV (excluding the AI skills) were relevant for a(n) ${e://Field/occupation} assistant role (12)	○	○	○	○	○

*** Values of the variable *AI_skills* conditional on *occupation*

marketing	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using AI / Machine Learning algorithms for personalised customer recommendations (...)".
engineering	This included experience such as the candidate having "[d]eveloped a machine learning (ML) approach to anomaly detection in the production process".

finance	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using AI / Machine Learning algorithms for fraud detection, risk modeling, and financial forecasts and planning (...)".
IT	This included skills such as "Azure AI" and a bullet point stating the candidate had experience "Implementing AI knowledge mining solutions with Azure AI".
HR	This included skills such as "IBM Watson" and a bullet point stating the candidate had experience "Using machine learning algorithms and tools for AI-supported virtual HR assistants (...)".
Supply Chain	This included skills such as "Machine Learning" and a bullet point stating the candidate had experience "Using AI tools to develop dynamic supply planning (...)".

*** Embedded variable *occupation* ∈ {marketing, finance, IT, engineering, supply chain, HR} ***

---

time_Q6.4 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q450
AI-related digital skills are employed to apply the methods and concepts of artificial intelligence in business functions. Such skills often include advanced programming and algorithm development, and applying machine learning methods.
How applicable do you think AI-related skills are for the following organizational functions? Please drag and drop to rank the organizational functions from 1 (Most applicable) to 6 (Least applicable).
______ Marketing (1)
______ HR (2)
______ Finance (3)
______ IT (4)
______ Supply Chain (5)
______ Engineering (6)

Q464 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Part III Both CVs

Start of Block: Transition message

transition_text **Part IV**
The first three parts of this study are now over. In the final fourth section, which should only take a few minutes, we will ask you a few questions about yourself.

time_transition Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Transition message

Start of Block: Demographics

Q7.1 What sex were you assigned at birth, such as on an original birth certificate?

- ◯ Male (1)
- ◯ Female (2)
- ◯ Prefer not to say (3)

JS *

Q7.2 How old are you?

_______ years (1)

Q7.3 What is the highest degree or level of education you have completed?

- ◯ Some high school, no diploma (9)
- ◯ High school graduate (10)
- ◯ Some college, no degree (11)
- ◯ Associate degree (12)
- ◯ Bachelor's degree (13)
- ◯ Master's degree (14)
- ◯ Professional degree (15)
- ◯ Doctorate degree (16)

---

time_Q7 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q7.4 Please describe your work.

- ○ Employee of a for-profit company or business or of an individual, for wages, salary, or commissions (1)
- ○ Employee of a not-for-profit, tax-exempt, or charitable organization (2)
- ○ Local government employee (city, county, etc.) (3)
- ○ State government employee (4)
- ○ Federal government employee (5)
- ○ Self-employed in own not-incorporated business, professional practice, or farm (6)
- ○ Self-employed in own incorporated business, professional practice, or farm (7)
- ○ Working without pay in family business or farm (8)
- ○ None of the above (9)

Q452 Please indicate which industry you work in.

- Accommodation and food services (1)
- Agriculture (2)
- Arts, culture, entertainment and design (3)
- Construction (4)
- Education and training (5)
- Fashion (6)
- Financial and insurance (7)
- Fitness and sport (8)
- Healthcare and social assistance (9)
- Marketing and advertising (10)
- Mining (11)
- Other (please describe) (12) ____

Q7.5 What is your current employment status?

- Employed (full time) (1)
- Employed (part time) (2)
- Studying (4)
- Unemployment (5)
- Retired (6)
- Other (please describe) (7) ____

Q7.6 How many years of work experience do you have?
_______ years (1)

Q415 How many times were you involved in hiring decisions during your professional career?

- ○ 0-3 (1)
- ○ 4-10 (2)
- ○ 11-20 (3)
- ○ 21 or more (4)

---

time_Q7.6 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Demographics

Start of Block: AI Personal Experience Questions

Q431 **Please identify on a scale from 'none' to 'very high' the level of your experience with AI (artificial intelligence).**

	None (1)	Very low (3)	Somewhat low (4)	Neither low nor high (5)	Somewhat high (6)	Very high (7)
General work experience (1)	○	○	○	○	○	○
General knowledge of AI topics (3)	○	○	○	○	○	○

Q470 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q433 Does your current employer provide courses or encourage employees to acquire AI skills?

- ◯ Yes (1)
- ◯ No (3)
- ◯ I don't know (4)

Q469 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Q434 Have you personally experienced AI adoption in your field?

- ◯ Yes (1)
- ◯ No (2)
- ◯ I don't know (4)

Q465 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: AI Personal Experience Questions

Start of Block: Subjects comments

Q8.1 What did you think of this survey? Do you have any comments for us?

____

____

____

____

____

---

time_Q8 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Q416 Thank you for taking part in our study. Please click on the button to end your participation. You will receive your payment and the additional bonus (if applicable) automatically within 72 hours.

Q466 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Subjects comments

Start of Block: End

end_text Thank you for taking the time to participate in our study. Please click on "Next" to finish the survey and receive your payment.

Q467 Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: End

References

Alekseeva, L., Azar, J., Giné, M., Samila, S., & Taska, B. (2021). The Demand for AI Skills in the Labor Market. *Labour Economics*, *71*(102002). https://doi.org/10.1016/j.labeco.2021.102002

Bachinskiy, A. (2019). *The growing impact of ai in financial services: Six examples*. Towards Data Science, Inc.

Baker, M. (2020). *Ai shows value and gains traction in HR* [Gartner].

Balakrishnan, T., Chui, M., Hall, B., & Henke, N. (2020). *Global survey: The state of AI in 2020.* McKinsey Global Publishing.

Bryan, J. (2018). *How ai will transform financial management applications.* Gartner, Inc.

Costello, K. (2020). *Gartner's Top 10 Strategic Technology Trends for 2021*.

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217. https://doi.org/10.1002/smj.3286

Forrester Inc. (2019). *The Forrester New Wave™: Conversational AI for Customer Service, Q2 2019.* https://www.forrester.com/report/The-Forrester-New-Wave-Conversational-AI-For-Customer-Service-Q2-2019/RES144416

Forrester Inc. (2021). *The Forrester Wave™: Cognitive Search, Q3 2021*. https://www.forrester.com/report/the-forrester-wave-tm-cognitive-search-q3-2021/RES161551

Gartner. (2019). *The Top 8 Supply Chain Technology Trends*.

Gartner, Inc. (2019). *Gartner Identifies Three Most Common AI Use Cases in HR and Recruiting*.

Håkanson, C., Lindqvist, E., & Vlachos, J. (2021). Firms and Skills: The Evolution of Worker Sorting. *Journal of Human Resources*. https://doi.org/10.3368/jhr.56.2.0517-8801R2

Heinz, M., & Schumacher, H. (2017). Signaling cooperation. *European Economic Review*, *98*, 199–216. https://doi.org/10.1016/j.euroecorev.2017.06.017

McKinsey & Company. (2021). *Succeeding in the AI supply-chain revolution*.

McKinsey Global Institute. (2019). *Driving impact at scale from automation and AI*.

Pemberton, C. (2018). *How ai will drive transformative change in marketing*. Gartner, Inc.

Verhaest, D., Bogaert, E., Dereymaeker, J., Mestdagh, L., & Baert, S. (2018). Do Employers Prefer Overqualified Graduates? A Field Experiment. *Industrial Relations*, *57*(3), 361–388. https://doi.org/10.1111/irel.12212

Wang, X. (2018). *Nine ai marketing use cases that have the potential to deliver business value*. Forrester Research, Inc.