Supplementary Online Appendices

for

Forced to Change? Media Exposure of Labor Issues and Firm Artificial Intelligence Investment

APPENDIX A. LITERATURE REVIEW OF AI–HUMAN RELATIONS IN INFORMATION SYSTEMS (IS) RESEARCH

We looked for all articles that contained the keywords “artificial intelligence,” “robot,” “automation,” “augmentation,” “machine learning,” and their close variants as well as acronyms (e.g., “AI” for “artificial intelligence” and “ML” for “machine learning”) published in *Information Systems Research* and *MIS Quarterly* starting from 2008. We excluded studies on the technical dimension of AI algorithm development and focused on only conceptual and empirical papers. We then manually checked whether an article related to AI–human relations. Table A1 lists the details of the literature review.

Table A1. Literature Review on AI–Human Relations in IS Research

Reference	**Journal**	**Unit of Analysis**	**Theory Frame**	**Dependent Variable**	**Independent Variable**	**Sample/Context**	**Method**	**Key Takeaways**
Bauer et al. (2023)	ISR	Individual	—	Posterior belief	Prior beliefs	607 participants on Prolific	Lab experiment	This paper shows the potential downstream consequences of the broad employment of modern explainable AI methods. Side effects of mental model adjustments present a potential risk of manipulating user behavior, promoting discriminatory inclinations, and increasing noise in decision making.

Berente et al. (2021)	MISQ	Conceptual	—	—	—	—	—	The paper discusses how the frontiers of AI have changed over time, and speculates about future frontiers in managing AI.
Fügener et al. (2022)	ISR	Individual	—	—	—	902 subjects from Open Science Foundation	Lab experiment	Humans and AI who work together can outperform the AI that outperforms humans when it works on its own. However, the combined performance improves only when the AI delegates work to humans but not when humans delegate work to the AI.
Ge et al. (2021)	ISR	Individual	—	Robo-advisor adoption probability	Previous investment characteristics	4,374 lenders from a P2P lending company with the complete history of their bids across from January 2015 to June 2016	Field experiment	Investors who need more help from robo-advisors are less likely to adopt such services; investors tend to adjust their usage of the service in reaction to recent robo-advisor performance.
Gnewuch et al. (2024)	ISR	Individual	—	Customer communication style, employee workload	Human involvement disclosure	8,966 customers	Field experiment	Disclosing human involvement before or during an interaction with the hybrid service agent leads customers to adopt a more human-oriented communication style.
Han et al. (2023)	ISR	Individual	—	—	—	Participants from U.S. universities	Lab experiment	AI-expressed positive emotion can influence customers via dual

								pathways: an affective pathway of emotional contagion and a cognitive pathway of expectation–disconfirmation.
Jain et al. (2021)	ISR	Conceptual	—	—	—	—	—	This editorial paper presents a framework to guide research on AI augmentation and advocates for the important implications of AI augmentation for the future of work, organizations, and society.
Jussupow et al. (2021)	ISR	Individual	—	—	—	68 novice and 12 experienced physicians	Lab experiment	Physicians use second-order cognitive processes — namely, metacognitions —, to monitor and control their reasoning while assessing AI advice. These metacognitions determine whether physicians are able to reap the full benefits of AI.
Li et al. (2021)	MISQ	Organizational	Upper echelons theory	AI orientation	Chief information officers' presence	1,454 public firms in China	Archival data analysis	The presence of a CIO positively influences AI orientation; board educational diversity, R&D experience, and AI experience positively moderate the CIO's effect on AI orientation.
Lou and Wu (2021)	MISQ	Organizational	Resource-	Number of	AI innovation	644 firms from	Archival	A firm's discovery of

			based view	drugs	capability	2020 to 2019	data analysis	new drug–target pairs for preclinical studies. Developing and improving AI tools is an iterative process requiring synthesizing inputs from both AI and domain experts.
Lysyakov and Viswanathan (2023)	ISR	Individual	Protection motivation theory	The number of contests per designer per day, number of submissions per designer per contest; second set, emotions and complexity of designs	AI introduction treatments	Crowdsourcing platform	Field experiment	In response to the threat from the introduction of an AI system; in response to the threat from the introduction of an AI system.
Rai et al. (2019)	MISQ	Conceptual	—	—	—	—	—	Emerging next-generation digital platforms, arising from the application of AI technologies, offer new possibilities for the relationship between humans and machines to perform tasks on digital platforms and for the effective design and governance of platforms.
Schanke et al. (2021)	ISR	Individual	—	The probability of conversion	Anthropomorphic treatments	323 subjects who initiated a conversation with the chatbot between November 16 and December 31 of 2018	Field experiment	Chatbot anthropomorphism is beneficial for transaction outcomes. As a chatbot becomes more human-like, consumers shift to a fairness evaluation or negotiating mindset.

Sturm et al. (2021)	MISQ	Organizational	—	—	—	—	Agent-based simulation	ML can reduce an organization’s demand for human exploratory learning that is aimed at uncovering new ideas. Adjustments to ML systems made by humans are largely beneficial. Reliance on knowledge created by ML systems can facilitate organizational learning in turbulent environments.
Teodorescu et al. (2021)	MISQ	Conceptual	—	—	—	—	—	The paper introduces a typology of human–ML augmentation for fairness consisting of four quadrants: reactive oversight, proactive oversight, informed reliance, and supervised reliance.
Van den Broek et al. (2021)	MISQ	Conceptual	—	—	—	ML vendor “NeuroYou”	Field experiment	Contrary to common views that imply an opposition between ML and domain expertise, this study foregrounds their interdependence and as such shows the dialectic nature of developing ML.
Zhang et al. (2024)	ISR	Organizational	—	Value-added for industries	IT capital, non-IT capital, high-education labor, middle-education labor, low-education labor	BEA and BLS data	Archival data analysis	IT is a complement (substitute) for high-education labor in industries with lower (higher) AI exposure and remains a net

								substitute for low- and middle-education labor, regardless of their AI exposure.
Our study	ISR	Organizational	Institutional theory	AI investment	Media labor issue exposure	U.S. public firms from 2010 to 2018	Archival data analysis	Media exposure of labor issues exerts reputational pressure on firms, heightening their awareness of the necessity and urgency to address these problems. Consequently, firms are driven to invest in AI as a potential solution, given its capabilities in both augmenting and automating human tasks.

Notes: Unit of analysis is classified as conceptual, individual, and organizational. Acronyms: ISR = *Information Systems Research*; MISQ = *MIS Quarterly*; AI = artificial intelligence; ML = machine learning.

APPENDIX B: SAMPLE AND DATA

Table B1. Sample Construction Process for Baseline Analyses (Table 2)

Sample construction steps	Number of observations
U.S. firms recorded in Compustat from 2010 to 2018	71,779
Less: Observations not covered by Burning Glass Technologies	(48,666)
Less: Observations with missing values of controls	(4,067)
Less: Singleton observations detected	(287)
Sample for OLS regression with controls	18,759

Table B2. Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 *AI investment (log)*	1.00																
2 *Non-AI investment (log)*	0.36	1.00															
3 *Traditional automation investment (log)*	0.40	0.35	1.00														
4 *Non-AI IT investment (log)*	0.43	0.90	0.38	1.00													
5 *Media labor issue exposure*	0.28	0.29	0.21	0.25	1.00												
6 *Firm size (log)*	0.28	0.66	0.32	0.58	0.33	1.00											
7 *ROA*	0.03	0.13	0.05	0.12	0.04	0.28	1.00										
8 *Resource slack*	–0.04	–0.17	–0.04	–0.15	–0.06	–0.28	–0.06	1.00									
9 *R&D expenditure (log)*	0.13	–0.08	0.13	0.11	–0.01	–0.08	–0.04	0.11	1.00								
10 *Blank R&D reporting*	–0.07	–0.01	–0.10	–0.03	0.00	0.05	0.03	–0.24	0.21	1.00							
11 *Labor intensity*	–0.05	–0.06	–0.04	–0.11	–0.02	–0.07	–0.20	0.55	–0.03	–0.07	1.00						
12 *Employee ownership*	0.08	0.12	0.12	0.13	0.10	0.19	0.03	–0.07	0.05	0.01	–0.03	1.00					
13 *Collective bargaining*	0.05	0.13	0.18	0.14	0.14	0.28	0.06	–0.07	0.10	0.00	–0.06	0.27	1.00				
14 *Technology governance orientation*	0.16	0.10	0.12	0.17	0.04	0.05	–0.01	0.15	0.21	–0.23	0.02	0.03	0.01	1.00			
15 *Managerial ability*	0.15	0.11	0.06	0.13	0.13	0.03	0.04	0.04	–0.01	–0.14	–0.03	–0.05	–0.09	0.13	1.00		
16 *Common ownership with media*	0.10	0.28	0.13	0.28	0.07	0.35	0.13	0.00	0.07	–0.06	–0.05	0.04	0.13	0.18	0.03	1.00	
17 *Industry AI exposure*	0.14	–0.03	–0.05	0.10	–0.12	–0.24	–0.05	–0.05	0.17	0.09	–0.05	–0.06	–0.27	0.06	0.09	–0.10	1.00

APPENDIX C. ROBUSTNESS CHECKS

C1. Alternative Independent Variable Measures

We first check the robustness of our results by using alternative independent variable measures—that is, by converting our binary independent variable to continuous measures. In Model 1 in Table C1, the independent variable is measured as the natural logarithm of 1 plus the number of labor-related incidents reported by RepRisk. We find that the coefficient for *media labor issue exposure* is positive and statistically significant ($\beta = 0.221$; $p < 0.01$), supporting our argument. In Model 2, the independent variable is measured as the natural logarithm of 1 plus the number of labor-related incidents, where each incident is weighted by a severity score. We follow the guidance of RepRisk (2021) to use the severity score (1 = low level; 2 = medium level; 3 = high level) as the weight. We find that the coefficient for *media labor issue exposure* is still positive and statistically significant ($\beta = 0.174$; $p < 0.01$), supporting our argument. In Model 3, we use the reach score provided by RepRisk to weigh each incident when measuring the independent variable (1 = low level; 2 = medium level; 3 = high level). We find that the coefficient for *media labor issue exposure* is still positive and statistically significant ($\beta = 0.159$; $p < 0.01$), consistent with our argument. These analyses collectively confirm that our results are robust to different operationalizations of independent variable.

Table C1. Alternative Independent Variable Measures

	Model 1		Model 2		Model 3	
Variable			*AI investment (log)*			
	Number of incidents		Weighted by severity		Weighted by reach	
Media labor issue exposure (log)	0.221***	(0.043)	0.174***	(0.036)	0.159***	(0.033)
Firm size (log)	0.029*	(0.015)	0.029*	(0.015)	0.029*	(0.015)
ROA	0.009	(0.012)	0.009	(0.012)	0.009	(0.012)
Resource slack	–0.099	(0.088)	–0.100	(0.088)	–0.099	(0.088)
R&D expenditure (log)	–0.002	(0.003)	–0.002	(0.003)	–0.002	(0.003)
Blank R&D reporting	0.032	(0.042)	0.032	(0.042)	0.031	(0.042)
Labor intensity	0.458**	(0.228)	0.460**	(0.228)	0.458**	(0.228)
Employee ownership	–0.392	(0.562)	–0.387	(0.564)	–0.385	(0.566)
Collective bargaining	0.003	(0.043)	0.003	(0.043)	0.004	(0.043)
Technology governance orientation	0.012***	(0.004)	0.012***	(0.004)	0.012***	(0.004)
Managerial ability	0.099	(0.062)	0.101	(0.062)	0.100	(0.062)
Common ownership with media	–0.167**	(0.067)	–0.169**	(0.067)	–0.169**	(0.067)
Non-AI investment (log)	0.020***	(0.004)	0.020***	(0.004)	0.020***	(0.004)
Constant	0.167	(0.131)	0.168	(0.132)	0.168	(0.132)
Firm FE	Yes		Yes		Yes	
Industry-year FE	Yes		Yes		Yes	
Observations	18,759		18,759		18,759	
Within *R*-squared	0.145		0.144		0.144	

Notes: This table shows the effect of media labor issue exposure on AI investment with different independent variable measures.

The sample period ranges from 2010 to 2018. The dependent variable in all models is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). The independent variable in all models is *media labor issue exposure*. In Model 1, the independent variable is measured as the natural logarithm of 1 plus the number of labor-related incidents in a year, as reported by RepRisk. In Model 2, the independent variable is measured as the natural logarithm of 1 plus the number of labor-related incidents in a year weighted by a severity score (1 = low level; 2 = medium level; 3 = high level). In Model 3, the independent variable is measured as the natural logarithm of 1 plus the number of labor-related incidents in a year weighted by a reach score (1 = low level; 2 = medium level; 3 = high level). Results are based on OLS regressions. Standard errors are clustered at the firm level and reported in parentheses. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.

C2. Alternative Job Posting Cutoffs in Identifying AI Investment

In our baseline results, we use the relatively strict threshold of 15% to identify AI-related jobs and precisely capture firms' AI orientation (Babina et al. 2024). In Table C2, we use alternative cutoffs (10% and 5%) to construct our AI investment measure. We find that the coefficient for *media labor issue exposure* is positive and statistically significant in predicting AI investment with either a 10% cutoff ($\beta = 0.227$; $p < 0.01$) or a 5% cutoff ($\beta = 0.226$; $p < 0.01$), supporting our argument.

Table C2. Alternative AI Job Posting Identification Cutoffs

Variable	Model 1		Model 2	
	AI investment (log)			
	10% cutoff		5% cutoff	
Media labor issue exposure	0.227***	(0.054)	0.226***	(0.056)
Firm size (log)	0.041**	(0.020)	0.056**	(0.022)
ROA	0.007	(0.010)	0.020	(0.018)
Resource slack	–0.078	(0.117)	–0.101	(0.134)
R&D expenditure (log)	–0.000	(0.006)	0.000	(0.007)
Blank R&D reporting	0.011	(0.070)	0.024	(0.081)
Labor intensity	0.765**	(0.361)	0.773**	(0.379)
Employee ownership	–0.816	(0.841)	–0.427	(0.946)
Collective bargaining	0.009	(0.055)	–0.016	(0.062)
Technology governance orientation	0.014***	(0.005)	0.016***	(0.005)
Managerial ability	0.092	(0.075)	0.113	(0.078)
Common ownership with media	–0.233***	(0.085)	–0.228**	(0.097)
Non-AI investment (log)	0.046***	(0.006)	0.065***	(0.006)
Constant	0.226	(0.169)	0.172	(0.188)
Firm FE	Yes		Yes	
Industry-year FE	Yes		Yes	
Observations	18,759		18,759	
Within *R*-squared	0.190		0.217	

Notes: This table shows the effect of media labor issue exposure on AI investment with different job posting cutoffs. The sample period ranges from 2010 to 2018. The dependent variable is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year with a 10% or 5% cutoff in identifying AI-related job postings, as developed by Babina et al. (2024). The independent variable is *media labor issue exposure,* assigned a value of 1 if a firm was the target of at least one media labor issue exposure in a given year and 0 otherwise, as reported by RepRisk. Results are based on OLS regressions. Standard errors are clustered at the firm level and reported in parentheses. *** and ** denote significance levels of 1% and 5%, respectively.

C3. Alternative AI Investment Proxies and Independent Variable Time Lags

We check whether our results are driven by a specific AI investment proxy (job postings) and a specific time lag between our dependent variable and independent variable (one year). We construct two alternative proxies for AI investment, as discussed next.

Employee-resume–proxied AI investment. In addition to the AI investment measure derived from job postings, Babina et al. (2024) create a measure of AI investment based on the AI skills mentioned in firm employees' resumes, using data from approximately 535 million individual profiles obtained from Cognism. Cognism gathers resumes from a variety of sources, including third-party resume aggregators, recruiting agencies, and user contributions. Its extensive resume database encompassed 64% of the entire U.S. workforce as of 2018, making it an appropriate choice for generating a representative sample within the context of our study. Specifically, based on the "skill-level AI-relatedness measure" ω_s^{AI} developed from the Burning Glass Technologies data, Babina et al. (2024) aggregate the number of employee resumes that directly involve AI-related skills to construct the employee-resume–based AI investment measure. Following Babina et al. (2024), we calculate our second dependent variable measure as the change in the number of AI-skilled resumes recorded by Cognism in a year. The sample period for this dependent variable measure is set to 2009–2018 to reflect the availability of data from Cognism (before 2018) and RepRisk (whose data are used to construct our independent variable and are available only after 2007; with a two-year lag used in our regression).

Compared to the job-posting–proxied measure, the employee-resume–proxied AI investment measure offers two primary advantages. First, the number of AI-skilled employee resumes can serve as an indicator of a firm's actual AI-related employees; this measure reflects "AI stock" rather than solely "AI demand" (Babina et al. 2024). Consequently, the change in the number of AI-skilled resumes is a net measure of firms' AI human capital, comprehensively encompassing both AI investment and AI divestment (in cases of reductions in "AI stock"). Second, because existing employees can develop their AI skills over time, changes in the number of AI-skilled resumes can capture firms' internal AI investment activities through, for example, AI-related employee training programs. In comparison, job postings capture firms' orientations toward external investment, though firms may not necessarily resort to external sources when investing in AI.

It is also important to acknowledge the limitations of the employee-resume–proxied measure. First, resumes contain self-disclosed information, allowing for possible misrepresentations of skills by employees. In contrast, job postings directly reflect firms' AI orientation and commitment (instead of employees' self-claims) and may be associated with less self-reporting bias. Second, resume data can potentially overestimate the number of high-skilled workers (and underestimate the number of low-skilled workers), so it may not accurately represent firms' employee compositions (Fedyk and Hodson 2023). In comparison, job-posting data appear to be relatively stable representations of firms' occupational compositions (Hershbein and Kahn 2018).

Patent-proxied AI investment. Our third measure is AI investment proxied by the number of AI-related patent applications made by firms. Specifically, we obtain data on AI-related patent applications from the Artificial Intelligence Patent Dataset (AIPD) provided by the U.S. Patent and Trademark Office (USPTO) (Giczy et al. 2022).[1] Based on the abstracts, claims, and citations of 13,244,037 patent applications reported by the USPTO Pre-Grant Publication (PGPubs) dataset from 1976 to 2020, AIPD identifies patents associated with AI technology by using machine-learning–based textual analyses. Specifically, AIPD decomposes AI technologies into eight nonexclusive components:[2] machine learning, natural language processing, computer vision, speech, knowledge processing, AI hardware, evolutionary computation, and planning/control. It then uses eight binary classification models to separately estimate the extent to which each patent includes one of these eight components.[3] AIPD also assesses its model predictions against evaluations made by several experienced USPTO patent experts, who have confirmed the validity of AIPD's machine learning algorithm. The final output includes eight scores between 0 and 1 indicating the probability that a patent application includes one of the eight AI components. Reflecting on

---

[1] These data are officially provided by USPTO. For details, see https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset (extracted March 19, 2023).

[2] Because the eight components are identified using different classification methods, they are nonexclusive and overlaps are possible. That is, a specific patent application can include more than one of these components. For example, Giczy et al. (2022) note that "an invention in natural language processing may rely on an underlying machine learning method."

[3] Giczy et al. (2022) use algorithms based on "word2vec" embedding (for abstracts and claims) and "one-hot" encoding (for citations) to translate textual data into machine-readable format. These authors adopt a four-step method to implement the machine learning methodology. The first step is to construct a seed set (an anti-seed set) that provides positive (negative) examples of AI. The second step is to implement logistic regressions that train a classification model. Based on the classification model, the third step classifies the universe of the patent documents in PGPubs. The final step is to manually evaluate a random sample of model predictions against evaluations made by USPTO patent experts.

the balance between the precision and recall of their model predictions, Giczy et al. (2022) conclude that a probability score of 0.5 is a reasonable threshold to determine whether a patent application focuses on AI technology.[4]

We therefore follow Giczy et al. (2022) and define a patent application as AI-related if one of its eight prediction scores is greater than 0.5. We then link each AI-related patent application to the corresponding firm applicant (if the application was filed by a firm) based on the firm–patent linkage data (Kogan et al. 2017).[5] For each firm, we aggregate the number of its AI-related patent applications that were ultimately granted in the year of application, and then take a natural logarithm to measure *AI investment*. We aggregate firms' granted AI-related patents in the year of application instead of the year of grant because we aim to capture firms' AI innovation efforts (i.e., applications) rather than their final outcomes (i.e., granted patents) — an approach that aligns with the strategies used in prior studies (e.g., Ravichandran et al. 2017, Saldanha et al. 2020). Our sample ends in 2016 for the patent data, instead of 2018 as with the prior two dependent variable measures, to avoid measurement errors arising from the possibility that patent applications may have not been granted as of 2024, the time of this paper.[6]

Compared to the first two measures (i.e., job-posting–proxied and employee-resume–proxied AI investment), granted AI-related patent applications capture firms' AI investment output in an alternative way. Thus, they serve as an ideal complement to the aforementioned measures of AI investment, which largely capture firms' AI input. However, given that firms may not necessarily engage in AI-related patenting activities in response to media exposure of their labor issues, we use the patent-proxied AI investment measure only in a complementary analysis.

We report our results with the respective dependent variable measures in Table C3. In addition, we

---

[4] "Precision" refers to the ratio of the number of true positives to the number of predicted positives; "recall" is defined as the ratio of the number of true positives to the number of actual positives.

[5] The linkage data are provided by Kogan et al. (2017). For details, see https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data (extracted March 19, 2023).

[6] According to our calculations based on the USPTO PatentsView database, after filing patent applications with the USPTO, firms wait an average of 2.8 years before a patent is granted. Notably, 94% of patent applications are granted within 5 years, 97% within 6 years, and 98% within 7 years. As of 2024, we can precisely capture only patent applications that were ultimately granted for those filed before 2017. We thus use the year 2016 as the end year for our sample with the patent-proxied dependent variable.

include multiple time lags for our independent variable to check whether our results are driven by a specific time lag. Model 1 uses the job-posting–proxied AI investment measure as the dependent variable. We find that while contemporaneous *media labor issue exposure* does not predict *AI investment* ($\beta = -0.016$; *n.s.*), *media labor issue exposure* taking place in the last year ($\beta = 0.142$; $p < 0.01$) and two years earlier ($\beta = 0.162$; $p < 0.01$) does positively predict *AI investment*. In terms of economic magnitude, when *media labor issue exposure* increases from 0 to 1, the level of AI investment increases by 15.3% (= exp [0.142] – 1) in the next year and by 17.6% (= exp [0.162] – 1) in the second year following the media exposure. These results suggest that when faced with media exposure of their labor issues, firms lean toward making AI investments within the next two years.

Model 2 adopts the employee-resume–proxied AI investment measure as the dependent variable. We find that the contemporaneous independent variable ($\beta = -0.150$; *n.s.*) and the one-year lagged independent variable ($\beta = 0.396$; *n.s.*) do not predict *AI investment*, but *media labor issue exposure* taking place in year $t - 2$ does positively predict *AI investment* in year t ($\beta = 1.011$; $p < 0.01$). In terms of economic magnitude, when *media labor issue exposure* in year $t - 2$ increases from 0 to 1, the change in the number of AI-skilled employee resumes increases from 0.21 to 1.22 in year t. These results further support our findings in Model 1. As it takes one year for the AI investment to be manifested in "AI stock," which can then be captured by existing AI-related employee resumes, it is reasonable to assume that only the two-year lagged independent variable can significantly predict AI investment. Furthermore, the employee-resume–proxied measure can capture firms' internal AI development (e.g., AI-related employment training programs) in addition to the external AI investment proxied by job postings, which may take a longer time to implement than job postings.

Model 3 uses the patent-proxied AI investment measure as the dependent variable. We find that the contemporaneous independent variable ($\beta = -0.015$; *n.s.*) and the one-year lagged independent variable ($\beta = 0.030$; *n.s.*) do not predict *AI investment*, whereas *media labor issue exposure* taking place in year $t - 2$ does positively predict *AI investment* in year t ($\beta = 0.038$; $p < 0.05$). In terms of economic magnitude, when *media labor issue exposure* in year $t - 2$ increases from 0 to 1, the level of AI investment increases by 3.9% (= exp [0.038] – 1) in year *t*. These results provide additional support for our findings in Model 1. It is not surprising that only the independent variable with a two-year lag can significantly predict patent-based AI investment, as it takes time for AI investment to translate into AI innovation output. This finding is also consistent with prior research on innovation (e.g., Fang et al. 2014, Aghion et al. 2023), which indicates that innovation as an organizational reaction typically requires at least two years for implementation. Furthermore, the economic magnitude is smaller than that associated with the job-posting–proxied AI investment measure because the patent-proxied measure captures only a limited subset of AI investment, and not all firms regard AI innovation as a feasible or meaningful choice in response to media exposure of their labor issues. Overall, we find consistent support for our core argument that media exposure of labor issues can lead firms to engage in AI investment.

Table C3. Alternative AI Investment Proxies and Independent Variable Time Lags

Variable	Model 1		Model 2		Model 3	
			AI investment			
	Job-posting–proxied measure		Employee-resume–proxied measure		Patent-proxied measure	
Media labor issue exposure (t)	–0.016	(0.022)	–0.150	(0.109)	–0.015	(0.011)
Media labor issue exposure (t – 1)	0.142***	(0.046)	0.396	(0.286)	0.030	(0.021)
Media labor issue exposure (t – 2)	0.162***	(0.044)	1.011***	(0.351)	0.038**	(0.019)
Firm size (log)	0.030*	(0.016)	0.187*	(0.102)	0.004	(0.004)
ROA	0.008	(0.012)	–0.023	(0.015)	–0.000	(0.000)
Resource slack	–0.096	(0.088)	0.222	(0.254)	0.004	(0.026)
R&D expenditure (log)	–0.003	(0.003)	–0.022*	(0.011)	0.000	(0.001)
Blank R&D reporting	0.035	(0.042)	0.248***	(0.096)	–0.000	(0.016)
Labor intensity	0.453**	(0.229)	0.120	(0.577)	–0.043	(0.066)
Employee ownership	–0.355	(0.566)	1.086	(1.055)	0.135	(0.255)
Collective bargaining	0.001	(0.043)	–0.028	(0.064)	0.009	(0.018)
Technology governance orientation	0.012***	(0.004)	0.003	(0.011)	0.006**	(0.002)
Managerial ability	0.102	(0.064)	0.710*	(0.366)	0.007	(0.023)
Common ownership with media	–0.171**	(0.068)	–0.398	(0.258)	–0.016	(0.033)
Non-AI investment (log)	0.020***	(0.004)	0.001**	(0.001)	0.205***	(0.015)
Constant	0.165	(0.134)	–0.774	(0.615)	0.075***	(0.028)
Firm FE	Yes		Yes		Yes	
Industry-year FE	Yes		Yes		Yes	
Observations	18,759		29,511		31,477	
Within *R*-squared	0.141		0.023		0.091	

Notes: This table shows the effect of media labor issue exposure on AI investment. The sample period ranges from 2010 to 2018 in Model 1, from 2009 to 2018 in Model 2, and from 2009 to 2016 in Model 3. The dependent variable is *AI investment*. In Model 1, the dependent variable is measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). In Model 2, the dependent variable is measured as the change in the number of AI-skilled resumes in a year, as developed by Babina et al. (2024). In Model 3, the dependent variable is measured as the natural logarithm of 1 plus the number of AI-related patent applications, as developed by Giczy et al. (2022). The independent variable is *media labor issue exposure*, assigned a value of 1 if a firm was the target of at least one media labor issue exposure in a given year and 0 otherwise, as reported by RepRisk. Results are based on OLS regressions. Standard errors are clustered at the firm level and reported in parentheses. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.

APPENDIX D. MITIGATING ENDOGENEITY CONCERNS

D1. Heteroskedasticity-Based Instrumental Variable Regression

To address potential omitted variable bias (e.g., managers' personal preference in the selection between AI and labor, which is largely unobservable, may simultaneously drive media labor issue exposure and AI investment), we resort to the heteroskedasticity-based instrumental variable estimation process (Lewbel 2012). This econometric technique uses a two-stage structural system of equations that is designed to address endogeneity when traditional instruments are unavailable. It first predicts the endogenous independent variable (*media labor issue exposure*) using a first-stage model with selected covariates as predictors. It then leverages the heteroskedastic nature of the residuals of the first-stage model to transform the heteroskedastic residuals into appropriate instruments that are used in the second-stage model (for details, see: Lewbel 2012, 2018, Baum and Lewbel 2019). Due to its applicability when external instruments are unavailable, especially regarding its fitness for binary endogenous variables (Lewbel 2018), this method has been widely used in prior studies (e.g., Hong et al. 2018, Yang et al. 2020, Wang et al. 2023). The regression results from this analysis are reported in Table D1. The coefficient for *media labor issue exposure* is positive and statistically significant ($\beta = 0.289$; $p < 0.1$), consistent with our main results. Interpreting the diagnostics, we find that the Cragg–Donald Wald F-statistic is 36.89, supporting the assumption that our instruments can be considered relevant. The Hansen J-statistic is 12.69 ($p = 0.242$), suggesting that our instruments can be considered exogenous.

Table D1. Heteroskedasticity-Based Instrumental Variable Regression

	Model 1	
Variable	*AI investment (log)*	
Media labor issue exposure	0.289*	(0.175)
Firm size (log)	0.030*	(0.016)
ROA	0.009	(0.012)
Resource slack	–0.095	(0.088)
R&D expenditure (log)	–0.002	(0.003)
Blank R&D reporting	0.034	(0.042)
Labor intensity	0.453**	(0.229)
Employee ownership	–0.355	(0.565)
Collective bargaining	0.005	(0.043)
Technology governance orientation	0.012***	(0.004)
Managerial ability	0.103	(0.064)
Common ownership with media	–0.172**	(0.068)
Non-AI investment (log)	0.020***	(0.004)
Firm FE	Yes	
Industry-year FE	Yes	

Observations	18,759

Notes: This table shows the effect of media labor issue exposure on AI investment based on heteroskedasticity-based instrumental variable estimation; the instrumental variable is generated from a function of the model's existing data (for details, please refer to Lewbel [2012]). The model specifies *firm size*, *ROA*, *resource slack*, *R&D expenditure*, *blank R&D reporting*, *labor intensity*, *employee ownership*, *collective bargaining*, *technology governance orientation*, *managerial ability*, and *common ownership with media* as exogenous to generate instruments. The sample period ranges from 2010 to 2018. The dependent variable is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). The independent variable is *media labor issue exposure*, assigned a value of 1 if a firm was the target of at least one media labor issue exposure in a year and 0 otherwise, as reported by RepRisk. Results are based on OLS regressions. Standard errors are clustered at the firm level and reported in parentheses. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.

D2. Entropy Balancing Method

To mitigate potential non-random treatment issues (e.g., firms' relationships with the media may influence whether they are subjected to media exposure of their labor issues, resulting in a non-random treatment), we apply an entropy balancing approach (Hainmueller 2012). Entropy balancing directly addresses the covariate imbalance between treated and control firms by reweighting observations in the sample of control firms such that the distributional moments of the matching variables for the reweighted sample are indistinguishable from the moments of the distributions of these variables for the sample of treated firms. Consequently, entropy balancing can be useful to create balanced samples in observational studies with a binary treatment, thereby alleviating non-random treatment bias.

Compared with other matching approaches (e.g., propensity score matching [PSM] and coarsened exact matching [CEM]), entropy balancing has several advantages. First, entropy balancing solves a constrained optimization problem to identify a set of continuous weights, as opposed to discrete weights used in nearest-neighbor matching techniques, in the balancing process. As a result, entropy balancing allows unit weights to vary more smoothly. By reducing covariate imbalance, entropy balancing can also reduce the degree of model dependence, thereby reducing inefficiency and bias (Hainmueller 2012). Additionally, entropy balancing does not discard observations in the matching process; instead, it reweights all control observations to reach the same distributional moments as all treated observations. In comparison, PSM and CEM often involve dropping treated and control observations that cannot be matched to reach covariate balance. Thus, entropy balancing demonstrates a significant advantage when the number of treated observations is relatively small, as in the context in our study (Hainmueller 2012).

In implementing the entropy balancing approach, we reweight our control firms (*media labor issue exposure* = 0) to reach the same first moments (means) as treated firms (*media labor issue exposure* = 1) in terms of all covariates. We then use the weights generated from the entropy balancing approach to perform weighted regressions. Table D2A presents the mean values of the control variables in our study, partitioned by *media labor issue exposure*. As shown in the first two columns, compared with firms without media labor issue exposure (i.e., control firms; *media labor issue exposure* = 0), firms faced with such exposure (i.e., treated firms; *media labor issue exposure* = 1) have a relatively larger firm size, higher ROA, greater engagement in collective bargaining, and so on.

Table D2A. Mean Values of Covariates Before and After Weighting

	Mean values before weighting		Mean values after weighting	
Variable	Treated group	Control group	Treated group	Control group
Firm size	10.59	7.60	10.59	10.59
ROA	0.10	0.01	0.10	0.10
Resource slack	–0.04	–0.02	–0.04	–0.04
R&D expenditure	14.49	14.79	14.49	14.49
Blank R&D reporting	0.40	0.41	0.40	0.40
Labor intensity	0.00	0.01	0.00	0.00
Employee ownership	0.01	0.00	0.01	0.01
Collective bargaining	0.33	0.11	0.33	0.33
Technology governance orientation	1.57	1.12	1.57	1.57
Managerial ability	0.07	–0.01	0.07	0.07
Common ownership with media	0.25	0.20	0.25	0.25
Non-AI investment	7.64	4.84	7.64	7.64

Notes: This table shows the mean values of the control variables before and after weighting using the entropy balancing method.

Upon reweighting our control firms (*media labor issue exposure* = 0) to reach the same first moments (means) as for the treated firms (*media labor issue exposure* = 1), the mean values of all variables in the treated group are exactly equal to the mean values of these variables in the reweighted control group, creating a balanced sample in our analyses. We thus integrate the weights generated from the entropy balancing approach into our regression. As shown in Table D2B, *media labor issue exposure* still has significant predictive power for *AI investment* after the reweighting ($\beta = 0.121$; $p < 0.1$), further supporting our argument and confirming our results may not be subject to non-random treatment issues.

Table D2B. Results with Weights Generated from Entropy Balancing

	Model 1	
Variable	*AI investment (log)*	
Media labor issue exposure	0.121*	(0.069)
Firm size (log)	0.092	(0.099)
ROA	0.348	(0.427)
Resource slack	–2.068	(2.035)

R&D expenditure (log)	–0.039*	(0.020)
Blank R&D reporting	0.362	(0.238)
Labor intensity	5.719	(14.130)
Employee ownership	0.112	(2.438)
Collective bargaining	–0.078	(0.209)
Technology governance orientation	0.012	(0.010)
Managerial ability	0.018	(0.208)
Common ownership with media	–0.529*	(0.314)
Non-AI investment (log)	0.084***	(0.026)
Constant	1.144	(1.038)
Firm FE	Yes	
Industry-year FE	Yes	
Observations	18,759	
Within *R*-squared	0.495	

Notes: This table shows the effect of media labor issue exposure on AI investment. All observations are reweighted using the entropy balancing method. The sample period ranges from 2010 to 2018. The dependent variable is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). The independent variable is *media labor issue exposure*, assigned a value of 1 if a firm was the target of at least one media labor issue exposure in a given year and 0 otherwise, as reported by RepRisk. Results are based on OLS regressions. Standard errors are clustered at the firm level and reported in parentheses. *** and * denote significance levels of 1% and 10%, respectively.

D3. Heckman Selection Model

To further strengthen our empirical identification of non-random treatment issues, we apply a Heckman selection model following Lennox et al. (2012). This approach addresses non-random treatment bias by adopting a two-stage method: (1) All covariates and an external exclusion restriction are used to predict the likelihood of being treated, and then (2) an inverse Mills ratio is generated as a treatment correction term and included in the second-stage regression. Specifically, we first perform a first-stage probit regression to estimate whether a firm has media labor issue exposure in a year. We include all controls and industry-year fixed effects in the first-stage regression. We also include an exclusion restriction that can predict *media labor issue exposure* but cannot directly influence *AI investment*: *media attention*, measured as the number of media outlets covering the focal firm. This variable is likely to positively predict *media labor issue exposure*, as firms that receive more attention from the media are more likely to have their labor issues exposed by the media. Yet, this variable is unlikely to directly influence our outcome variable.

Model 1 in Table D3 displays the results from the first-stage probit regression. In this regression, the coefficient for *media attention* is positive ($\beta = 0.039$; $p < 0.05$), suggesting this variable is a significant predictor of *media labor issue exposure*. However, it does not have significant predictive power if included in the second-stage regression ($\beta = 0.001$; $p = 0.839$). We then calculate the *inverse Mills ratio* from the first-stage probit regression and include it in the second-stage OLS regression. As shown in Model 2 in Table D3, we continue to find support for our arguments ($\beta = 0.089$; $p < 0.1$). This result suggests that our results are unlikely to be subject to non-random treatment issues.

Table D3. Results from Heckman Selection Model

	Model 1		Model 2	
	First-stage probit		Second-stage OLS	
Variable	*Media labor issue exposure*		*AI investment (log)*	
Media labor issue exposure			0.089*	(0.048)
Firm size (log)	0.666***	(0.047)	–0.007	(0.015)
ROA	–0.344***	(0.046)	0.017	(0.012)
Resource slack	3.965***	(1.197)	–0.181**	(0.085)
R&D expenditure (log)	0.004	(0.014)	–0.000	(0.003)
Blank R&D reporting	0.109	(0.126)	0.004	(0.039)
Labor intensity	–55.724***	(14.591)	0.610***	(0.223)
Employee ownership	2.733**	(1.368)	–0.654	(0.560)
Collective bargaining	0.130	(0.105)	–0.047	(0.040)
Technology governance orientation	0.003	(0.014)	0.010**	(0.004)
Managerial ability	1.189***	(0.249)	–0.134**	(0.058)
Common ownership with media	–0.025	(0.290)	–0.116*	(0.065)
Non-AI investment (log)	0.036	(0.024)	0.017***	(0.004)
Media attention	0.039**	(0.018)		
Inverse Mills ratio			–3.333***	(0.342)
Constant	–6.041***	(0.648)	2.988***	(0.309)
Firm FE	No		Yes	
Industry-year FE	Yes		Yes	
Observations	18,759		18,759	
Pseudo/within *R*-squared	0.474		0.166	

Notes: This table shows the effect of media labor issue exposure on AI investment with Heckman selection model. The sample period ranges from 2010 to 2018. Model 1 shows the results from the first-stage probit regression estimating the likelihood of media labor issue exposure. The dependent variable is *media labor issue exposure*, assigned a value of 1 if a firm was the target of at least one media labor issue exposure in a year and 0 otherwise, as reported by RepRisk. The exclusion restriction included in Model 1 is *media attention*, measured as the number of media outlets covering the focal firm in a year, as reported by RavenPack. Model 2 shows the results from the second-stage OLS regression. The dependent variable is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). The independent variable is *media labor issue exposure*, which is the dependent variable in the first stage. Standard errors are clustered at the firm level and reported in parentheses. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.

D4. Staggered Difference-in-Differences (DiD) Model

We conduct a DiD analysis, which builds on a setting of staggered firm-year treatments. We define a firm-year as treated if at least one labor-related negative incident takes place as recorded by RepRisk. Accordingly, all other observations are regarded as untreated. In Model 1 of Table D4, we apply the DiD approach developed by Callaway and Sant'Anna (2021) to estimate the "group-time" average treatment effect. As shown, the average treatment effect is 0.276 ($p < 0.01$), suggesting that firms have experienced labor issue exposure are more inclined to engage in AI investment compared to those have not experienced such exposure.

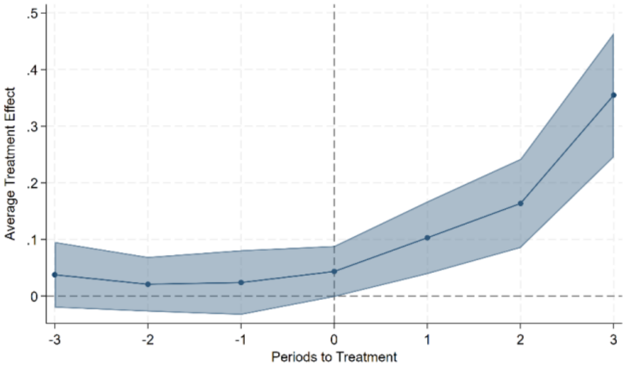
Next, we decompose the average treatment effect into dummy variables indicating the specific year around the year of treatment, and report the results of the three years pre- and post-treatment in Model 2 of Table D4. We find that the pre-treatment effects are statistically insignificant, which implies that treated and untreated firms experienced a parallel trend during the three years before the treatment. We also find that the post-treatment effects are positive and statistically significant, indicating that firms that experienced labor issue exposure continue to engage in AI investment in the following two years. This is aligned with our results in Table C3 examining different time lags. We also plot the treatment effect around the year of treatment in Figure D1, which intuitively confirms our arguments. Overall, our results are robust to DiD model specifications.

Table D4. Results from DiD Regressions

	Model 1		Model 2	
Variables	*AI investment (log)*			
Average treatment effect	0.276***	(0.041)		
Pre-treatment (t – 3)			0.038	(0.030)
Pre-treatment (t – 2)			0.021	(0.025)
Pre-treatment (t – 1)			0.024	(0.029)
Post-treatment (t + 0)			0.043*	(0.023)
Post-treatment (t + 1)			0.103***	(0.033)
Post-treatment (t + 2)			0.164***	(0.040)
Post-treatment (t + 3)			0.355***	(0.056)
Firm FE	Yes		Yes	
Year FE	Yes		Yes	
Observations	18,759		18,759	

Notes: This table shows the effect of media labor issue exposure on AI investment with DiD regressions (Callaway and Sant'Anna, 2021). The sample period ranges from 2010 to 2018. The dependent variable is *AI investment*, measured as the natural logarithm of 1 plus the number of AI-related job postings in a year, as developed by Babina et al. (2024). Model 1 reports the average treatment effect from DiD analysis based on the approach developed by Callaway and Sant'Anna (2021), where treatment is assigned a value of 1 if a firm-year has experienced at least one labor-related negative incident reported by RepRisk and 0 otherwise. Model 2 decomposes the average treatment effect in Model 1 into dummy variables indicating specific years around the treatment, and only the three years before and after the treatment are reported. Robust and asymptotic standard errors are reported in parentheses (Callaway and Sant'Anna, 2021). *** and * denote significant levels of 1% and 10%, respectively.

Figure D1. Average Treatment Effect with 95% Confidence Intervals



REFERENCES FOR APPENDICES

Aghion P, Bergeaud A, Van Reenen J (2023) The impact of regulation on innovation. *American Economic Review* 113(11):2894-2936.

Babina T, Fedyk A, He A, Hodson J (2024) Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics* 151:103745.

Bauer K, von Zahn M, Hinz O (2023) Expl (AI) ned: The impact of explainable artificial intelligence on users' information processing. *Information Systems Research* 34(4):1582-1602.

Baum CF, Lewbel A (2019) Advice on using heteroskedasticity-based identification. *The Stata Journal* 19(4):757-767.

Berente N, Gu B, Recker J, Santhanam R (2021) Managing artificial intelligence. *MIS Quarterly* 45(3):1433-1450.

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

Fang VW, Tian X, Tice S (2014) Does stock liquidity enhance or impede firm innovation? *Journal of Finance* 69(5):2085-2125.

Fedyk A, Hodson J (2023) Trading on talent: Human capital and firm performance. *Review of Finance* 27(5):1659-1698.

Fügener A, Grahl J, Gupta A, Ketter W (2022) Cognitive challenges in Human–Artificial Intelligence Collaboration: Investigating the path toward productive delegation. *Information Systems Research* 33(2):678-696.

Ge R, Zheng Z, Tian X, Liao L (2021) Human–robot interaction: When investors adjust the usage of robo-advisors in peer-to-peer lending. *Information Systems Research* 32(3):774-785.

Giczy AV, Pairolero NA, Toole AA (2022) Identifying artificial intelligence (AI) invention: A novel AI patent dataset. *Journal of Technology Transfer* 47(2):476-505.

Gnewuch U, Morana S, Hinz O, Kellner R, Maedche A (2024) More than a bot? The impact of disclosing human involvement on customer interactions with hybrid service agents. *Information Systems Research* 35(3):936-955.

Hainmueller J (2012) Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20(1):25-46.

Han E, Yin D, Zhang H (2023) Bots with feelings: Should AI agents express positive emotion in customer service? *Information Systems Research* 34(3):1296-1311.

Hershbein B, Kahn LB (2018) Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review* 108(7):1737-1772.

Hong Y, Hu Y, Burtch G (2018) Embeddedness, pro-sociality, and social influence: Evidence from online crowdfunding. *MIS Quarterly* 42(4):1211-1224.

Jain H, Padmanabhan B, Pavlou PA, Raghu T (2021) Editorial for the special section on humans, algorithms, and augmented intelligence: The future of work, organizations, and society.

Information Systems Research 32(3):675-687.

Jussupow E, Spohrer K, Heinzl A, Gawlitza J (2021) Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence. *Information Systems Research* 32(3):713-735.

Kogan L, Papanikolaou D, Seru A, Stoffman N (2017) Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics* 132(2):665-712.

Lennox CS, Francis JR, Wang Z (2012) Selection models in accounting research. *The Accounting Review* 87(2):589-616.

Lewbel A (2012) Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics* 30(1):67-80.

--- (2018) Identification and estimation using heteroscedasticity without instruments: The binary endogenous regressor case. *Economics Letters* 165:10-12.

Li J, Li M, Wang X, Thatcher JB (2021) Strategic Directions for AI: The Role of CIOs and Boards of Directors. *MIS Quarterly* 45(3):1603-1644.

Lou B, Wu L (2021) AI on Drugs: Can Artificial Intelligence Accelerate Drug Development? Evidence from a Large-Scale Examination of Bio-Pharma Firms. *MIS Quarterly* 45(3):1451-1482.

Lysyakov M, Viswanathan S (2023) Threatened by AI: Analyzing Users' Responses to the Introduction of AI in a Crowd-sourcing Platform. *Information Systems Research* 34(3):1191-1210.

Rai A, Constantinides P, Sarker S (2019) Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids. *MIS Quarterly* 43(1):iii-ix.

Ravichandran T, Han S, Mithas S (2017) Mitigating diminishing returns to R&D: The role of information technology in innovation. *Information Systems Research* 28(4):812-827.

RepRisk (2021) RepRisk methodology overview. https://wrds-www.wharton.upenn.edu/documents/1476/RepRisk_Methodology_Overview_-_September_2021_version.pdf.

Saldanha TJ, Sahaym A, Mithas S, Andrade-Rojas MG, Kathuria A, Lee H-H (2020) Turning liabilities of global operations into assets: IT-enabled social integration capacity and exploratory innovation. *Information Systems Research* 31(2):361-382.

Schanke S, Burtch G, Ray G (2021) Estimating the impact of "humanizing" customer service chatbots. *Information Systems Research* 32(3):736-751.

Sturm T, Gerlach JP, Pumplun L, Mesbah N, Peters F, Tauchert C, Nan N, Buxmann P (2021) Coordinating Human and Machine Learning for Effective Organization Learning. *MIS Quarterly* 45(3):1581-1602.

Teodorescu MH, Morse L, Awwad Y, Kane GC (2021) Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation. *MIS Quarterly* 45(3):1483-1500.

Van den Broek E, Sergeeva A, Huysman M (2021) When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly* 45(3):1557-1580.

Wang Z, Yang L, Hahn J (2023) Winner takes all? The blockbuster effect on crowdfunding platforms. *Information Systems Research* 34(3):935-960.

Yang L, Wang Z, Hahn J (2020) Scarcity strategy in crowdfunding: An empirical exploration of reward limits. *Information Systems Research* 31(4):1107-1131.

Zhang D, Peng G, Yao Y, Browning TR (2024) Is a college education still enough? The IT-labor relationship with education level, task routineness, and artificial intelligence. *Information Systems Research* 35(3):992-1010.