SKILL DEPRIORITIZATION: REORGANIZING IN THE AGE OF GENERATIVE AI

Matteo Devigli
matteo.devigili@insead.edu
INSEAD
1 Ayer Rajah Avenue
138676 Singapore

Vibha Gaba
vibha.gaba@insead.edu
INSEAD
1 Ayer Rajah Avenue
138676 Singapore

Henrich Greve
henrich.greve@insead.edu
INSEAD
1 Ayer Rajah Avenue
138676 Singapore

Erdem Dogukan Yilmaz
erdemdogukan.yilmaz@insead.edu
Bd de Constance,
77300 Fontainebleau, France

Because Generative Artificial Intelligence (GenAI) can effectively retrieve data, perform analysis, and convey information, it can be applied to enhance the capabilities of individuals involved in these tasks. This matters for the structure and processes of firms because GenAI can be used to extend the capabilities of each manager, potentially allowing similar management tasks to be executed by fewer people or increasing the range of management tasks performed. We draw on a theoretically grounded classification of the skills associated with the problems of organizing—task division, task allocation, information provision, reward distribution, and exception management—and a queuing theory model of organizing efficiency to predict the immediate reduction in skilled workers required when GenAI is available to organizations. Using a quasi-experimental design that leverages the introduction of ChatGPT as an exogenous shock, we analyze 1,774 publicly listed U.S. companies and track changes in their hiring patterns over a period of ±12 months following the shock. While we observe unchanged hiring patterns for Task Allocation, we find a significant decline in demand for Task Division and Information Provision. Exception Management and Reward Distribution exhibit a more nuanced response. These findings demonstrate that firms engage in immediate and selective skill deprioritization, suggesting that the potential for improvement varies across different organizational functions. Our paper advances our understanding of AI-driven transformation and its implications for managing organizations.

INTRODUCTION

Emerging literature on artificial intelligence (AI), and more recently on Generative AI (GenAI), has primarily focused on how AI substitutes or complements humans at the task level, particularly highlighting productivity gains and automation efficiencies (Choudhury et al 2020, Choi et al 2022, Yilmaz et al 2023, Wang et al 2024, Demirci et al 2025). A natural next question is how these task-level changes collectively reshape workforce composition at an organizational level. For instance, Moderna employees without coding backgrounds now leverage GenAI tools to create software code[1], and IBM's CEO announced in 2023 a hiring freeze for certain back-office roles, forecasting that up to 30% of such tasks could soon be automated by GenAI[2]. These examples highlight the importance of examining how organizations systematically adjust their workforce composition in response to GenAI capabilities. This paper addresses this critical yet underexplored mechanism of organizational adaptation to GenAI: skill deprioritization, which is the systematic reduction in hiring human skills effectively addressed by GenAI. Specifically, we investigate how GenAI reshapes organizational skill demands across key organizing functions and identify emerging patterns of skill deprioritization as organizations adapt to these new capabilities.

To understand these patterns theoretically, we revisit foundational organizing problems discussed by scholars since March and Simon (1958): division of labor and integration of effort. Division of labor refers to breaking organizational goals into discrete tasks and assigning resources, while integration of effort encompasses coordination, information sharing, motivation, and conflict resolution (Puranam et al 2014, Puranam 2018, Raveendran et al 2020, Raveendran et al 2022, Tonellato et al 2024). These organizing functions directly influence workforce hiring decisions and provide a solid theoretical lens for analyzing how firms systematically respond to GenAI technologies.

Central to understanding whether such effects are present is the need to elucidate AI's impact on the nature of work, particularly with respect to GenAI—a subset of AI based on large-language models (LLMs)

[1] https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Accenture-Work-Can-Become-Era-Generative-AI.pdf
[2] https://www.reuters.com/technology/ibm-pause-hiring-plans-replace-7800-jobs-with-ai-bloomberg-news-2023-05-01

such as OpenAI's GPT-4, Google's Gemini, and Meta's LlaMa. These models are trained on massive datasets and use billions of parameters to predict words in response to user prompts. Unlike previous technologies, they do not require the explicit codification of knowledge as they instead implicitly learn from a variety of inputs, including textual data. As just a casual inspection of news about GenAI technologies will reveal, they can do a variety of tasks in the blink of an eye: solve mathematical and optimizing problems correctly, perform computer coding tasks with accuracy comparable to a human coder, search the internet for answers to complex questions and produce plain-text answers that are easy to comprehend, and rewrite text to communicate with different sentiment and detail. It would be naïve to presume that organizations would respond to such novel computer-assisted capabilities by having the same individuals perform the same tasks in the same way.

We expect that LLMs will have a profound impact on how organizations address the challenges of division of labor and integration of effort through the skill deprioritization that becomes possible when GenAI capabilities make human problem-solvers unnecessary for certain tasks. While skill deprioritization responds to specific opportunities through reimagining certain organizing functions and the related job categories, the total effect on an organization or the job market may be larger than the sum of discrete managerial decisions. This paper aims to quantify and characterize this overall effect on skill demand.

There are two reasons why empirically documenting this phenomenon is crucial. First, it provides insights into how GenAI currently impacts the nature of work and labor market dynamics, beyond productivity gains on specific tasks, establishing a foundation for understanding future developments as these technologies evolve (Dixon et al 2021, Yilmaz et al 2023, Hui et al 2024, Law and Shen 2024). Second, it can guide firms and policymakers in addressing the shifts in skill demand caused by GenAI.

We analyze firm-level labor market data on skills associated with the problems of organizing (Puranam 2018): task division, task allocation, information provision, reward distribution, and exception management. This provides a theoretically grounded and systematic lens for examining GenAI's impact on essential organizing functions. To make our theoretical predictions on these effects, we view organizations as engaging in problem-solving and apply a queuing model consisting of a problem arrival process, a set of problem solvers with associated problem-solving processes, and probability distributions for the behavior of each of these (see Adan and Resing 2015). This approach has previously been used to model decision-making in organizations by Glynn, Greve, and Rao (2020), and we extend it to model how the added capabilities offered by GenAI can enhance task-skill matching in organizing functions.

To empirically test how LLMs affect problems of organizing, we adopt a quasi-experimental design that leverages the public release of ChatGPT as an exogenous availability[3] shock to traditional AI and GenAI. Our sample comprises 1,774 publicly listed U.S. companies, with hiring data tracked over a ±12-month window surrounding ChatGPT's launch. Using a difference-in-differences approach, we compare changes in job postings for skills related to organizing functions in firms with high and low GenAI exposure. To define GenAI exposure, we use the AI Exposure measure originally proposed by Felten et al (2021) and subsequently updated to account for occupations exposed to GenAI (Felten et al 2023). We find that following the launch of ChatGPT, firms with high GenAI Exposure experienced an immediate statistically significant decline in demand for skills related to Task Division and Information Provision. We find significant reductions across all Information Provision subcategories—coordination, synchronous communication, and asynchronous communication. We also observe notable declines in skills associated with Reward Distribution monitoring and the operational side of Exception Management. In contrast, job postings for skills related to Task Allocation (staffing and mapping), Reward Distribution incentives, and Exception Management conflict resolution did not exhibit any statistically significant changes. We interpret the implications of these findings for theory and for the practical use of GenAI.

Our study makes several contributions. First, we empirically demonstrate that a theoretical model of organizations solving problems optimally provides a useful foundation for predicting the reduction of human skills following the availability of GenAI. While we do not intend to claim that the firms are acting fully optimally, our results suggest that the optimal queuing model is a good first approximation because

[3] ChatGPT represents an availability shock to GenAI and AI capabilities by reducing barriers across multiple dimensions. For example, it lowers knowledge boundaries, enabling non-technical audiences to leverage complex GenAI and AI capabilities; it lowers prototyping costs by fostering rapid experimentation; and it may accelerate skill acquisition and learning curve—all the above, even from a smartphone.

firms change their hiring in ways that are directionally the same as what it predicts. Specifically, our queuing model predicts that with enhanced GenAI capabilities, organizations would reduce problem-solvers in functions where GenAI can: (1) process information more quickly, (2) reduce task interdependencies, and (3) automate standardized monitoring tasks. Our empirical findings closely match these theoretical predictions. For instance, the model anticipated reductions in information provision and task division skills, which we observed in a statistically significant decline of 21.9% in coordination skills and 23.5% in task division job postings. The model's prediction that GenAI would impact structured, rule-based organizational tasks most strongly proved remarkably accurate, suggesting that even without intentional optimization, firms naturally gravitate toward skill configurations that approximate rational resource allocation.

Second, we document the immediate effects of GenAI availability on industry-wide changes in the demand for human skills. As GenAI renders certain skills redundant or less valuable, firms reorganize resource portfolios to improve fit with AI (Lieberman et al 2017, Tandon et al 2024, Wuebker et al 2023). In doing so, we contribute to research on AI-driven organizational transformation by highlighting skill deprioritization—a mechanism that has received comparatively less attention than skill augmentation or reskilling. Our analysis of workforce configurations shows how AI-enabled changes can systematically deprioritize certain skills for individuals, with implications for organizational structures and processes. Our findings imply that certain organizing skills may become scarcer in the labor pool because they are less important to organizations that use AI systems to solve them with lower labor input than previously.

Finally, our study builds on existing models of categorizing skills into substantive clusters (e.g., Autor et al 2003, Deming and Kahn 2018), proposing a theoretically grounded taxonomy for organizing them. In identifying which jobs accomplish which organizing tasks, we develop a structured methodological framework that combines a hand-curated analysis of meanings with natural language processing scaling capabilities. Our work thus advances theory, tests theory, and offers a description of the new GenAI-enabled world of organizing, and it offers tools for new research on these topics.

THEORETICAL BACKGROUND

Foundations of Organizing Behavior

Research on organization design examines how formal and informal organizational attributes—such as structures, processes, and systems—address coordination, control, and motivation challenges (Simon 1947, March and Simon 1958, Burns and Stalker 1961, Thompson 1968, Weber, 1978, McEvily et al 2015, Joseph and Sengul 2025). This research has long emphasized that successful organizational configurations must achieve external and internal fit (Lawrence and Lorsch 1967, Mintzberg 1979, Burton and Obel 2004), and accordingly, structure, processes, and staffing need to adapt to environmental and strategic conditions (Donaldson and Joffe 2014). Recent studies further refine this notion of "fit" by focusing on misfit dynamics—how gaps between design elements and environmental demands prompt rebalancing of resource portfolios (Burton and Obel 2004, Puranam, 2018).

The organization design literature conceptualizes the problem of organizing along two dimensions: division of labor and integration of effort (March and Simon 1958, Lawrence and Lorsch 1967, Mintzberg 1979, Puranam 2018). Division of labor involves breaking down the organization's overarching goals into tasks (task division) and assigning those tasks to individuals or units (task allocation). Task division involves structuring goals into manageable components and deciding which tasks should be grouped or separated (Puranam 2018). Task allocation then matches those tasks to the most suitable individuals or units, considering skill requirements and interdependencies. Effective task allocation hinges on individual proficiencies and how tasks fit together to generate synergies that align with both internal capabilities and external demands.

Integration of effort ensures that agents coordinate and work collectively toward organizational goals. The first component is information provision, which is how relevant information flows among interdependent agents. The latter may happen through real-time or asynchronous communication channels—a distinction further elaborated in the literature (e.g., Rerup 2009, Joseph and Ocasio 2012), highlighting how organizational design can shape managerial attention and feedback mechanisms. Reward distribution entails designing incentive systems to motivate desired behaviors and maintain goal alignment, incorporating both tangible (e.g., compensation) and intangible rewards (e.g., social recognition and status). Exception management addresses unanticipated situations or conflicts, and its effective integration often depends on robust coordination mechanisms (Martin and Eisenhardt 2010, Foss et al 2013) that support knowledge flows and shared understanding within the organization.

Human-AI Coordination as an Organizing Challenge

The availability of GenAI – and, consequently, AI in general – alters the problems of organizing by enabling individual agents to perform information processing and decision making that they could not do earlier and by introducing more effective communication along with opportunities for automation that substitute for communication among human employees. Early AI technologies primarily excelled at large-scale pattern recognition and automation of routine tasks, offering clear advantages over human performance in these areas (Brynjolfsson and Mitchell 2017). However, these systems were limited in their ability to interpret context and engage in ethical reasoning, which requires human oversight for broader strategic decision-making (Lebovitz et al 2022). Recent advances in AI, particularly in LLMs, now challenge this division by exhibiting capabilities that increasingly mimic facets of human creativity and communication (Kellogg et al 2020). This evolution invites a reexamination of traditional task division and coordination between human and AI agents, suggesting that the boundaries of human labor are shifting as organizations reconfigure work processes to exploit these new capabilities (Anthony et al 2023). Such a shift may require a recalibration of internal coordination processes as organizations reallocate responsibilities from roles associated with routine supervision, conflict resolution, or performance assessment.

For instance, Dixon et al (2021) demonstrate that investing in robotics can expand managerial spans of control and flatten hierarchical structures, partly by reallocating monitoring and coordination functions to machines. While earlier AI systems managed real-time data analysis and automated structured decisions, they largely depended on human judgment for strategic direction and resolution of unforeseen or unstructured issues (Choudhury et al 2020). In contrast, the emergence of GenAI suggests that tasks once considered uniquely human may become automated, while other tasks remain human but are made more efficient through AI tools. It could lead organizations to reallocate responsibilities away from organizational roles that are solved more efficiently using AI or substituted by AI processes.

To understand how such reallocation of responsibility will change organizations, we apply a queuing model that allows for comparative analysis of organizational capabilities under technological change (Glynn et al 2020). We focus on analyzing the first-order effect of GenAI, enhancing the problem-solving capabilities of individuals by examining how this technological advancement immediately impacts organizational skill requirements. We also provide some proposals on the second-order effect of GenAI, allowing reorganization of information flow and reassignment of some problem-solving from humans to AI decision making, but we note that these second-order effects are unlikely to be observable soon after the introduction of GenAI.

The Glynn et al (2020) model views the organization as a queue that faces a Poisson distributed arrival process of problems. Problems are addressed by having problem-solving individuals and assigning to each problem an expected solution time following an exponential distribution. The individuals do not necessarily solve problems independently, as they may instead work in teams or work in hierarchies where supervisors approve solutions proposed by subordinates. They may also have functional specialization that requires problems to be sorted and correctly assigned to a qualified problem-solver, or they may have different levels of expertise requiring problems to be sorted by difficulty level and correctly assigned to a qualified problem-solver. Because these organizing approaches can all be modeled as queues with different structures, this framework allows the modeling of many organizational forms (Glynn et al 2020).

The first-order effects of GenAI on organizing functions are relatively simple to predict. Because each individual problem-solver has enhanced capabilities, giving the exponential distribution describing problem-solving duration a lower expected value, fewer problem-solvers can accomplish the same tasks. However, the decrease will be nonlinear in the increase of capabilities, as fewer problem-solvers lead to greater risk in terms of a higher standard deviation of solution times. This effect may be equal across organizing tasks and cannot be used to propose that specific organizing tasks are more vulnerable than others.

More specific effects are also likely. First, because AI-enabled individuals solve problems more quickly and can draw on GenAI's stock of knowledge to address a broader range of problems, task division and scheduling tasks are simplified. Scheduling becomes less important because faster problem-solving requires less scheduling, whether it is done individually or in teams. Traditionally, managers spend considerable time breaking down complex tasks, but GenAI may make this process smoother. First, it allows for automated task decomposition, saving managerial resources. For example, the project management software Jira uses a GenAI assistant to turn high-level project descriptions into detailed issues and sub-tasks, recommending work items and dependencies.[4] In software teams, GenAI assistants accelerate the planning phase by drafting user stories and to-do lists. Jira's AI can instantly create a set of subtasks from a plain-language goal or generate a work breakdown structure for an agile sprint.[5] Second, task decomposition also becomes simpler and faster because problem-solvers themselves are more flexible. With enhanced capabilities, they can take on broader, less specialized roles, thus deemphasizing the need for breaking down tasks. As the theory proposes and the examples illustrate, AI can reduce the resources needed for task division as an organizing task. Because task division is accompanied by behavior and outcome control to ensure that operations are smooth, the exception management created by this organizing task should also require fewer resources. This reduction in resource use occurs because AI systems enable organizations to automatically spot and solve operational flaws, while also facilitating the process of task decomposition in the first place. Stripe, for example, built GPT-4 into its internal tools for fraud detection and developer support, digesting large documentation and spotting out-of-pattern issues. Stripe's GPT-4 analyzes support tickets and code to troubleshoot integration problems, answering developer questions and highlighting likely bugs.[6]

A second specific effect is related to improving information gathering by AI-enabled individuals. To the extent that GenAI facilitates the collection and interpretation of information both externally and

[4] https://www.atlassian.com/platform/artificial-intelligence
[5] https://thejiraguy.com/2023/12/11/atlassian-intelligence-is-here/
[6] https://www.techcircle.in/2023/03/16/11-companies-using-gpt-4-in-consumer-products

internally within the organization, this allows for an effective information-pull mechanism governed by each problem-solver, which in turn lowers the need for an information-push mechanism governed by the organization. As a result, the information provision organizing task becomes easier to perform. It is also plausible that AI enables direct retrieval, integration, and analysis of information, enabling the organization to delegate some decision-making tasks to automated decision-makers instead of human ones. Such a delegation would be an additional reduction in the resources needed for information gathering. For example, in the public sector, large agencies are tapping GenAI for translation and research assistance – forms of knowledge provision crucial in government. The U.S. State Department, for instance, launched an internal GenAI chatbot in 2024 for a thousand diplomats and staff. Officials reported using it for tasks like summarizing policy documents and translating text between languages, which helps diplomats digest information quickly and communicate across language barriers.[7]

A third specific effect is related to the automatic collection of information that can be used to monitor individual effort and outcomes, and hence can be applied for control and reward purposes. In many enterprises, monitoring technology capabilities (e.g., performance tracking software, automated metrics collection) have been ahead of information-gathering and analysis technology for some time. However, GenAI allows organizations to utilize this richer data in an automated manner, thereby reducing the need for human problem-solvers to interpret the performance of other employees and assign rewards to them. To the extent that such automation of reward distribution occurs, the resources needed for this task would also be reduced. For instance, the HR analytics platform Culture Amp has experimented with ChatGPT to summarize multiple sources of employee feedback. Instead of a manager relying only on memory, an AI agent can retrieve data like kudos from Slack, customer comments, and peer feedback collected over months and use it to generate a coherent performance summary.[8]

We see less foundation for predicting a reduction in task allocation as an organizing effort, the design of incentives as a reward distribution effort, or conflict resolution as an exception management

[7] https://fedscoop.com/state-department-encouraging-workers-to-use-chatgpt
[8] https://www.shrm.org/topics-tools/news/technology/how-hr-using-generative-ai-performance-management

effort, as individual selection and motivation would presumably be similar when individuals are AI-enabled, and the potential for conflicts is also likely to remain. Nevertheless, we posit that reductions in these organizing efforts are also possible through the second-order effects of GenAI not only enabling individuals to perform a greater variety of tasks with more information and greater speed but also allowing reorganization of the firm and changes in the division of labor between human and automated information distribution and decision making. If such effects occur, they should be observed later than the first-order effects. We expect organizations first to change work processes in more straightforward ways and later engage in more complex reorganization. We will examine whether these second-order effects can be detected in our data, but the longer duration of implementing such changes could mean that they cannot be detected.

Our theory thus predicts that firms will initially reduce human resources assigned to task division and scheduling, information provision, and monitoring for reward distribution, and will later reduce human resources assigned to task allocation, designing incentives, and conflict resolution. To test whether the changes follow our expectations, we empirically investigate how GenAI influences organizational skill demand by leveraging the public release of ChatGPT as an exogenous shock, as we detail in the next section.

METHODS

To identify AI's impact on skill demand, we employ a quasi-experimental design using real-world observational data. We take advantage of ChatGPT's public release as a natural experiment - an unexpected event that allows us to examine how companies with different levels of GenAI Exposure responded to this technological advancement.

The launch of ChatGPT serves as an ideal quasi-experimental setting for three reasons. First, the introduction of ChatGPT represented a significant advance in AI capabilities that was simultaneously revealed to all market participants. This technological visibility shift transformed how managers perceived AI's capabilities - while previously AI may have seemed abstract or limited in scope, ChatGPT provided concrete demonstrations of traditional AI models and GenAI's ability to handle a wider range of tasks, including complex cognitive tasks. This widespread exposure likely led decision-makers to reassess which organizing tasks could be automated or augmented by AI, ultimately influencing the workforce planning and hiring decisions regarding skill requirements. Second, the introduction of ChatGPT is empirically relevant due to its heterogeneous effects across our sample. Following Felten et al (2023), we identified companies that are more likely to be affected by this technological advancement based on their exposure to AI-relevant tasks. Specifically, we define treated companies as those predominantly operating in task environments where AI's natural language processing and generative capabilities have the most significant impact. Third, the ChatGPT introduction is also credibly exogenous, reducing the concern for anticipatory adaptation. Managers likely had neither well-formed priors about AI's scope nor advance knowledge of the chatbot's release. Additionally, companies lacked the incentives and capacity to self-select into either treatment or control groups.

To operationalize this design, we implement a difference-in-differences (DiD) approach:

$$y_{it} = \alpha_{it} + \beta PostChatGPT_t \, x \, Treated_i + \gamma X_{it} + \delta_i + \delta_t + \varepsilon_{it}$$

where *i* indicate firms, *t* refers to year-month, δ_i and δ_t refer to the company and year-month fixed effects, and ε_{it} represents the error component. Our dependent variable y_{it} is the count of job postings requiring a specific subset of skills by a firm in a given month. The interaction term $PostChatGPT_t \, x \, Treated_i$ is our regressor of interest, identifying treated companies after the ChatGPT release by OpenAI. We do not add $PostChatGPT_t$ and $Treated_i$ first-order terms as separate regressors because these are collinear with time and firm fixed effects. We define the post-treatment period as a ±12-month window around December 2022, the month of ChatGPT's launch (specifically, November 30, 2022). We chose this ±12-month window around ChatGPT's release because the pre-period (December 2021 - November 2022) provides sufficient historical data to establish baseline hiring patterns while being recent enough to avoid contamination from earlier AI developments. Also, the post-period (December 2022 - November 2023) allows enough time for organizations to adjust their hiring practices in response to the shock while being short enough to minimize the risk of confounding events like subsequent GenAI releases or broader economic changes affecting our results.

We employ a Poisson Pseudo-Likelihood estimator with high-dimensional fixed effects, clustering standard errors at the company level.

Data and Variables

We analyze a sample of 1,774 publicly listed US companies, tracking hiring decisions over ±12 months from the exogenous shock, which results in 39,122 firm-month observations. Our dataset integrates three sources: Lightcast, Compustat, and Revelio. Lightcast maps the entire US job market daily, with detailed information at the job-posting level, while Compustat and Revelio provide company-level control variables. Section A of the Online Appendix[9] provides a detailed description of the sample construction process, including the matching procedures and the number of companies matched at each step.

Treated Companies. We identified treated companies using the AI Exposure measure, originally proposed by Felten et al (2021), updated specifically to account for occupations exposed to GenAI (Felten et al 2023). This measure has two main components: (1) the relationship between 52 human abilities and more than 800 occupations, as listed in the Occupational Information Network (O*NET) database developed by the United States Department of Labor; and (2) a crowd-sourced matrix quantifying the perceived exposure of these 52 human abilities to various AI applications (Felten et al 2021). For our research, we focus specifically on the 'language modeling' AI application and use the O*NET database updated to 2023.

We applied the following steps to classify companies into the treated and control groups. First, we replicated Felten et al (2023) AI Occupational Exposure (AIOE) as follows:

$$AIOE_k = \frac{\sum_{j=1}^{52} A_j \times L_{jk} \times I_{jk}}{\sum_{j=1}^{52} L_{jk} \times I_{jk}}$$

Where *k* refers to an occupation and *j* indexes each of the 52 occupational abilities. A_j represents the ability *j* exposure to the *'language modeling'* AI application, while L_{jk} and I_{jk} are weights for ability *j* importance and prevalence by occupation *k*, respectively. Second, using Lightcast data, we collected monthly counts of job postings categorized by four-digit NAICS industries and O*NET occupations for the period 2021-2023. We then calculated the ratio of each O*NET occupation's monthly count to the total monthly job postings within the corresponding four-digit NAICS industry. This ratio represents the relative prevalence

[9] Online Appendix (anonymized): https://osf.io/4ymcq?view_only=69cebabc1ab844268d69a8c262194b8d

of an occupation within each industry-month combination. Then, we obtained the monthly four-digit NAICS GenAI Exposure as follows:

$$GenAI\ Exposure\ (monthly)_n = \sum_{k=1}^{K} AIOE_k \times S_{kn}$$

Where n indexed a four-digit NAICS sector and k refers to an occupation. $AIOE_k$ is discounted by S_{nk} which refers to the prevalence of occupation k in sector n. Lastly, we averaged these monthly GenAI Exposure values over the entire study period to obtain a final sector-level GenAI Exposure measure. Companies in sectors whose average GenAI Exposure exceeded the overall sample mean were classified into the treated group.

Our classification strategy assigns 966 companies to the treatment group and 808 to the control group. This distribution emerged naturally from our data, resulting in a division reasonably close to an ideal balanced design. Section A of the Online Appendix provides a snapshot of the top and bottom 20 sectors in terms of GenAI Exposure in our sample. Additionally, Figure A1 shows the distribution of GenAI skills demand between treated and control groups, considering the 12 months after the ChatGPT release, which demonstrates that treated companies posted substantially more jobs requiring GenAI skills compared to control companies, providing support for the validity of our treatment and control group assignment

Problems of Organizing. Our dependent variables measure the number of job postings requiring skills related to the problems of organizing (Puranam et al 2014, Puranam 2018, 2024) by a company each month. To operationalize this, we leverage Lightcast's skill taxonomy, which includes 33,832 unique skills mapped to job postings. For instance, a product manager job in a software company might be associated with Product Management, Engineering Design Processes, and Prioritization. Each skill in the taxonomy is also accompanied by a short synopsis (~70 words) describing its function. For example, "Prioritization" is defined as: *'the ability to effectively plan and prioritize by assessing the relative importance and urgency of a task, activity, or event [...]'*. We use these descriptions to classify skills into five *a priori* categories reflecting the problems of organizing: task division, task allocation, information provision, reward distribution, and exception management. Only job postings that include any of these management skills are analyzed, so the classic examples of labor jobs (e.g., lathe operator) are not analyzed.

To systematically map the 33,832 skills onto *a priori* categories, we employed a multi-step process that combines topic modeling, manual coding, and GenAI-based categorization. First, we used topic modeling to group skills into coherent thematic clusters, from which we selected four clusters closely aligned with our categories of interest. Next, by manually reviewing the descriptions of the resulting 1,765 skills, we developed ten distinct categories mapping onto the five problems of organizing. Finally, we scaled our classification approach to the full dataset of 33,832 skills using GPT-4o.

In the first step, we conducted a Topic Modeling analysis to filter relevant skills. Specifically, we applied a Latent Dirichlet Allocation (LDA) algorithm, generating 20 alternative models (varying the number of topics from 5 to 100 in increments of 5). We evaluated these models based on Coherence Scores and qualitative inspection of top keywords and skills' descriptions per topic. The 65-topic solution provided the best discrimination of themes. Four topics were highly associated with our five categories of interest. We selected all skills with high posterior probabilities for those topics, yielding a filtered pool of 1,765 skills.

In the second step, we read the descriptions of the filtered skills, defining ten labels as summarized in Table 1. This manual labeling process identified 397 skills that matched our categories (True Positives), and 1,368 that did not (True Negatives). Each resulting label is minimal, thus containing the smallest possible set of semantic attributes of the concept to measure, and independent, as each label does not rely on any other label to explain the target meaning. These properties increase confidence that the subsequent ML-based projection will accurately capture the intended meanings. We validated the above classification by asking a Research Assistant to reproduce the hand-curated labeling over the 1,765 skills. In doing so, we achieved an average Precision, Recall, and F1-Score across all labels of 0.83, 0.81, and 0.83 suggesting a satisfactory agreement. On the True Positive subset alone, we reached a Cohen's Kappa of 0.7928. Lastly, we discussed cases of disagreement to improve label definitions and identify boundary cases.

Insert Table 1 about here

In the final step, we scale our classification to all 33,832 skills, leveraging GPT-4o, to determine whether a skill reflects one or more[10] of our organizing problem categories. Each skill is labeled by three independent GPT-4o agents and assigned one or more labels when at least two agents reach an agreement on those dimensions. In essence, the GPT-4o agents act as filters to determine the final set for manual screening. Thus, we validated the model on the manually labeled sample, where GPT-4o achieved 93% Accuracy, 86% Recall, and 89% F-1 score in identifying True Negatives. The average Precision, Recall, and F1-Score across all labels is of 0.82, 0.80, and 0.81, respectively. On the full dataset, the model identified 1,369 positive instances, which were then manually reviewed for accuracy. This process identified 657 True Positives mapping on 719 skill-label relationships.

We used these skill-problem pairings to count the number of job postings related to each of the ten organizing problem categories, forming our dependent variables. Through this step, we processed 11,154,915 jobs posts.

Section B of the Online Appendix presents a topical map derived from the 65-topic model. It also includes the prompt used to configure GPT-4o agents and a bar chart showing skill frequencies by label.

Control Variables. Our DiD includes the following controls: size, return on assets (ROA), leverage, the ratio of R&D over total assets (R&D/TA), an R&D dummy to further distinguish companies with no R&D investments, the ratio of capital expenditures over total assets (Capex/TA), and a Workforce Change indicator.

The first set of control variables used in the analysis were computed using quarterly financial data from Compustat. Size is measured as the natural logarithm of one plus total assets. Return on Assets (ROA) is defined as quarterly net income divided by total assets. Leverage is calculated as the ratio of total liabilities (sum of short-term and long-term debt) to total assets. The ratio of R&D expenditures to total

[10] Multilabel classification.

assets (R&D/TA) is computed as quarterly R&D expenses divided by total assets (missing R&D values are set to zero). To differentiate firms without R&D activities, an R&D dummy is introduced, taking the value of one for firms reporting positive R&D expenditures and zero otherwise. Finally, the ratio of capital expenditures to total assets (Capex/TA) is obtained by dividing quarterly capital expenditures by total assets.

Additionally, we control for the Workforce Change measured as the month-over-month percentage difference in headcount, which can be negative (contraction) or positive (expansion). For example, firms undergoing substantial layoffs may exhibit temporarily depressed hiring activity or shifts in skill demand due to restructuring, rather than GenAI. We used Revelio to track changes in each company's headcount by month:

$$Workforce\ Change_{it} = \frac{HC_t - HC_{t-1}}{HC_{t-1}}$$

where *i* refers to a company, *t* a month-year. HC_t refers to the total number of employees associated with a company in month t, while HC_{t-1} is the total for the previous month.

In Figure A2 in section A of the Online Appendix, we compare the evolution of Workforce Change between the treated and control groups across the ±12-month window around November 2022. The trends in workforce change between treated and control groups are similar across the study period, indicating no substantial differences in workforce reduction patterns.

Descriptives. Table 2 presents descriptive statistics for key variables across 39,122 firm-month observations. On average, firms in our sample have a size of 7.85 (measured as the log of total assets), profitability close to zero (ROA), leverage of 0.30, and a capital expenditures ratio of 2% relative to total assets. Regarding innovation activity, 44% of firms report positive R&D expenditures, with an average R&D intensity of 2%. 55% of companies in the sample are classified as highly exposed to GenAI, with an average monthly GenAI Exposure score of 0.34.

In terms of job postings, companies issue an average of 92 postings requiring Task Division skills. Demand for Information Provision (IP) roles is also substantial, with an average of 77 job postings related to synchronous communication, 61 to asynchronous communication, and 54 to coordination tasks. Exception Management (EM) also shows strong demand, particularly for operational-related skills, averaging 65 job postings per firm month. In contrast, Reward Distribution (RD) and Task Allocation (TA) categories exhibit lower levels of monthly job postings.

Table 3 reports correlations among the ten variables representing the demand for the Problems of Organizing. Correlations are generally high, ranging from 0.21 to 0.91. All coefficients are significant at the 5% level.

Insert Tables 2 and 3 about here

RESULTS

Baseline Results

Table 4 reports the baseline DiD estimates examining the impact of GenAI Exposure on hiring for skills associated with the Problems of Organizing. We do not include any control variables in these specifications to avoid the potential issue of "bad controls".[11] All estimates include firm and year-month fixed effects with standard errors clustered at the firm level. The results indicate a statistically significant decline in job postings for roles related to Task Division, with an estimated coefficient of -0.268 ($p < 0.001$). This corresponds to a reduction of 23.5%, suggesting that, following exposure to advanced GenAI, treated firms decreased their demand for positions focused on breaking down, structuring, and organizing work before execution. Similarly, the results show a substantial reduction in job postings related to Information Provision. Postings for coordination skills declined by 21.9% ($\beta = -0.246$, $p < 0.01$), synchronous communication skills by 19.0% ($\beta = -0.210$, $p < 0.05$), and asynchronous communication skills by 22.1% ($\beta = -0.250$, $p < 0.01$). For Reward Distribution, the estimates indicate a significant decrease in postings related to monitoring activities, with a reduction of 24.1% ($\beta = -0.275$, $p < 0.001$). There is no statistically

[11] A *bad control* is a variable that is affected by the treatment and also affects the outcome. Including it as a control can "soak up" part of the treatment effect, leading to biased or misleading estimates. In DiD designs, controlling for post-treatment variables (or variables that may be endogenous to treatment) can distort causal inference.

significant change in roles tied to incentive provision. Within Exception Management, operational exception management shows a significant decline of 19.5% ($\beta = -0.217$, $p < 0.01$), while conflict management shows no significant difference post-ChatGPT introduction. Lastly, neither the staffing nor mapping dimensions of Task Allocation results were affected by the GenAI availability shock.

Insert Table 4 about here

Table 5 presents the estimates from the baseline specification with the inclusion of control variables. The results remain consistent with the previous analysis. Demand for Task Division declines by 23.7% following GenAI Exposure. Similarly, job postings related to Information Provision show significant reductions, with coordination roles decreasing by 21.7%, synchronous information sharing by 19.1%, and asynchronous information sharing by 22.3%. In addition, the monitoring component of Reward Distribution experiences a decline of 23.8%, while demand for Operational Exception Management skills decreases by 19.6%.

Insert Table 5 about here

Figure 1 presents a coefficient plot with 95% confidence intervals from a dynamic DiD model. The specification is consistent with our DiD approach, incorporating the same control variables, fixed effects, and clustered standard errors. However, we include month-year dummies rather than relying on a single post-ChatGPT indicator, using November[12] 2022 as the reference period (i.e., t-1). The plotted estimates support the validity of the parallel trend assumption for most of our dependent variables in the pre-treatment period. The only exception is Task Allocation (staffing), where the pre-ChatGPT distribution (December 2021 – November 2022) reveals a small divergence between the treatment and control groups. To formally assess the assumption of parallel trends, we conduct a χ^2 test to evaluate the joint significance of the pre-treatment coefficients. We fail to reject the null hypothesis for all outcomes of interest, indicating no

[12] ChatGPT was introduced the 30th of November 2022. For this reason, we consider November as our reference period (t-1) and December as the month in which the treatment starts.

statistically significant differences in pre-treatment trends, except that Task Allocation showed a χ^2 test significant at 5%. Finally, Figure C1 in Online Appendix C reports the results of a sensitivity analysis at the monthly level following Rambachan and Roth (2023), examining how our estimates change when allowing for bounded deviations from strict parallel trends. The analyses suggest that even under high levels of deviation from the parallel trends assumption, our results would remain robust.

Insert Figure 1 about here

Continuous Industry-level GenAI Exposure

Next, we relax the assumption of a dichotomous treatment versus control group classification by leveraging the continuous measure of industry-level GenAI Exposure: GenAI Exposure (monthly). We estimate the interaction effect between PostGPT and GenAI Exposure (monthly), as reported in Table 6. Overall, the findings largely mirror those reported in Tables 4 and 5, except for the results for Information Provision (synchronous), where the effect differs.

The estimates indicate that firms with higher GenAI Exposure experience a statistically significant reduction in job postings related to the Task Division, with a coefficient estimate of -0.337 ($p < 0.01$). Similarly, there is a substantial and statistically significant decline in demand for skills associated with Information Provision (coordination) ($\beta = -0.297$, $p < 0.05$) and Information Provision (asynchronous) ($\beta = -0.324$, $p < 0.05$). In contrast, the demand for Information Provision (synchronous) skills remains statistically unchanged. For Reward Distribution, we find a significant decline in postings related to monitoring activities ($\beta = -0.349$, $p < 0.01$), while the demand for roles focused on the design of incentives shows no statistically significant variation. For Exception Management, demand for operational skills decreases significantly ($\beta = -0.305$, $p < 0.05$), whereas postings for conflict resolution roles remain unaffected. Finally, we find no significant changes in demand for Task Allocation skills, whether related to staffing or mapping.

Insert Table 6 about here

Distinguishing the effect of GPT-3.5 versus GPT4

This research rests on the premise that the introduction of ChatGPT may have helped managers form expectations regarding the scope and applicability of AI technologies. Based on this argument, we posit that managers have updated their hiring decisions in response to their own interactions with ChatGPT or those of their peers. To test the robustness of this claim, we examine the subsequent upgrade from GPT-3.5 to GPT-4 in March 2023. GPT-4 was regarded as a major improvement over GPT-3.5 with enhanced accuracy across a broad set of tasks. GPT-4 remained state-of-the-art for over a year and has been available for over two years as a legacy model. Consequently, we expect GPT-4 to exert a greater impact on managers' expectations and firm decisions on workforce restructuring.

To validate our claim, we split our treatment into GPT-3.5 (30th of November 2022) and GPT-4 (14th of March 2023), so we created a GPT3.5 dummy that takes the value of 1 between December 2022 and February 2023 and zero otherwise, while a GPT4 dummy with a value of 1 from March 2023 onwards and zero otherwise. Table 7 replicates the DiD specification of Table 5, while Table 8 replicates Table 6 leveraging the continuous GenAI Exposure (monthly) measure. The results in Table 7 reveal a marked increase in the magnitude of the coefficients for Information Provision (asynchronous), Reward Distribution (monitoring), and Exception Management (operational) during the GPT-4 period, alongside a more modest increase for Task Division. Table 8 also provides strong support for our claim: all coefficients exhibit substantial increases in magnitude, suggesting that the release of GPT-4 intensified the observed effects of GenAI Exposure on firms' hiring decisions. This increase in the effect when GPT-4 became available is consistent with a change in expectations modifying hiring decisions and is inconsistent with the gradual changes that would result from anticipation of GPT-4 capabilities or improved implementation of GPT-3 capabilities.

Insert Tables 7 and 8 about here

Robustness

We conducted several robustness checks to validate our findings. First, we performed a placebo analysis to assess whether the observed effects could be explained by random fluctuations rather than the release of ChatGPT. Specifically, we selected three pseudo-intervention dates set at 6, 9, and 12 months prior to the actual launch of ChatGPT. For each pseudo-intervention, we constructed samples spanning ±12 months around the assigned date and replicated the DiD analysis presented in Table 5. The resulting estimates are reported in section D of the Online Appendix, Tables D1, D2, and D3. We find no significant effects for the pseudo-interventions at 9 and 12 months. For the 6-month pseudo-intervention, the only significant coefficient is smaller in magnitude and exhibits lower statistical significance. Overall, the placebo estimates provide evidence that our main results are unlikely to be driven by random fluctuations, which reinforces our causal interpretation.

As a second robustness test, we excluded technology companies from our DiD analysis, given that they may have had prior knowledge or expectations regarding the applicability of AI and GenAI. Following the approach of Babina et al (2024), we removed all firms classified under the 2-digit NAICS codes 51 (Information) and 54 (Professional, Scientific, and Technical Services) and re-estimated our models. Tables E1 and E2 in Section E of the Online Appendix replicate Tables 4 and 5 from the main text. The estimates remain largely consistent, suggesting that our main results are not driven by this subset of firms. The only exception is the coefficient for Information Provision (synchronous), which in Table E1 is not statistically significant.

We also replicated our analysis using an alternative method for classifying treated and control units. We identified treated sectors based on the share of job postings in non-routine cognitive occupations, following the classification approach of Jaimovich and Siu (2012). Sectors where more than 50% of job postings in the year preceding the shock were for non-routine cognitive roles were assigned to the treatment group. This method resulted in 1,027 companies classified as treated and 747 as controls (see section F of the Online Appendix for additional details on the classification process). We then replicated the analyses reported in Tables 4 and 5, with the results presented in Tables F1 and F2 of section F of the Online

Appendix. The findings remain largely consistent with our main results, with the exception of the Exception Management conflict category, which shows a negative impact following the introduction of ChatGPT. On the one hand, these findings provide additional reassurance regarding the robustness of our main results. On the other hand, the observed expansion in GenAI's scope of influence further supports the use of GenAI Exposure measure by Felten et al (2023). In particular, while the classification based on non-routine cognitive occupations captures a relevant dimension of AI impact, it may overestimate or underestimate the true effect. This is because it does not explicitly account for the 'language model' capabilities of GenAI systems like GPT-3.5 and GPT-4. Thus, performing a non-routine cognitive occupation is a necessary but not sufficient condition for exposure to GenAI: within these occupations, certain tasks and abilities are more susceptible to automation or augmentation by language models, which ultimately shapes the degree of exposure.

Additionally, we replicated our main analysis using Coarsened Exact Matching (CEM) to further validate the robustness of our findings. We implemented a restrictive k-to-k matching procedure based on the simple moving average of key control variables measured between January 2021 and November 2022: Size, ROA, Leverage, R&D/TA, R&D dummy, Capex/TA, and Workforce Change indicator. This matching process resulted in a balanced sample of 345 treated and 345 control units, totaling 690 companies. The CEM-based estimates are reported in section G of the Online Appendix, where Tables G1 and G2 correspond to the results in Tables 4 and 5 of the main text. The results remain consistent with our baseline analysis, with two exceptions: Information Provision (synchronous) and Reward Distribution (monitoring). For Information Provision (synchronous), the CEM specifications replicate the effect found when using the continuous GenAI Exposure (monthly) measure as per Table 6.

Lastly, we tested three additional specifications reported in Online Appendix H. First, we adjusted the main estimates by controlling for the total hiring of a company in a given month. As shown in Tables H1 and H2, the results remain consistent. Second, we replicated the estimates by taking a seemingly unrelated regression approach to account for potential correlation among the unobserved components across dependent variables. The results in Table H3 confirm our main results, except for a divergence observed in

Exception Management (conflict). Finally, Tables H4 and H5 present aggregate effects across the five organizational problem areas. Task Division, Task Allocation, and Information Provision results remain stable. However, the negative coefficients for Reward Distribution and Exception Management mask underlying heterogeneity across their respective sub-dimensions.

DISCUSSION

This study examines how GenAI affects the set of skills that organizations rely on to address the general problems of organizing. It is based on our theoretical framework that generates three primary predictions about GenAI's impact on organizational problem-solving capabilities. First, GenAI will enable individuals to solve problems more quickly and with broader knowledge application, which reduces the resources needed for traditional task division and scheduling. Second, information gathering and interpretation will become more efficient, which allows a shift from organizational-push to individual-pull mechanisms of information access. Finally, monitoring and performance evaluation can be increasingly automated, reducing the human resources traditionally allocated to these oversight functions. These predictions stem from conceptualizing organizations as dynamic queuing processes with increasingly capable problem-solvers enabled by AI (Glynn et al 2020). Rather than assuming perfect optimization, we recognize that organizations can incrementally improve their work processes in response to technological capabilities, approximating more efficient organizational designs.

Our results show that organizations are mostly deprioritizing skills related to Monitoring for Reward Distribution and Task Division. The deprioritization of Monitoring skills likely stems from GenAI's suitability for performance monitoring tasks. Current AI systems are good at processing structured performance data, consistently applying standardized evaluation criteria, and generating appropriate feedback. These are capabilities that directly substitute for traditional human monitoring skills. Furthermore, GenAI can analyze performance patterns at a scale and speed that human managers cannot match. The significant drop in Monitoring skills suggests that organizations perceive these functions as having lower barriers to AI substitution and potentially higher returns from automation than other organizing functions. Additionally, monitoring tasks, at least from an organization's point of view, often involves structured and predictable decisions based on predefined criteria, making them more amenable to algorithmic approaches than tasks requiring nuanced contextual judgment or complex social interaction.

It is plausible that the increase in automatic surveillance and assessment of work performance reduces the human role in monitoring for distributing rewards. However, it is also possible that GenAI has been applied as an investment in restructuring the division of labor in ways that simplify Monitoring for Reward Distribution by reducing or automating the handling of task interdependence. The significant decrease in Coordination skills supports this interpretation, suggesting that GenAI is changing how organizations manage task interdependencies. Interdependence of tasks is produced by dividing tasks into specialized sub-tasks and assigning expert problem-solvers to each. Yet, GenAI-enabled organizations may better integrate sub-tasks because GenAI enhances problem-solvers' capabilities, thus allowing reduction of interdependence and, as a consequence, the need to share information among problem-solvers. Reducing human effort devoted to Coordination and Monitoring suggests that AI is indeed used to reorganize in ways that reduce task interdependence.

While Monitoring showed a strong decline, Incentives remained largely unchanged. This suggests that while organizations may leverage AI systems to track and assess performance, they are continuing to rely on human judgment for designing and implementing reward systems. This asymmetry may reflect organizations' recognition that effective incentive structures still require contextual understanding and social calibration, which current AI systems lack. The significant impact on Task Division and Asynchronous Information Provision, coupled with the insignificant effects on Task Allocation, suggests a selective set of changes in organizational skill demands. It seems that GenAI is primarily influencing formal and documentation-heavy processes rather than real-time interactions and staffing decisions that involve more nuanced human judgment.

These findings produce clear and distinct contributions. First, they reveal that organizations engage in selective skill deprioritization rather than uniform capability enhancement. This extends research on technological change and organizational adaptation (Barley 1986, Orlikowski 1992, Zammuto et al 2007) by showing how firms strategically reallocate resources away from activities where GenAI provides the strongest complementarities or substitution effects. The significant decreases in job postings related to Task Division, Information Provision, Reward Distribution (monitoring), and Exception Management (operational) indicate that these skill domains are particularly susceptible to AI-induced transformation. This selective pattern aligns with recent work by Shrestha et al (2019) and Waardenburg et al (2021) on how organizations adapt differentially to GenAI capabilities.

Second, our results advance queuing theory applications in organizational contexts (Cohen et al 1972, Glynn et al 2020) by demonstrating how AI-enhanced problem-solving capabilities affect different organizing functions. The substantial decline in Information Provision skills suggests that GenAI is transforming how organizations handle information flows, supporting predictions from queuing models that enhanced individual capabilities reduce the need for organizational information-push mechanisms. Similarly, the decrease in Monitoring skills validates theoretical predictions about how GenAI enables more efficient quality control processes with fewer human resources, supporting recent work on algorithmic management (Bailey et al 2022, Kellogg et al 2020).

Lastly, our study contributes to the emerging literature on human-AI collaboration (Dell'Acqua et al 2023, Choudhary et al 2025, Kellogg et al 2024). While existing research offers micro-level insights into how human and AI agents may collaborate by complementing or substituting each other, it says less about the meso- and macro-level consequences of novel divisions of labor. By examining meso-level organizational adaptations to human-AI ensembles, this study provides insights into the changing nature of work and how organizations are responding with new hiring practices. In so doing, it also considers macro-level implications by reflecting on how GenAI is reshaping the broader labor market. In turn, this is consequential for the organizing and functioning of those ensembles.

Our analysis of organizational responses to GPT-3.5 versus GPT-4 provides additional insights into the mechanisms driving skill deprioritization. The intensification of effects following GPT-4's introduction—particularly in Information Provision (asynchronous), Reward Distribution (monitoring), and Exception Management (operational)—suggests that capability improvement rather than mere awareness drives organizational adaptation. This natural experiment helps distinguish between two potential mechanisms: (1) changes in managerial perception of AI's potential, which would be primarily triggered by the initial GPT-3.5 release; and (2) actual capability improvement enabling more substantial organizational restructuring, which would be more pronounced with GPT-4's superior capabilities. Our results support the latter mechanism, aligning with technology adoption models that emphasize performance expectancy as a key driver of organizational change (Venkatesh et al 2003, Beaudry and Pinsonneault 2005). The stronger effects observed with GPT-4 suggest that organizations respond more decisively when AI demonstrates more advanced capabilities, indicating that skill deprioritization is driven by substantive capability enhancement rather than hype or novelty effects.

In sum, our study offers an examination of the immediate effects of how a broad sample of firms respond to the inflow of AI algorithms through selective deskilling of their labor force. This is a crucial first step, and in future research, we expect to identify the processes that underlie each of these effects. As AI tools continue to improve, we expect to find additional effects of their availability, including the second-order effects predicted by queueing models for organizations that use AI to fundamentally reorganize to allow non-human problem solvers to handle greater parts of their operations. While AI is poised to enhance productivity and create new roles, it will likely necessitate a significant workforce adaptation and likely raise strategic considerations about skills relevant to the organization.

In the medium term, the restructuring of labor division and integration of effort driven by GenAI could still intensify intra-organizational conflict, even though our data currently shows no significant short-run effect on conflict-resolution roles. This prediction aligns with Kellogg et al.'s (2020) argument regarding employee resistance to algorithmic management. We anticipate tensions arising from organizations' efforts to automate labor processes, which then shift accountability for work outcomes onto individual employees. Though we do not observe an immediate spike in conflict-management hiring, such pressures could accumulate over time as GenAI-based surveillance and reduced interdependence among workers become more apparent—conditions employees are likely to resist. Eventually, organizations may respond to this heightened conflict either by temporarily increasing demand for conflict-resolution skills or by employing alternative strategies, such as restructuring internal workflows, to proactively minimize human-AI conflict occurrences. This deserves further investigation.

In the longer term, we could also expect the emergence of novel organizational forms that blend human and AI capabilities in ways that differ from traditional hierarchical structures. The reduced need for coordination and monitoring might enable organizations to operate with significantly flatter hierarchies (Lee and Edmondson 2017) or network-based structures that were previously impractical due to coordination costs. We may expect an increased managerial span of control, diminishing the need for multiple hierarchical layers and overall managerial intensity, because AI systems can handle coordination and monitoring tasks. This can allow organizational structures in which humans focus primarily on strategic decision-making, with fewer operational decisions made by humans. Such structures raise the question of how organizations will develop and maintain unique capabilities when key knowledge-generating activities become increasingly mediated by GenAI. For example, as organizations like Google potentially outsource coding tasks to GenAI, how will they ensure the development of proprietary technical capabilities that serve as sources of competitive advantage? The risk of organizational knowledge hollowing—where critical tacit knowledge erodes as tasks are delegated to GenAI—represents a significant strategic challenge. One solution may be for organizations to develop new approaches to knowledge management that explicitly account for human-AI collaborative learning, where AI enhances rather than substitutes for organizational learning processes. This could involve designating certain knowledge domains as strategically critical and maintaining human expertise in these areas, even as other domains are increasingly delegated to AI systems.

While our study provides important insights, some limitations should be noted. First, our analysis is based on publicly listed U.S. companies, which may not fully capture the dynamics in privately held firms or in other national contexts. It is possible that the effects are particularly pronounced for large companies such as those we study. Publicly listed companies are generally large and are likely to have more formalized hiring processes and better access to cutting-edge technologies compared to smaller or non-U.S. firms. This could limit the extent to which our findings apply to smaller or privately held firms or those operating in different regulatory and economic environments. Future research should consider extending the analysis to different types of organizations and geographical and regulatory regimes to assess the broader applicability of our results.

Second, our classification of skills relies on a method using topic modeling and manual labeling. While we achieved high validation accuracy, there is always a risk of misclassification or incomplete capture of the nuanced organizing functions of skills in different contexts. Additional research is needed to assess whether alternative approaches to skill classification produce the same results.

Third, our identification strategy uses the release of ChatGPT as an exogenous shock, but there may be potential spillovers between the treatment and control groups. If control firms are also exposed to information about AI capabilities, estimated effects could be attenuated. This means that our approach may be conservative, as the effects would likely be even stronger with a proper control group.

CONCLUSION

Our study opens several exciting avenues for future research. Our findings raise questions about the broader structural implications of AI adoption. For instance, studies on flat organizations (Lee and Edmondson 2017, Foss and Klein 2022) and multiple organizational goals (Audia and Greve 2021, Battilana et al 2022) may offer alternative explanations for why certain skills, such as those in conflict resolution, remain unchanged. Future research could investigate whether and how firms are reconfiguring their organizational structures—perhaps moving toward flatter models—to better align with the evolving technological landscape. The new forms of organizing allowed by the availability of AI should be a central topic of management research, and we expect and encourage work drawing from the theoretical ideas and empirical methods employed in this paper.

TABLE 1
Universal Problems of Organizing

Category	Synopsis	Example of Skills
Task Division	*The process of breaking down work into smaller components, focusing on sorting, sequencing, and structuring tasks before execution.*	Process Sequencing, Planning, Scheduling, Activity Sequencing, Prioritization.
Task Allocation		
• Staffing	*Recruiting and selecting individuals for specific tasks based on skill fit and job requirements.*	Human Resource Planning, Human Resource Strategy, Capacity Planning, Staff Planning, Workforce Planning Operations.
• Mapping	*Assigning specific human or non-human resources (tools, technology) to tasks.*	Assigning Employees, Competency Mapping, Organizational Architecture, Resource Allocation, Delegated Authority.
Information Provision		
• Coordination	*Information sharing to synchronize efforts among different individuals or teams.*	Cross-Functional Coordination, Process Integration, Collaboration, Coordinating, Stakeholder Coordination.
• Synchronous	*Information sharing via real-time communication (e.g., meetings, calls).*	Verbal Communication Skills, Speech Fluency, Social Communications, Professional Speaking, Interactive Communications.
• Asynchronous	*Information sharing via asynchronous, documented communication such as emails, reports, and wikis*	Report Writing, Professional Writing, Structured Writing, Project Documentation, Writing Outlines.
Reward Distribution		
• Monitoring	*Monitoring or measuring employees' tasks and behaviors to assess employees' performance.*	Employee Monitoring, Plan Of Action And Milestones (POA&M), Employee Performance Management, Workforce Productivity, Performance Review.
• Incentives	*Designing and implementing motivational rewards (monetary or non-monetary) to encourage desired behavior.*	Reward Management, Executive Compensation Strategy, High Potential Programs, Benefits Strategies, Compensation Strategy.
Exception Management		
• Conflict	*Identifying and addressing conflict between employees.*	Employee Conflict Resolution, Conflict Resolution, Conflict Transformation, De-escalation Techniques, Organizational Conflict.
• Operational	*Identifying, analyzing, and resolving unexpected operational problems.*	Overcoming Obstacles, A3 Problem Solving Techniques, Creative Problem Solving, Root Cause Corrective Action, Deviation Investigations.

TABLE 2
Summary Statistics

	N	Mean	SD	p10	Median	p90
EM (conflict)	39122	13.98	72.92	0	1	24
EM (operational)	39122	65.21	249.93	0	6	132
IP (asynchronous)	39122	61.47	239.63	0	5	117
IP (coordination)	39122	54.25	235.78	0	4	100
IP (synchronous)	39122	77.42	333.10	0	6	144
RD (incentives)	39122	1.48	8.05	0	0	3
RD (monitoring)	39122	25.24	109.02	0	1	46
TA (mapping)	39122	13.32	50.96	0	1	27
TA (staffing)	39122	5.29	45.47	0	0	7
Task Division	39122	91.55	338.04	0	8	186
GenAI Exposure (dummy)	39122	0.55	0.50	0	1	1
GenAI Exposure (monthly)	39122	0.34	0.31	-0.08	0.36	0.72
Size	39122	7.85	2.15	5.16	7.83	10.66
ROA	39122	-0.01	0.08	-0.08	0.01	0.04
Leverage	39122	0.30	0.27	0.03	0.27	0.59
R&D/TA	39122	0.02	0.04	0.00	0.00	0.06
R&D (dummy)	39122	0.44	0.50	0.00	0.00	1.00
CAPEX/TA	39122	0.02	0.03	0.00	0.01	0.05
WFC	39122	-0.05	0.09	-0.12	-0.04	0.00

TABLE 3
Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) EM (conflict)	1.00									
(2) EM (operational)	0.53*	1.00								
(3) IP (asynchronous)	0.68*	0.74*	1.00							
(4) IP (coordination)	0.61*	0.64*	0.83*	1.00						
(5) IP (synchronous)	0.70*	0.65*	0.80*	0.81*	1.00					
(6) RD (incentives)	0.33*	0.44*	0.41*	0.43*	0.43*	1.00				
(7) RD (monitoring)	0.50*	0.69*	0.80*	0.75*	0.59*	0.38*	1.00			
(8) TA (mapping)	0.47*	0.72*	0.70*	0.66*	0.60*	0.43*	0.75*	1.00		
(9) TA (staffing)	0.26*	0.47*	0.41*	0.39*	0.33*	0.21*	0.55*	0.44*	1.00	
(10) Task Division	0.70*	0.77*	0.91*	0.88*	0.81*	0.45*	0.82*	0.73*	0.43*	1.00

TABLE 4
Demand for Skills after the GenAI availability shock

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task Division	Task Allocation		Information Provision			Reward Distribution		Exception Management	
		(staffing)	(mapping)	(coord)	(sync)	(async)	(monitor)	(incentive)	(conflict)	(operational)
PostGPT* Treated	-0.268***	0.165	-0.102	-0.246**	-0.210*	-0.250**	-0.275***	-0.151	-0.236	-0.217**
	(0.068)	(0.315)	(0.082)	(0.084)	(0.088)	(0.082)	(0.078)	(0.105)	(0.132)	(0.075)
Constant	6.286***	3.988***	4.405***	5.938***	6.232***	5.927***	5.190***	2.400***	4.766***	5.937***
	(0.012)	(0.064)	(0.019)	(0.016)	(0.017)	(0.015)	(0.014)	(0.022)	(0.024)	(0.015)
N	37405	29667	33193	36397	37236	37115	34540	25740	32683	36780
Pseudo R^2	0.929	0.819	0.865	0.919	0.920	0.919	0.903	0.685	0.882	0.918

Notes. ± 12 Months from ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. * p<0.05, ** p<0.01, *** p<0.001

TABLE 5
Demand for Skills after the GenAI availability shock with control variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task Division	Task Allocation		Information Provision			Reward Distribution		Exception Management	
		(staffing)	(mapping)	(coord)	(sync)	(async)	(monitor)	(incentive)	(conflict)	(operational)
PostGPT*Treated	-0.270***	0.177	-0.096	-0.245**	-0.212*	-0.252**	-0.272***	-0.150	-0.232	-0.218**
	(0.068)	(0.311)	(0.082)	(0.083)	(0.087)	(0.083)	(0.079)	(0.102)	(0.132)	(0.076)
Size	0.316	-0.079	0.001	0.046	0.214	0.177	0.125	0.328	0.388	0.068
	(0.198)	(0.232)	(0.184)	(0.247)	(0.256)	(0.218)	(0.227)	(0.334)	(0.296)	(0.234)
ROA	-0.126	-1.543	0.352	0.199	0.616	0.471	-0.121	-0.397	0.221	0.311
	(0.588)	(2.034)	(0.681)	(0.705)	(0.619)	(0.614)	(0.731)	(1.126)	(0.685)	(0.561)
Leverage	0.084	-1.576	-0.469	0.520	0.214	0.023	-0.182	-0.453	0.153	0.148
	(0.439)	(1.093)	(0.461)	(0.584)	(0.567)	(0.506)	(0.512)	(0.676)	(0.731)	(0.472)
R&D/TA	-2.887	-15.521	-5.633	-4.823	-6.145	-1.543	-5.825	-7.451	-4.782	-1.361
	(3.448)	(10.028)	(4.668)	(3.860)	(4.708)	(2.980)	(3.982)	(5.609)	(4.408)	(4.081)
R&D (dummy)	0.013	0.306	0.060	0.059	0.093	0.089	0.101	0.300*	0.051	-0.034
	(0.065)	(0.160)	(0.074)	(0.073)	(0.090)	(0.064)	(0.077)	(0.118)	(0.116)	(0.095)
CAPEX/TA	0.264	-0.040	0.533	0.377	-0.185	-0.089	0.159	0.701	0.377	-0.611
	(0.958)	(1.384)	(1.110)	(1.081)	(1.119)	(0.901)	(0.832)	(1.254)	(0.936)	(0.930)
WFC	0.416*	0.047	-0.055	0.235	0.231	0.306	0.091	0.484	0.166	0.263
	(0.174)	(0.298)	(0.238)	(0.161)	(0.183)	(0.170)	(0.159)	(0.400)	(0.262)	(0.158)
Constant	3.008	5.356*	4.546*	5.267*	3.950	4.082	3.950	-0.825	0.639	5.216*
	(2.047)	(2.477)	(1.939)	(2.573)	(2.631)	(2.252)	(2.341)	(3.419)	(3.173)	(2.442)
N	37405	29667	33193	36397	37236	37115	34540	25740	32683	36780
Pseudo R^2	0.929	0.819	0.865	0.919	0.921	0.919	0.903	0.686	0.883	0.918

Notes. ± 12 Months from ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. * p<0.05, ** p<0.01, *** p<0.001

TABLE 6
Demand for Skills after the GenAI availability shock with Industry-level GenAI exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task Division	Task Allocation		Information Provision			Reward Distribution		Exception Management	
		(staffing)	(mapping)	(coord)	(sync)	(async)	(monitor)	(incentive)	(conflict)	(operational)
PostGPT*GenAI Exposure	-0.337**	0.417	-0.185	-0.297*	-0.200	-0.324*	-0.349**	-0.158	-0.263	-0.305*
	(0.114)	(0.539)	(0.111)	(0.136)	(0.141)	(0.132)	(0.117)	(0.151)	(0.164)	(0.119)
GenAI Exposure	1.466***	0.787	2.341***	1.355***	1.460***	1.504***	2.222***	1.490*	1.225*	1.567***
	(0.317)	(0.913)	(0.397)	(0.407)	(0.432)	(0.372)	(0.364)	(0.600)	(0.544)	(0.348)
Size	0.283	-0.067	-0.010	0.038	0.200	0.146	0.102	0.321	0.389	0.054
	(0.185)	(0.217)	(0.164)	(0.231)	(0.241)	(0.204)	(0.205)	(0.329)	(0.301)	(0.214)
ROA	-0.207	-1.575	0.139	0.116	0.490	0.413	-0.234	-0.485	0.182	0.220
	(0.575)	(1.916)	(0.601)	(0.684)	(0.610)	(0.602)	(0.697)	(1.093)	(0.683)	(0.530)
Leverage	-0.086	-1.729	-0.718	0.363	0.050	-0.113	-0.416	-0.614	0.086	-0.047
	(0.446)	(1.033)	(0.422)	(0.608)	(0.570)	(0.506)	(0.501)	(0.676)	(0.765)	(0.435)
R&D/TA	-2.906	-14.067	-4.893	-4.696	-6.070	-1.540	-5.455	-7.023	-4.725	-1.403
	(3.345)	(9.308)	(4.377)	(3.666)	(4.604)	(2.782)	(3.673)	(5.544)	(4.354)	(4.061)
R&D (dummy)	0.020	0.288	0.058	0.065	0.102	0.097	0.105	0.301*	0.063	-0.028
	(0.065)	(0.149)	(0.068)	(0.072)	(0.089)	(0.062)	(0.075)	(0.118)	(0.116)	(0.095)
CAPEX/TA	0.131	-0.282	0.417	0.196	-0.379	-0.251	-0.054	0.658	0.245	-0.773
	(0.951)	(1.280)	(1.070)	(1.070)	(1.099)	(0.898)	(0.772)	(1.235)	(0.972)	(0.931)
WFC	0.425*	0.027	-0.070	0.234	0.242	0.309	0.109	0.472	0.158	0.251
	(0.175)	(0.310)	(0.223)	(0.158)	(0.187)	(0.166)	(0.163)	(0.396)	(0.259)	(0.151)
Constant	2.908	4.950*	3.820*	4.957*	3.645	3.923	3.546	-1.275	0.205	4.858*
	(1.890)	(2.304)	(1.740)	(2.378)	(2.452)	(2.079)	(2.103)	(3.298)	(3.187)	(2.233)
N	37405	29667	33193	36397	37236	37115	34540	25740	32683	36780
Pseudo R^2	0.930	0.820	0.869	0.920	0.922	0.920	0.906	0.687	0.883	0.919

Notes. ± 12 Months from ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. * p<0.05, ** p<0.01, *** p<0.001

TABLE 7
Demand for Skills post GPT3 and GPT4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task Division	Task Allocation		Information Provision			Reward Distribution		Exception Management	
		(staffing)	(mapping)	(coord)	(sync)	(async)	(monitor)	(incentive)	(conflict)	(operational)
GPT3*Treated	-0.268**	-0.271	-0.098	-0.275**	-0.209	-0.227*	-0.171	-0.120	-0.220	-0.154
	(0.093)	(0.165)	(0.096)	(0.107)	(0.115)	(0.109)	(0.091)	(0.132)	(0.136)	(0.102)
GPT4*Treated	-0.270***	0.295	-0.095	-0.235**	-0.213*	-0.260**	-0.306***	-0.162	-0.236	-0.241**
	(0.069)	(0.365)	(0.087)	(0.085)	(0.087)	(0.080)	(0.083)	(0.113)	(0.136)	(0.073)
Size	0.316	-0.099	0.001	0.045	0.214	0.178	0.130	0.330	0.388	0.070
	(0.198)	(0.226)	(0.184)	(0.247)	(0.256)	(0.218)	(0.228)	(0.336)	(0.296)	(0.234)
ROA	-0.126	-1.510	0.351	0.201	0.617	0.470	-0.129	-0.398	0.222	0.306
	(0.589)	(1.958)	(0.677)	(0.704)	(0.618)	(0.615)	(0.735)	(1.127)	(0.684)	(0.562)
Leverage	0.084	-1.601	-0.470	0.515	0.214	0.027	-0.164	-0.450	0.154	0.159
	(0.437)	(1.101)	(0.460)	(0.581)	(0.563)	(0.503)	(0.512)	(0.672)	(0.731)	(0.474)
R&D/TA	-2.887	-15.448	-5.633	-4.824	-6.145	-1.539	-5.880	-7.447	-4.777	-1.371
	(3.449)	(9.992)	(4.668)	(3.859)	(4.708)	(2.980)	(3.973)	(5.612)	(4.409)	(4.081)
R&D (dummy)	0.013	0.302	0.060	0.059	0.093	0.089	0.102	0.301*	0.051	-0.033
	(0.065)	(0.159)	(0.074)	(0.073)	(0.090)	(0.064)	(0.077)	(0.118)	(0.116)	(0.094)
CAPEX/TA	0.265	-0.384	0.532	0.373	-0.184	-0.081	0.200	0.696	0.384	-0.604
	(0.952)	(1.450)	(1.110)	(1.073)	(1.114)	(0.903)	(0.845)	(1.261)	(0.936)	(0.932)
WFC	0.416*	0.054	-0.055	0.234	0.231	0.305	0.090	0.488	0.166	0.265
	(0.175)	(0.288)	(0.239)	(0.161)	(0.183)	(0.170)	(0.158)	(0.396)	(0.262)	(0.159)
Constant	3.008	5.585*	4.547*	5.274*	3.949	4.075	3.899	-0.846	0.638	5.192*
	(2.045)	(2.437)	(1.940)	(2.567)	(2.631)	(2.253)	(2.349)	(3.437)	(3.174)	(2.447)
N	37405	29667	33193	36397	37236	37115	34540	25740	32683	36780
Pseudo R^2	0.929	0.820	0.865	0.919	0.921	0.919	0.903	0.686	0.883	0.918

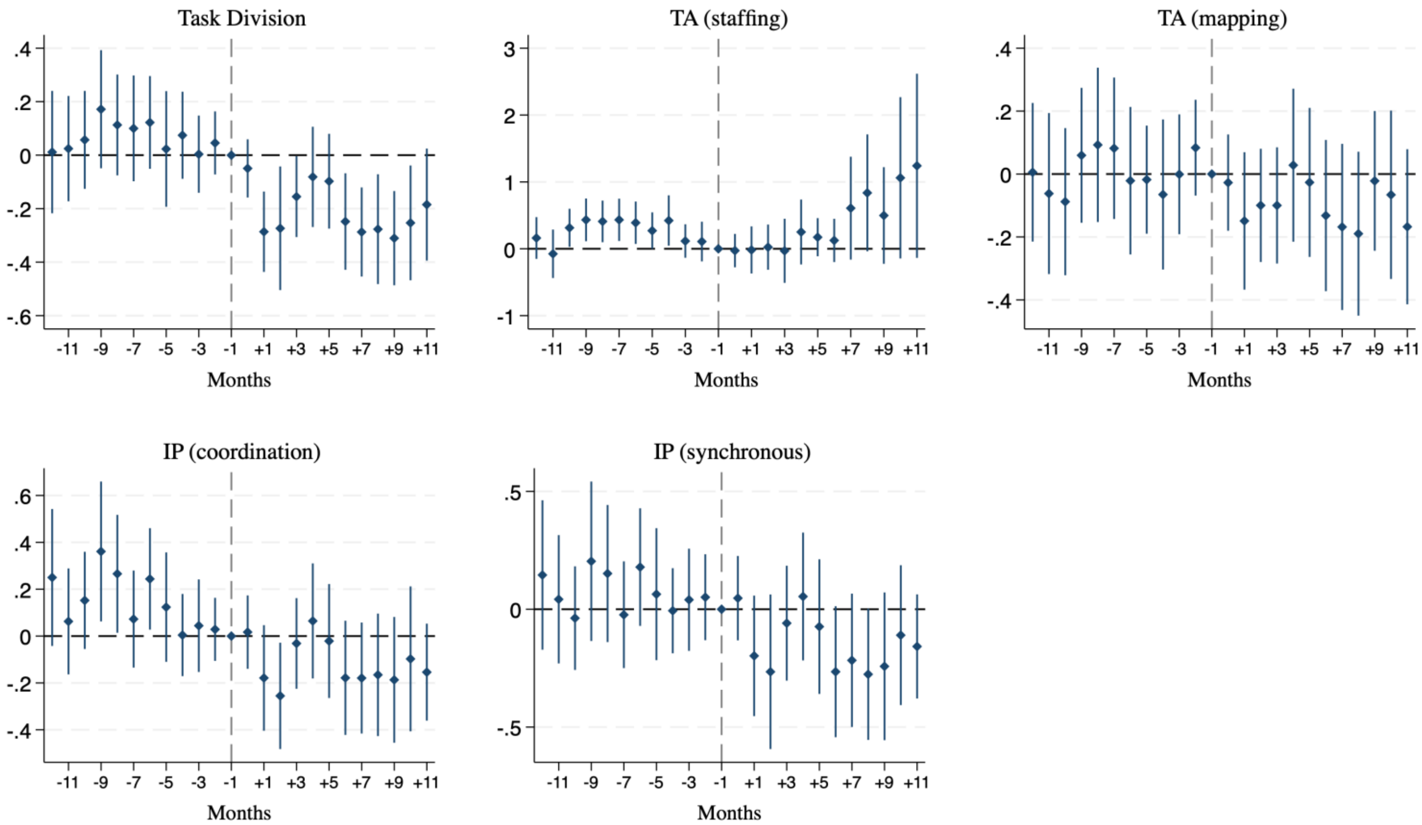
Notes. ± 12 Months from ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

TABLE 8
Demand for Skills post GPT3.5 and GPT4 with Industry-level GenAI exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task Division	Task Allocation		Information Provision			Reward Distribution		Exception Management	
		(staffing)	(mapping)	(coord)	(sync)	(async)	(monitor)	(incentive)	(conflict)	(operational)
GPT3.5*GenAI Exposure	-0.277	-0.240	-0.009	-0.308	-0.163	-0.245	-0.152	0.016	-0.211	-0.142
	(0.165)	(0.259)	(0.148)	(0.175)	(0.181)	(0.187)	(0.147)	(0.225)	(0.195)	(0.177)
GPT4*GenAI Exposure	-0.357**	0.589	-0.242*	-0.293*	-0.212	-0.350**	-0.411***	-0.222	-0.280	-0.359***
	(0.110)	(0.648)	(0.119)	(0.139)	(0.140)	(0.124)	(0.117)	(0.160)	(0.166)	(0.109)
GenAI Exposure	1.476***	0.741	2.373***	1.353**	1.467***	1.519***	2.249***	1.529*	1.232*	1.598***
	(0.321)	(0.925)	(0.396)	(0.413)	(0.436)	(0.377)	(0.363)	(0.600)	(0.545)	(0.343)
Size	0.283	-0.080	-0.009	0.038	0.200	0.146	0.103	0.327	0.389	0.054
	(0.185)	(0.213)	(0.165)	(0.231)	(0.241)	(0.204)	(0.206)	(0.330)	(0.301)	(0.215)
ROA	-0.214	-1.551	0.147	0.117	0.490	0.410	-0.256	-0.506	0.183	0.211
	(0.580)	(1.858)	(0.615)	(0.686)	(0.611)	(0.605)	(0.704)	(1.096)	(0.683)	(0.533)
Leverage	-0.085	-1.761	-0.701	0.363	0.052	-0.108	-0.404	-0.609	0.089	-0.031
	(0.446)	(1.049)	(0.419)	(0.605)	(0.568)	(0.504)	(0.501)	(0.670)	(0.764)	(0.435)
R&D/TA	-2.901	-14.123	-4.841	-4.697	-6.060	-1.523	-5.459	-6.960	-4.712	-1.379
	(3.344)	(9.348)	(4.376)	(3.665)	(4.600)	(2.776)	(3.651)	(5.537)	(4.353)	(4.037)
R&D (dummy)	0.020	0.288	0.057	0.065	0.102	0.096	0.105	0.302*	0.063	-0.028
	(0.065)	(0.150)	(0.068)	(0.072)	(0.089)	(0.062)	(0.075)	(0.118)	(0.116)	(0.094)
CAPEX/TA	0.135	-0.502	0.407	0.199	-0.385	-0.254	-0.047	0.607	0.256	-0.802
	(0.955)	(1.316)	(1.083)	(1.073)	(1.109)	(0.908)	(0.781)	(1.254)	(0.973)	(0.937)
WFC	0.425*	0.032	-0.064	0.233	0.244	0.310	0.112	0.482	0.161	0.255
	(0.175)	(0.306)	(0.225)	(0.158)	(0.186)	(0.166)	(0.163)	(0.391)	(0.259)	(0.151)
Constant	2.900	5.124*	3.800*	4.958*	3.638	3.918	3.518	-1.348	0.202	4.842*
	(1.890)	(2.282)	(1.751)	(2.376)	(2.452)	(2.079)	(2.105)	(3.309)	(3.185)	(2.236)
N	37405	29667	33193	36397	37236	37115	34540	25740	32683	36780
Pseudo R^2	0.930	0.821	0.869	0.920	0.922	0.920	0.906	0.687	0.883	0.920

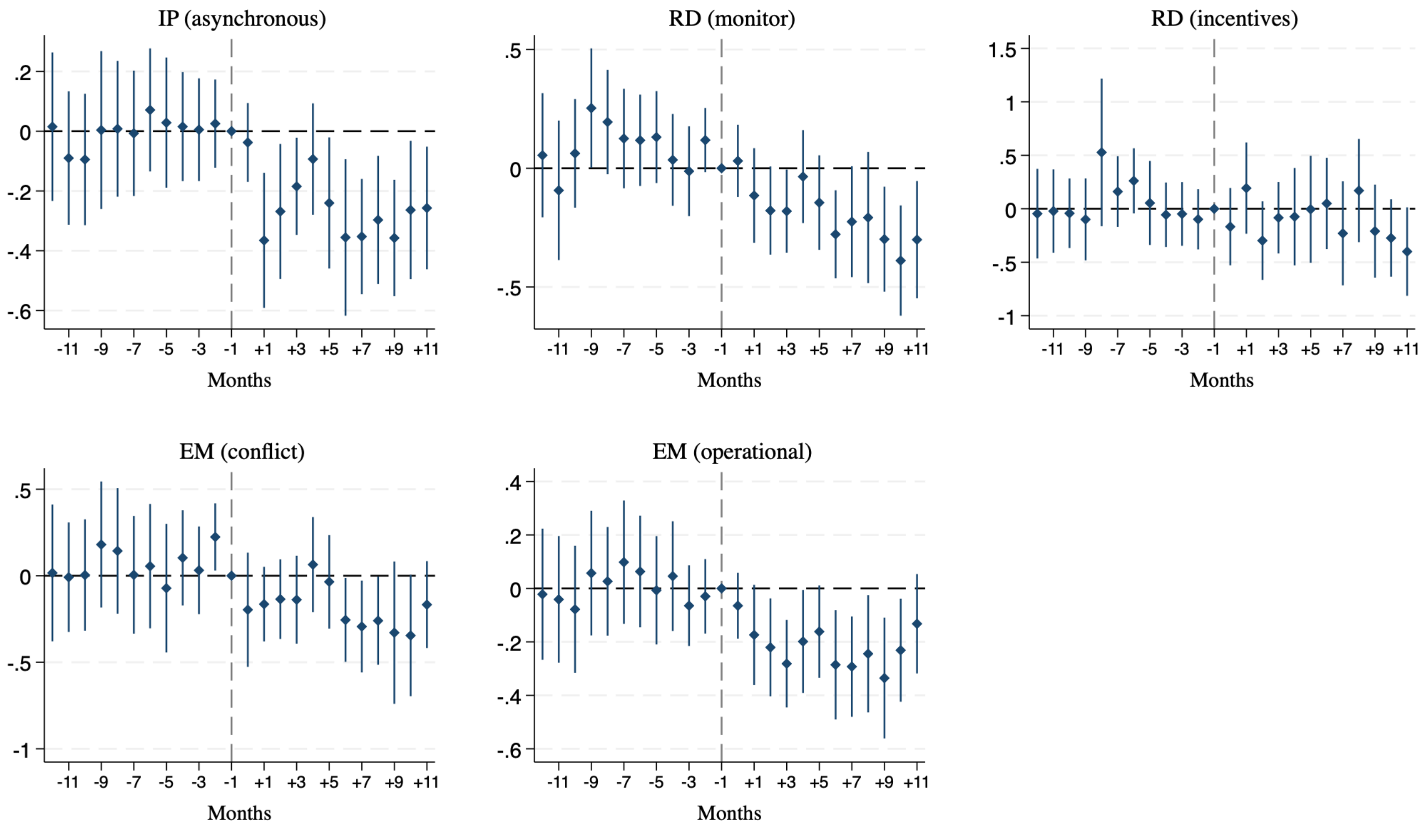
Notes. ± 12 Months from ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. * p<0.05, ** p<0.01, *** p<0.001

FIGURE 1
Dynamic Effects



Notes. These graphs show the effect of ChatGPT on problems of organizing. The vertical axis shows Poisson coefficients (and 95 percent confidence intervals). The horizontal axis shows time relative to treatment. The coefficients are obtained from a regression of $Y = \alpha + \Sigma z\, \beta_1 * 1(z) + \Upsilon_i + \delta_t + G_i + \varepsilon_{it}$ where Υ_i is Firm FE, δ_t is Year-Month FE, G_i is the set of controls, and ε_{it} is the error term. z represents the "lag," or the months relative to a "zero month," which marks the year when ChatGPT is introduced.

FIGURE 1 (cont'd)
Dynamic Effects



REFERENCES

Adan I, Resing J (2015) *Queueing Systems*. Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, The Netherlands. Unpublished manuscript, March 26.

Anthony C, Bechky BA, Fayard AL (2023) Collaborating with AI: Taking a system view to explore the future of work. *Organ. Sci.* 34(1):123–140.

Audia PG, Greve HR (2021) *Organizational Learning from Performance Feedback: A Behavioral Perspective on Multiple Goals* (Cambridge University Press, Cambridge, UK).

Autor DH, Levy F, Murnane RJ (2003) The skill content of recent technological change: An empirical exploration. *Quart. J. Econom.* 118(4):1279–1333.

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

Bailey DE, Faraj S, Hinds P, von Krogh G, Leonardi PM (2022) We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organ. Sci.* 33(5):1123–1140.

Barley SR (1986) Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Admin. Sci. Quart.* 31(1):78–108.

Battilana J, Yen J, Ferreras I, Ramarajan L (2022) Democratizing work: Redistributing power in organizations for a democratic and sustainable future. *Organ. Theory* 3(1):26317877221084714.

Beaudry A, Pinsonneault A (2005) Understanding user responses to information technology: A coping model of user adaptation. *MIS Quart.* 29(3):493–524.

Brynjolfsson E, Mitchell T (2017) What can machine learning do? Workforce implications. *Science* 358(6370):1530–1534.

Burns T, Stalker GM (1961) *Mechanistic and Organic Systems* (Tavistock Publications, London, UK).

Burton RM, Obel B (2004) What is an organizational design? Burton RM, Obel B, eds. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit* (Springer, New York), 43–85.

Choi S, Kang H, Kim N, Kim J (2022) How does artificial intelligence improve human decision-making? Evidence from the AI-powered Go program. *Strategic Management J.* Forthcoming.

Choudhary V, Marchetti A, Shrestha YR, Puranam P (2025) Human-AI ensembles: When can they work? *J. Management* 51(2):536–569.

Choudhury P, Starr E, Agarwal R (2020) Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management J.* 41(9):1381–1411.

Cohen MD, March JG, Olsen JP (1972) A garbage can model of organizational choice. *Admin. Sci. Quart.* 17(1):1–25.

Dell'Acqua F, Kogut B, Perkowski P (2023) Super Mario meets AI: Experimental effects of automation and skills on team performance and coordination. *Rev. Econom. Statist.* Forthcoming.

Deming D, Kahn LB (2018) Skill requirements across firms and labor markets: Evidence from job postings for professionals. *J. Labor Econom.* 36(S1):S337–S369.

Demirci O, Hannane J, Zhu X (2025) Who is AI replacing? The impact of generative AI on online freelancing platforms. *Management Sci.* Forthcoming.

Dixon J, Hong B, Wu L (2021) The robot revolution: Managerial and employment consequences for firms. *Management Sci.* 67(9):5586–5605.

Donaldson L, Joffe G (2014) Fit–the key to organizational design. *J. Organ. Design* 3(3):38–45.

Felten E, Raj M, Seamans R (2021) Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management J.* 42(12):2195–2217.

Felten EW, Raj M, Seamans R (2023) Occupational heterogeneity in exposure to generative AI. Working paper, available at SSRN 4414065.

Foss NJ, Klein PG (2022) *Why Managers Matter: The Perils of the Bossless Company* (Hachette, London, UK).

Foss NJ, Lyngsie J, Zahra SA (2013) The role of external knowledge sources and organizational design in the process of opportunity exploitation. *Strategic Management J.* 34(12):1453–1471.

Glynn PW, Greve HR, Rao H (2020) Relining the garbage can of organizational decision-making: Modeling the arrival of problems and solutions as queues. *Indust. Corporate Change* 29(1):125–142.
Hui X, Reshef O, Zhou L (2024) The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organ. Sci.* 35(6):1977–1989.
Jaimovich N, Siu HE (2012) The trend is the cycle: Job polarization and jobless recoveries. NBER Working Paper 18334.
Joseph J, Ocasio W (2012) Architecture, attention, and adaptation in the multibusiness firm: General Electric from 1951 to 2001. *Strategic Management J.* 33(6):633–660.
Joseph J, Sengul M (2025) Organization design: Current insights and future research directions. *J. Management* 51(1):249–308.
Kellogg KC, Lifshitz-Assaf H, Randazzo S, Mollick ER, Dell'Acqua F, McFowland III E, Candelon F, Lakhani KR (2024) Don't expect juniors to teach senior professionals to use generative AI: Emerging technology risks and novice AI risk mitigation tactics. Harvard Business School Technology & Operations Management Unit Working Paper No. 24-074. https://doi.org/10.2139/ssrn.4857373.
Kellogg KC, Valentine MA, Christin A (2020) Algorithms at work: The new contested terrain of control. *Acad. Management Ann.* 14(1):366–410.
Law KK, Shen M (2024) How does artificial intelligence shape audit firms? *Management Sci.* Forthcoming.
Lawrence PR, Lorsch JW (1967) Differentiation and integration in complex organizations. *Admin. Sci. Quart.* 12(1):1–47.
Lebovitz S, Lifshitz-Assaf H, Levina N (2022) To engage or not to engage with AI for critical judgments: How professionals deal with opacity. *Organ. Sci.* 33(1):122–139.
Lee MY, Edmondson AC (2017) Self-managing organizations: Exploring the limits of less-hierarchical organizing. *Res. Organ. Behav.* 37:35–58.
Lieberman MB, Lee GK, Folta TB (2017) Entry, exit, and the potential for resource redeployment. *Strategic Management J.* 38(3):526–544.
March JG, Simon HA (1958) *Organizations* (John Wiley & Sons, New York).
Martin JA, Eisenhardt KM (2010) Rewiring: Cross-business-unit collaborations in multibusiness organizations. *Acad. Management J.* 53(2):265–301.
McEvily B, Soda G, Tortoriello M (2014) More formally: Rediscovering the missing link between formal organization and informal social structure. *Acad. Management Ann.* 8(1):299–345.
Mintzberg H (1979) *The Structuring of Organizations* (Pearson, Upper Saddle River, NJ).
Orlikowski WJ (1992) The duality of technology: Rethinking the concept of technology in organizations. *Organ. Sci.* 3(3):398–427.
Puranam P (2018) *The Microstructure of Organizations* (Oxford University Press, Oxford, UK).
Puranam P (2024) *Re-Humanize: How to Build Human-Centric Organizations in the Age of Algorithms* (Penguin Random House SEA, Singapore).
Puranam P, Alexy O, Reitzig M (2014) What's "new" about new forms of organizing? *Acad. Management Rev.* 39(2):162–180.
Rambachan A, Roth J (2023) A more credible approach to parallel trends. *Rev. Econom. Stud.* 90(5):2555–2591.
Raveendran M, Puranam P, Warglien M (2022) Division of labor through self-selection. *Organ. Sci.* 33(2):810–830.
Raveendran M, Silvestri L, Gulati R (2020) The role of interdependence in the micro-foundations of organization design: Task, goal, and knowledge interdependence. *Acad. Management Ann.* 14(2):828–868.
Rerup C (2009) Attentional triangulation: Learning from unexpected rare crises. *Organ. Sci.* 20(5):876–893.
Shrestha YR, Ben-Menahem SM, Von Krogh G (2019) Organizational decision-making structures in the age of artificial intelligence. *California Management Rev.* 61(4):66–83.
Simon HA (1947) A comment on "The Science of Public Administration." *Public Admin. Rev.* 7(3):200–203.

Tandon P, Gupta A, Khanna T (2024) Balancing allocative and dynamic efficiency with redundant R&D allocation. *Strategic Management J.* 45(1):50–78. https://doi.org/10.1002/smj.3645.

Thompson JD (1967) *Organizations in Action: Social Science Bases of Administrative Theory* (McGraw-Hill, New York).

Tonellato M, Tasselli S, Conaldi G, Lerner J, Lomi A (2024) A microstructural approach to self-organizing: The emergence of attention networks. *Organ. Sci.* 35(2):496–524.

Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: Toward a unified view. *MIS Quart.* 27(3):425–478.

Waardenburg L, Huysman M, Sergeeva AV (2021) In the land of the blind, the one-eyed man is king: Knowledge brokerage in the age of learning algorithms. *Organ. Sci.* 32(6):1478–1498.

Wang W, Gao G, Agarwal R (2024) Friend or foe? Teaming between artificial intelligence and workers with variation in experience. *Management Sci.* 70(9):5753–5775.

Weber M (1978) *Economy and Society*, 2 vols. Roth G, Wittich C, trans. and eds. (University of California Press, Berkeley, CA).

Wuebker R, Zenger T, Felin T (2023) The theory-based view: Entrepreneurial microfoundations, resources, and choices. *Strategic Management J.* 44(12):2922–2949.

Yilmaz ED, Naumovska I, Aggarwal VA (2023) AI-driven labor substitution: Evidence from Google Translate and ChatGPT. Working paper.

Zammuto RF, Griffith TL, Majchrzak A, Dougherty DJ, Faraj S (2007) Information technology and the changing fabric of organization. *Organ. Sci.* 18(5):749–762.