Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Skill Deprioritization: Reorganizing in the Age of Generative Artificial Intelligence

Matteo Devigili, Erdem Dogukan Yilmaz, Vibha Gaba, Henrich Greve

To cite this article:
Matteo Devigili, Erdem Dogukan Yilmaz, Vibha Gaba, Henrich Greve (2026) Skill Deprioritization: Reorganizing in the Age of Generative Artificial Intelligence. Management Science

Published online in Articles in Advance 24 Jul 2026

. https://doi.org/10.1287/mnsc.2025.01859

informs
https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–21
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Skill Deprioritization: Reorganizing in the Age of Generative Artificial Intelligence

Matteo Devigili,[a] Erdem Dogukan Yilmaz,[b,*] Vibha Gaba,[a] Henrich Greve[a]

[a] Entrepreneurship and Family Enterprise, INSEAD, Singapore 138676; [b] Entrepreneurship and Family Enterprise, INSEAD, Fontainebleau 77300, France

*Corresponding author

Contact: matteo.devigili@insead.edu, https://orcid.org/0009-0009-3680-4645 (MD); erdemdogukan.yilmaz@insead.edu, https://orcid.org/0000-0002-2727-3557 (EDY); vibha.gaba@insead.edu, https://orcid.org/0000-0002-0182-8948 (VG); henrich.greve@insead.edu, https://orcid.org/0000-0002-7927-6573 (HG)

Received: May 12, 2025
Revised: November 14, 2025; February 7, 2026; February 18, 2026
Accepted: February 23, 2026
Published Online in Articles in Advance: July 24, 2026

https://doi.org/10.1287/mnsc.2025.01859

Abstract. How does generative artificial intelligence (GenAI) reshape the skills that organizations seek as they adapt to a new general-purpose technology? GenAI effectively retrieves data, performs analysis, and conveys information, so it can substitute for workers doing these activities and complement workers relying on them. A natural consequence is skill deprioritization, a systematic reduction in firms' demand for human skills that GenAI can effectively address as organizations adjust the division of labor and integration of effort. We draw on a theoretically grounded classification of organizing skills—task division, task allocation, information provision, reward distribution, and exception management—and a queuing theory model of organizing efficiency to predict which skills firms will deprioritize first. Using a quasiexperimental design that leverages the introduction of ChatGPT as an exogenous shock, we analyze 1,820 publicly listed U.S. companies and track changes in their hiring demand patterns over a period of a ±12-month window surrounds the shock. We find significant declines in demand for monitoring (reward distribution), operational exceptions, and task division skills, with information provision also showing declines. Task allocation and conflict resolution showed greater stability, suggesting greater reliance on human judgment. These effects intensify following GPT-4's release, indicating that capability improvements also drive adaptation. Our findings demonstrate that firms engage in immediate and selective skill deprioritization, raising questions about longer-term hollowing out of human expertise in automated domains. We contribute a novel taxonomy of organizing skills, extend queuing theory to the GenAI context, and provide early empirical evidence on how GenAI is reshaping organizational skill demands.

History: Accepted by Anita McGahan, strategy.

Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2025.01859.

Keywords: artificial intelligence • workforce transformation • human-AI coordination • organizational adaptation

Introduction

Emerging research on artificial intelligence (AI), and more recently, on generative artificial intelligence (GenAI) has largely examined whether AI substitutes or complements humans at the task level, with particular attention to potential productivity gains and automation efficiencies (Choudhury et al. 2020, Yilmaz et al. 2023, Wang et al. 2024, Choi et al. 2025, Demirci et al. 2025). This focus is natural because GenAI is a general-purpose technology that can be deployed on its own or combined with other technologies to solve a wide range of problems (Bick et al. 2026). Yet, the breadth of general-purpose technologies also implies substantial uncertainty about their productivity consequences, in part because general-purpose technologies typically require complementary organizational changes before their benefits are fully realized (David 1990). Just as the steam engine reshaped the geography of production (Rosenberg and Trajtenberg 2004) and the dynamo catalyzed factory redesigns (David 1990), so may GenAI reorganize how firms collect, integrate, and apply information in decision making. A natural path to study such reorganization is through changes in the firm's workforce. We focus on an early and understudied mechanism of organizational adaptation to GenAI: skill deprioritization,[1] which we define as a systematic reduction in firms' demand for human skills that can be effectively addressed by GenAI. Specifically, we examine how GenAI reshapes organizational skill demands across core organizing functions and document emerging patterns of skill deprioritization as firms begin to adapt to the new technology. Because our empirical window is shortly after GenAI became widely accessible, we interpret these patterns as early signals of how firms *expect* GenAI to reshape organizing rather than as direct evidence of realized internal reorganization.

To theorize these patterns, we revisit the foundational organizing problems emphasized 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, whereas integration of effort encompasses coordination, information sharing, motivation, and conflict resolution (Puranam et al. 2014; Puranam 2018; Raveendran et al. 2020, 2022; Tonellato et al. 2024). These organizing functions shape what firms hire for and therefore, provide a concrete theoretical lens for analyzing how firms respond to GenAI technologies when adjusting skill demand.

Central to understanding such effects is the need to elucidate AI's impact on the nature of work, particularly with respect to the GenAI applications of large language models (LLMs), such as OpenAI's GPT, Google's Gemini, and Meta's Llama. These models are trained on massive data sets and use billions of parameters to predict words in response to user prompts. Unlike previous technologies, LLMs do not require the explicit ex ante codification of knowledge; instead, they implicitly learn statistical regularities from diverse inputs, especially textual data. LLM-based GenAI systems can generate, revise, summarize, and translate text; draft workplace communications (e.g., emails, memos, and reports); and support analytical and interpretive knowledge work by rapidly synthesizing information from multiple sources. These capabilities plausibly constitute a general-purpose technology shock to managerial problem-solving and by extension, to how organizations address the problems of division of labor and integration of effort. It is, therefore, unlikely that organizations will respond to such novel computer-assisted capabilities by having the same individuals perform the same tasks in the same way.

We expect that one immediate adjustment is skill deprioritization; as LLMs expand what can be accomplished by fewer human inputs, firms can reduce demand for skills tied to organizing functions that become less labor intensive. The long-run effects on organizations or the job market may be larger than the initial response because general-purpose technologies typically diffuse through a process of experimentation, learning, and complementary reorganization (David 1990). Accordingly, we emphasize that our study examines the *early stages* of GenAI adoption. We quantify and characterize immediate effects on skill demand shortly after ChatGPT's release while recognizing that the full impact may unfold over years as firms learn how to integrate GenAI capabilities into managerial organizing functions. We also do not assume instantaneous recognition or optimization; rather, we interpret our estimates as capturing initial firm reactions.

Empirically documenting these early adjustments is crucial for two reasons. First, it provides insights into how GenAI is beginning to reshape the nature of work and labor market dynamics beyond productivity gains on specific tasks, establishing a baseline for assessing how these impacts evolve as these technologies evolve (Dixon et al. 2021, Yilmaz et al. 2023, Hui et al. 2024, Law and Shen 2025). Second, it can inform firms and policymakers seeking to anticipate and respond to shifts in skill demand induced 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 mapping provides a theoretically grounded and systematic lens for assessing how GenAI relates to core managerial organizing functions. To derive our theoretical predictions, we view organizations as engaging in problem-solving and apply a queuing model comprising 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 et al. (2020); we extend this approach to characterize 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 quasiexperimental design that leverages the public release of ChatGPT as an exogenous availability[2] shock to traditional AI and GenAI. Our sample comprises 1,820 publicly listed U.S. companies, with hiring demand data tracked over a ±12-month window surrounding the launch of ChatGPT. Using a difference-in-differences (DiD) approach, we compare changes in job postings for skills related to organizing functions among firms with higher versus lower exposure to GenAI. To define GenAI exposure, we use the AI exposure measure introduced by Felten et al. (2021), which was subsequently updated to better capture GenAI-related occupational exposure (Felten et al. 2023). We find that following ChatGPT's release, firms with higher GenAI exposure exhibit an immediate statistically significant decline in demand for skills related to task division and information provision. We find a significant decline across all information provision subcategories—coordination, synchronous communication, and asynchronous communication. We also observe reductions in skills associated with monitoring within reward distribution and on the operational side of exception management. In contrast, we find no statistically significant changes in skill demand related to task allocation (staffing and mapping), reward distribution incentives, and exception management conflict resolution. We discuss the theoretical and practical implications of these findings for how GenAI may reshape managerial organizing functions.

Our study makes three main contributions. First, we show that a theoretical model of organizations solving problems optimally provides a useful foundation for predicting early reductions in human skill demand following GenAI availability. Although we do not claim that the firms are behaving fully optimally, our results suggest that the optimal queuing model is a good first approximation because firms adjust their hiring demand in ways that are directionally consistent with its predictions. In particular, the model predicts that with enhanced GenAI capabilities, organizations would anticipate reducing 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 21.3% in coordination skills and 23% in task division-related job postings. The model's prediction that GenAI would most strongly impact structured, rule-based organizational tasks proved remarkably accurate, suggesting that early hiring adjustments 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 (Yilmaz et al. 2023, Demirci et al. 2025), firms appear to adjust their skill portfolios to improve fit with the anticipated effects of GenAI. This reorganization allows firms to restore strategic alignment with the new technological environment rather than simply reacting to demand shifts (Lieberman et al. 2017, Wuebker et al. 2023, Tandon et al. 2024). Our study complements research on GenAI-driven organizational transformation by highlighting skill deprioritization—a mechanism that has received comparatively less attention than skill augmentation or reskilling—and by providing evidence on which organizing functions appear most exposed in the early period. Our findings also suggest that demand for certain organizing skills may become scarcer in the labor market, with implications for how managerial work and organizational structure and processes may evolve in the future.

Finally, we advance existing models of categorizing skills into substantive clusters (e.g., Autor et al. 2003, Deming and Kahn 2018) by proposing a theoretically grounded taxonomy for organizing them. To identify which jobs perform which organizing tasks, we develop a structured methodological framework that combines a hand-curated analysis of meanings with natural language processing scaling capabilities. The resulting framework both supports our empirical tests and offers tools for future research on GenAI and organizing, including how the technology may reshape the division of labor, coordination architecture, and managerial problem-solving.

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 1967, Weber 1978, McEvily et al. 2014, 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, highlighting 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 tasks to the most suitable individuals or units, considering skill requirements and interdependencies. Effective task allocation hinges on individual skills and how tasks fit together to generate synergies that align with both internal capabilities and external demands.

Integration of effort ensures that interdependent agents coordinate and work collectively toward organizational goals. The first component is information provision, which refers to how relevant information flows among interdependent agents. Information provision can occur through synchronous channels (real-time interactions) or asynchronous channels (such as dashboards, documenting processes, and archiving past decisions)—a distinction further elaborated on in the literature (e.g., Rerup 2009, Joseph and Ocasio 2012), highlighting how organizational design can shape managerial attention and feedback mechanisms. The second component is reward distribution, which includes designing incentive systems to motivate desired behaviors and maintain goal alignment, incorporating both tangible rewards (e.g., compensation) and intangible rewards (e.g., social recognition and status). A third component is exception management, which addresses unanticipated situations, coordination failures, or conflicts. Effective exception management depends on robust coordination mechanisms that maintain knowledge flows and shared understanding when routines break down within the organization (Martin and Eisenhardt 2010, Foss et al. 2013).

Human-GenAI Coordination as an Organizing Challenge

The availability of GenAI alters the problems of organizing in two ways. First, it expands the information processing and decision-making capabilities available to individuals. Second, it changes how information is produced and shared, including automation that can substitute for some forms of human-to-human communication. Early AI technologies primarily excelled at large-scale pattern recognition and the automation of routine tasks, often outperforming humans in those areas (Brynjolfsson and Mitchell 2017). At the same time, their limited ability to interpret context, justify decisions, or engage in normative reasoning meant that organizations typically retained human oversight for broader strategic decision making (Lebovitz et al. 2022). Recent advances in large language models blur this division by improving performance on tasks involving language, synthesis, and communication—exhibiting capabilities that increasingly mimic facets of human creativity and communication (Kellogg et al. 2020). This shift invites a re-examination of traditional task division and coordination between human-AI ensembles as organizations reconfigure work processes to exploit these new capabilities (Anthony et al. 2023). It may require a recalibration of internal coordination processes as organizations reallocate responsibilities from roles associated with routine supervision, conflict resolution, or performance assessment.

Evidence from adjacent technologies suggests plausible structural consequences. For example, 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. Although earlier AI systems handled real-time data analysis and automated structured decisions, they largely relied 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, whereas other tasks remain human but are completed more efficiently with GenAI tools. These changes could, in turn, lead organizations to reallocate responsibilities away from organizational roles that are solved more efficiently using GenAI or substituted by GenAI processes.

To understand how such reallocation of responsibility may shape the organizational skill requirements, we apply a queuing model that allows for comparative analysis of organizational capabilities under technological change (Glynn et al. 2020). We analyze the first-order effect of GenAI on the problem-solving capabilities of individuals by examining how this technological advancement immediately affects organizational skill requirements. We also provide some proposals on the second-order effects of GenAI, which could enable the reorganization of information flow and the reassignment of certain problem-solving tasks from humans to AI decision making. However, we note that these second-order effects require experimentation, learning, and implementation, and therefore, they 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 assigning problem-solving individuals, and the model specifies an expected solution time for each problem following an exponential distribution. Importantly, individuals do not necessarily solve problems independently as they may 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 assigned to a qualified problem-solver accordingly. Because these organizing approaches can all be modeled as queues with different structures, this framework allows for the modeling of many organizational forms (Glynn et al. 2020).

The first-order effects of GenAI on organizing functions are relatively easier 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 with increasing 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 plausible. Specifically, we highlight three. 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 as faster problem-solving requires less scheduling whether done individually or in teams. Traditionally, managers spend considerable time breaking down complex tasks, but GenAI may make this process smoother. It can enable automated task decomposition, saving managerial resources. For example, the project management software Jira uses a GenAI assistant to transform high-level project descriptions into detailed issues and subtasks, recommending work items and dependencies.[3] 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.[4] Likewise, 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, thereby reducing the need to break down tasks. As the theory proposes and the examples illustrate, GenAI can reduce the resources needed for task division as an organizing task. Because task division is accompanied by behavior and outcome control to ensure smooth operations, the exception management created by this organizing task should also require fewer resources. This reduction in resource use occurs because GenAI systems can enable organizations to automatically identify and fix operational flaws while facilitating task decomposition in the first place. Stripe, for example, built GPT-4 into its internal tools for fraud detection and developer support, enabling it to process large amounts of documentation and identify out-of-pattern issues. Stripe's GPT-4 analyzes support tickets and code to troubleshoot integration problems, answering developer questions and highlighting likely bugs.[5]

A second specific effect is related to improving information gathering by GenAI-enabled individuals. To the extent that GenAI facilitates the collection and interpretation of information both externally and internally within the organization, this enables 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 task becomes easier to perform. It is also plausible that GenAI enables direct retrieval, integration, and analysis of information, allowing 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 that are 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 such as summarizing policy documents and translating text between languages, which helps diplomats quickly digest information and communicate across language barriers.[6]

A third specific effect concerns 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 and automated metrics collection) have outpaced information-gathering and analysis technology for some time. GenAI now allows organizations to use these richer data in an automated manner, reducing the need for human problem-solvers to interpret employees' performance 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 human resources (HR) analytics platform Culture Amp experimented with ChatGPT to summarize multiple sources of employee feedback. Instead of a manager relying solely on memory, a GenAI agent can retrieve data, such as kudos from Slack, customer comments, and peer feedback, collected over months and use them to generate a coherent performance summary.[7]

In contrast, we have weaker grounds to predict immediate reductions in task allocation as an organizing effort, the design of incentives as a reward distribution effort, or conflict resolution as an exception management effort as individual selection and motivation would presumably be similar when individuals are GenAI enabled, and the potential for conflicts is also likely to remain. Nevertheless, we posit that reductions in these organizing efforts are also plausible 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; however, the longer duration of implementing such changes may make them difficult to detect.

Finally, although our theoretical model focuses on the immediate efficiency gains that GenAI offers, it is important to acknowledge that these first-order adaptations have potential costs. The same processes that broaden individual capabilities and speed up tasks may also introduce unintended negative externalities, such as the erosion of deep expertise or the emergence of new challenges related to algorithmic management. These could be problematic, especially because some of them would unfold over time. Substituting human experts with GenAI could reduce the effectiveness of responses to unusual challenges because human experts can reason from first principles and adapt causal models to unprecedented situations, whereas GenAI is constrained by patterns learned from training data. Reducing the middle management layer also weakens the promotion tournament by leaving fewer candidates competing for promotion and reducing the ability to test their capabilities as their work will be supported and partially substituted by GenAI. Although a full exploration of these issues is beyond the scope of what we directly measure, we revisit these important long-term challenges in the Discussion section.

From Skill Deprioritization to Reorganizing

This section explains why the selective *deprioritization* of task division and information provision skills implies a wider reconfiguration of the managerial function, thus reorganizing the firm. We formalize that a decrease in the demand for task division and information provision skills necessarily expands managerial span of control when the variance of problem-solving times declines.

Research on the span of control—the number of direct reports that a manager supervises—shows that it is jointly determined by (i) the arrival rate of decisions that a manager must endorse and (ii) the variance in the time required to process each decision (Keren and Levhari 1979, 1989; Garicano 2000; Garicano and Rossi-Hansberg 2006; Puranam 2018; Aghion et al. 2021). GenAI applications, such as ChatGPT, reduce both the arrival rate and the processing time for many coordination and monitoring tasks. For example, GenAI enables autogenerated progress updates, prevalidated crossfunctional handoffs, and anomaly flags from dashboards rather than human subordinates. Queuing theory, therefore, predicts that a manager's effective service rate increases as the variance of service times declines. When the coefficient of variation c falls, the expected waiting time in an $M/G/1$ queue declines convexly, so k additional subordinates ("customers" in the original model) can be absorbed without raising the expected delay (Glynn et al. 2020). Put differently, a selective fall in coordination labor mechanically widens the feasible span of control even if the total problem arrival rate is unchanged.

Let problems arrive at a manager m according to a Poisson process with rate λ and have service times that follow an exponential distribution with mean μ^{-1} and variance μ^{-2}. When GenAI reallocates a share $0 < \varphi \leq 1$ of coordination work to an LLM-enabled agent, the residual variance becomes

$$\sigma^2_{post} = (1-\phi)\sigma^2_{pre}.$$

For an $M/G/1$ queue, the expected waiting time is

$$E[W] = \frac{\lambda\sigma^2}{2(1-\rho)},$$

where $\rho = \lambda/\mu$. Holding μ constant and substituting σ^2, the manager can admit additional subordinates Δn such that

$$\frac{\lambda(1-\phi)\sigma^2_{pre}}{2\left[1-\rho+\Delta n.\overline{\lambda}\right]} \leq (\lambda\backslash\sigma^2_{pre})/\left(2\left(1-\rho\right)\right),$$

where $\overline{\lambda}^{-}$ is the average arrival rate generated by one subordinate. Solving yields

$$\Delta n \approx \frac{\phi}{1-\rho}\cdot\ (\sigma^2_{pre})/(\lambda^{-}).$$

This suggests that for any $0 < \varphi \leq 1$ and utilization $\rho < 1$, the feasible span of control increases monotonically in φ; the effect is amplified when pre-GenAI variance σ^2 and baseline utilization ρ are high.

If firms exploit this advantage, we expect to observe (i) a decrease in postings for task division and information provision skills and (ii) a parallel expansion in the required managerial span of control. Our empirics focus on the first set of predictions. Accordingly, our theory predicts that firms will initially reduce human resources allocated 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 detailed in the next section.

Methods

To assess the impact of AI on skill demand, we employ a quasiexperimental 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 a compelling quasiexperimental setting for three reasons. First, the introduction of ChatGPT marked a major advance in GenAI capabilities that was revealed to all market participants roughly at the same time. It also concretely reshaped managerial beliefs about what GenAI could do. Although earlier AI may have seemed abstract or narrow in scope, ChatGPT offered concrete demonstrations of natural language and generative capabilities across a wider range of tasks, including complex cognitive tasks. This broad and visible exposure plausibly prompted decision makers to update their expectations about which organizing tasks could be automated or augmented by GenAI, with implications for workforce planning and skill requirements embedded in hiring demand. Second, the introduction of ChatGPT as a shock is empirically meaningful because of 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 GenAI-relevant tasks. Specifically, we define treated companies as those predominantly operating in task environments where GenAI's natural language processing and generative capabilities have the most significant impact (see the Treated Companies section for operationalization). Third, the timing of ChatGPT's release is also credibly exogenous, limiting concerns about anticipatory adjustment. Managers plausibly lacked advance knowledge of its release and well-formed priors about its scope. 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 approach:

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

where i indicates 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 term. The dependent variable y_{it} is the count of job postings by firm i in month t that require a specific subset of skills. The interaction term $PostGPT_t \times Treated_i$ is our regressor of interest, identifying treated companies after the ChatGPT release by OpenAI. We do not add $PostGPT_t$ and $Treated_i$ first-order terms as separate regressors because these are collinear with time and firm fixed effects. We study a ±12-month window around ChatGPT's public release on November 30, 2022; operationally, December 2022 onward is the postperiod. We chose this ±12-month window because the preperiod (from December 2021 to November 2022) provides sufficient history to establish baseline hiring demand patterns while being recent enough to avoid contamination from earlier AI developments. Furthermore, the postperiod (from December 2022 to November 2023) allows sufficient time for organizations to adjust their hiring demand in response to the shock while being short enough to minimize the risk of confounding events, such as subsequent GenAI releases or broader macroeconomic changes that could potentially affect our results.

We employ a Poisson pseudomaximum likelihood (PPML) estimator with high-dimensional fixed effects and cluster standard errors at the company level. We chose PPML for three key reasons. (i) It accommodates count variables with a large number of zeros without requiring problematic transformations, (ii) it yields consistent estimates even when the underlying data are not truly Poisson distributed, and (iii) it is robust to heteroskedasticity (Santos Silva and Tenreyro 2006, Chen and Roth 2024).

Data and Variables

We analyze a sample of 1,820 publicly listed U.S. companies, tracking hiring demand decisions over ±12 months around the exogenous shock, which resulted in 40,465 firm-month observations. Our dataset integrates three sources: Lightcast, Compustat, and Revelio. Lightcast maps the entire U.S. job market daily, with detailed information at the job-posting level, whereas Compustat and Revelio provide company-level control variables. Online Appendix A 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) and later explicitly updated to account for occupations exposed to GenAI (Felten et al. 2023). This measure has two main components: (1) the mapping between 52 human abilities and more than 800 occupations as listed in the Occupational Information Network (O*NET) database developed by the U.S. Department of Labor and (2) a crowdsourced 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 classify firms into treatment and control groups in four steps. First, we replicated Felten et al. (2023) artificial intelligence 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 indexes occupations and j indexes each of the

52 occupational abilities. The term A_j captures the exposure of ability j to "language modeling" AI application, and L_{jk} and I_{jk} are weights for ability j's importance and prevalence by occupation k, respectively. Second, using Lightcast, we computed monthly counts of job postings by four-digit North American Industry Classification System (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 of an occupation within each industry-month combination. Third, we constructed the monthly four-digit NAICS GenAI exposure:

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

where n indexes four-digit NAICS industries and k refers to an occupation. The term S_{nk} is the prevalence of occupation k in industry n, and $AIOE_k$ weights exposure by occupational composition. Finally, we average monthly GenAI exposure over the pre-ChatGPT period to obtain an industry-level GenAI exposure measure. Firms in industries whose pre-ChatGPT GenAI exposure exceeds the pre-ChatGPT sample mean are classified as treated (we also replicate our results using the sample median) (see Table D3 in Online Appendix D).

Our classification procedure yields 966 treated firms and 854 control firms. This distribution emerged naturally from our data, resulting in a near-balanced split. Online Appendix A reports the top and bottom 20 industries in terms of GenAI exposure in our sample. Figure A1 in Online Appendix A shows that in the 12 months postrelease, treated firms post substantially more jobs requiring GenAI skills than control firms, providing support for the validity of our treatment and control group assignment (see also Figure C1 in Online Appendix C for monthly trends).

Problems of Organizing. Our dependent variables measure monthly firm-level job postings requiring skills related to the problems of organizing (Puranam et al. 2014; Puranam 2018, 2024). We operationalize these using Lightcast's skill-level taxonomy (version 9.25), which includes 33,832 unique skills linked to job postings and accompanied by short descriptions (~70 words) describing its function. For instance, a product manager job in a software company might be associated with product management, engineering design processes, and prioritization. Further, "prioritization" is defined by Lightcast's skill-level taxonomy 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. We restrict the analysis to job postings that include at least one of these management-related skills, excluding postings unlikely to involve organizing functions (e.g., lathe operator roles).

To map the 33,832 skills to these categories, we use a multistep process. First, we conducted topic modeling to explore the semantic landscape and filter for relevant skills. Specifically, we estimated 20 latent Dirichlet allocation models, varying the number of topics from 5 to 100 in increments of 5. We evaluated the models based on a bundle of coherence scores, top keywords, and skill descriptions for each topic. The 65-topic solution provides the best discrimination of themes for our scope, with four topics highly associated with our five organizing function categories. We selected all skills with high posterior probabilities for those topics, yielding a filtered set of 1,765 skills.

Second, we manually read the descriptions of the 1,765 filtered skills, defining 10 labels as summarized in Table 1. We split the Puranam (2018) five high-level categories into more granular subcategories[8] for two reasons. First, theoretically, more fine-grained subcategories allow for sharper predictions about where GenAI should matter (e.g., GenAI affecting operational exceptions versus conflict resolution). Second, empirically, it improves measurement by defining categories that are both minimal (containing the smallest possible set of semantic attributes of the concept) and independent (each label does not rely on any other label to explain the meaning). These properties increase confidence that the subsequent machine learning--based projection will accurately capture the intended meanings.

This manual labeling process identified 397 skills that matched our categories (true positives) and 1,368 skills that did not match our categories (true negatives) of the 1,765 skills filtered using topic modeling. We validated the above classification by having a research assistant produce hand-curated labeling over the same set of 1,765 skills. In doing so, we achieved average precision, recall, and $F1$ score across all labels of 0.83, 0.81, and 0.83, respectively, suggesting a satisfactory level of 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.

Third, we scale our classification of the 10 labels to all 33,832 skills using GPT-4o to determine whether each skill reflects one or more (multilabel classification) of our organizing problem categories. Each skill is labeled by three independent GPT-4o agents and assigned a label when at least two agents agree. Accordingly, we treat GPT-4o agents as filters to determine the final set for manual screening. Next, we

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, and 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, and workforce planning operations
Mapping	Assigning specific human or nonhuman resources (tools and technology) to tasks	Assigning employees, competency mapping, organizational architecture, resource allocation, and delegated authority
Information provision		
Coordination	Information sharing to synchronize efforts among different individuals or teams	Crossfunctional coordination, process integration, collaboration, coordinating, and stakeholder coordination
Synchronous	Information sharing via real-time communication (e.g., meetings and calls)	Verbal communication skills, speech fluency, social communications, professional speaking, and interactive communications
Asynchronous	Information sharing via asynchronous, documented communication, such as emails, reports, and wikis	Report writing, professional writing, structured writing, project documentation, and writing outlines
Reward distribution		
Monitoring	Monitoring or measuring employees' tasks and behaviors to assess employees' performance	Employee monitoring, plan of action and milestones, employee performance management, workforce productivity, and performance review
Incentives	Designing and implementing motivational rewards (monetary or nonmonetary) to encourage desired behavior	Reward management, executive compensation strategy, high-potential programs, benefits strategies, and compensation strategy
Exception management		
Conflict	Identifying and addressing conflict between employees	Employee conflict resolution, conflict resolution, conflict transformation, de-escalation techniques, and organizational conflict
Operational	Identifying, analyzing, and resolving unexpected operational problems	Overcoming obstacles, A3 problem-solving techniques, creative problem-solving, root cause corrective action, and deviation investigations

Note. A3, paper size (11.7 × 16.5 inches).

validated the model's filtering ability against the manually labeled sample from step 2. Here, GPT-4o achieved 93% accuracy, 86% recall, and 89% *F*1 score in identifying true negatives. On the full data set, the model identified 1,369 candidate instances.

Finally, we manually labeled the 1,369 candidate instances retrieved in the previous step. This process identified 657 true positives mapping onto 719 skill-label relationships.

We use these skill-problem pairings to count the number of job postings related to each of the 10 organizing function categories for our dependent variables. Through this step, we processed 11,154,915 job postings. The skill categories are not mutually exclusive, so a job post may refer to multiple skill categories (multilabel classification).

Online Appendix B presents the topical map derived from the 65-topic model. It also includes the prompt used to configure GPT-4o agents, a bar chart showing skill frequencies by label, and the full list of 657 skills by problem of organizing. Online Appendix B4 provides further justification for using GPT-4o filtering rather than traditional keyword-based approaches.

Control Variables. Our baseline DiD includes the following controls: size, return on assets (ROA), leverage, the ratio of research and development over total assets (R&D/TA), a research and development (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 is computed using quarterly financial data from Compustat. Size is measured as log (1 + *total assets*). ROA is quarterly net income divided by total assets. Leverage is the ratio of total liabilities (sum of short-term and long-term debt) over total assets. R&D/TA is computed as quarterly R&D expenses divided by total assets (missing R&D values are set to zero). We include an R&D dummy equal to one for firms reporting positive R&D expenditures and zero otherwise. Finally, Capex/TA is quarterly capital expenditures over total assets.

We also control for workforce change, which is measured as the month-over-month percentage difference in head count, which can be negative (contraction) or positive (expansion). For example, firms undergoing substantial layoffs may exhibit temporarily depressed demand for hiring or shifts in skill demand because of restructuring rather than GenAI. We used Revelio to track changes in each company's head count by month:

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

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

Figure A2 in Online Appendix A compares 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 reports descriptive statistics for key variables across 40,465 firm-month observations. The average firm size is 7.87 (measured as the log of total assets), profitability is close to zero (ROA), leverage is 0.30, and the capital expenditures ratio is 2% relative to total assets. Forty-two percent of firms report positive R&D expenditures, with an average R&D intensity of 2%. Fifty-three percent of firms in the sample are classified as highly exposed to GenAI, with an average monthly GenAI exposure score of 0.35.

Table 2. Summary Statistics

Variable	*N*	Mean	SD	p10	Median	p90
EM (Conflict)	40,465	13.67	71.77	0.00	1.00	23.00
EM (Operational)	40,465	63.59	246.10	0.00	5.00	128.00
IP (Asynchronous)	40,465	60.06	236.15	0.00	5.00	113.00
IP (Coordination)	40,465	52.85	232.13	0.00	4.00	96.00
IP (Synchronous)	40,465	75.51	327.83	0.00	6.00	139.00
RD (Incentives)	40,465	1.44	7.92	0.00	0.00	2.00
RD (Monitoring)	40,465	24.62	107.34	0.00	1.00	45.00
TA (Mapping)	40,465	12.98	50.18	0.00	1.00	26.00
TA (Staffing)	40,465	5.14	44.72	0.00	0.00	7.00
Task Division	40,465	89.27	332.98	0.00	8.00	180.00
GenAI Exposure (dummy)	40,465	0.53	0.50	0.00	1.00	1.00
GenAI Exposure (monthly)	40,465	0.35	0.31	−0.08	0.36	0.77
Size	40,465	7.87	2.15	5.15	7.87	10.66
ROA	40,465	−0.01	0.08	−0.07	0.01	0.04
Leverage	40,465	0.30	0.27	0.03	0.26	0.59
R&D/TA	40,465	0.02	0.04	0.00	0.00	0.05
R&D (dummy)	40,465	0.42	0.49	0.00	0.00	1.00
Capex/TA	40,465	0.02	0.03	0.00	0.01	0.05
WFC	40,465	0.00	0.04	−0.01	0.00	0.02

Note. EM, exception management; IP, information provision; p, percentile; RD, reward distribution; SD, standard deviation; TA, task allocation; WFC, workforce change.

On average, firms list task division skills in 89 postings per month. Demand for information provision is also substantial; synchronous communication skills appear in an average of 76 postings, asynchronous communication skills appear in an average of 60 postings, and coordination tasks appear in an average of 53 postings. Exception management skills also show strong demand, particularly for operational-related skills, and they appear on average in 64 postings per firm month. Reward distribution and task allocation skills are referenced in relatively fewer postings. Because job postings can contain multiple skill categories, these counts are not mutually exclusive.

Table D1 in Online Appendix D reports correlations among the 10 variables representing the demand for the organizing functions ranging from 0.21 to 0.91, with all significant at the 5% level. Figure C1 in Online Appendix C provides additional descriptive insights on the demand for skills relevant to GenAI. The trends indicate that during the study time period, demand for GenAI skills rose sharply driven primarily by firms with high GenAI exposure.

Results

Baseline Results

Panel A of Table 3 reports the baseline DiD estimates of GenAI exposure on hiring demand for skills associated with the organizing functions. We omit control variables in these specifications to avoid the potential issue of "bad controls."[9] All estimates include firm and year-month fixed effects with standard errors clustered at the firm level. We find a statistically significant decline in the demand for task division skills within job postings ($\beta = -0.261$, $p < 0.001$). This corresponds to a 23% reduction, suggesting that following exposure to GenAI, treated firms reduce their demand for skills related to decomposing, structuring, and sequencing work before execution. Similarly, we also find a substantial reduction in demand for information provision skills. Demand for coordination skills declined by 21.3% ($\beta = -0.240$, $p < 0.01$), demand for synchronous communication skills declined by 18.5% ($\beta = -0.204$, $p < 0.05$), and demand for asynchronous communication skills declined by 21.4% ($\beta = -0.241$, $p < 0.01$). For reward distribution, the estimates indicate a significant decrease in monitoring skills in job postings, with a 23.2% reduction ($\beta = -0.264$, $p < 0.001$). There is no statistically significant change in incentive provision skills. Within exception management, the

Table 3. Demand for Skills After the GenAI Availability Shock

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task	Task allocation		Information provision			Reward distribution		Exception management	
Variable	Division	(Staffing)	(Mapping)	(Coord)	(Sync)	(Async)	(Monitor)	(Incentive)	(Conflict)	(Operational)
					Panel A					
PostGPT × *Treated*	−0.261***	0.162	−0.097	−0.240**	−0.204*	−0.241**	−0.264***	−0.150	−0.224	−0.211**
	(0.067)	(0.315)	(0.082)	(0.084)	(0.087)	(0.081)	(0.078)	(0.105)	(0.131)	(0.074)
Constant	6.274***	3.976***	4.392***	5.925***	6.216***	5.913***	5.174***	2.387***	4.745***	5.924***
	(0.012)	(0.063)	(0.018)	(0.016)	(0.017)	(0.015)	(0.014)	(0.021)	(0.024)	(0.015)
N	38,551	30,302	33,997	37,368	38,336	38,204	35,373	26,344	33,501	37,905
Pseudo-R^2	0.929	0.818	0.865	0.919	0.921	0.919	0.902	0.685	0.882	0.918
					Panel B					
PostGPT × *Treated*	−0.260***	0.181	−0.087	−0.237**	−0.204*	−0.243**	−0.259**	−0.151	−0.218	−0.209**
	(0.068)	(0.308)	(0.081)	(0.082)	(0.086)	(0.082)	(0.079)	(0.103)	(0.130)	(0.075)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.780	5.121*	4.166*	4.953	3.753	3.772	3.286	−0.897	0.404	4.854*
	(2.014)	(2.357)	(1.896)	(2.536)	(2.574)	(2.242)	(2.291)	(3.395)	(3.103)	(2.401)
N	38,551	30,302	33,997	37,368	38,336	38,204	35,373	26,344	33,501	37,905
Pseudo-R^2	0.929	0.819	0.865	0.920	0.921	0.919	0.902	0.685	0.882	0.918

Notes. Hiring demand data were tracked over a ±12-month window from the ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. Refer to Online Appendix D for the full tables with controls. Async, asynchronous; Coord, coordination; Sync, synchronous.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

demand for operational exception management skills declined by 19% ($\beta = -0.211$, $p < 0.01$), whereas conflict management shows no significant difference following the introduction of ChatGPT. Finally, neither staffing nor mapping dimensions of task allocation were affected by the GenAI availability shock.

Panel B of Table 3 includes control variables in the baseline specification. The results remain consistent with the previous analysis (panel A of Table 3). Demand for task division declines by 22.9%. Similarly, job postings related to information provision skills show significant reductions, with coordination skills decreasing by 21.1%, synchronous information-sharing skills decreasing by 18.5%, and asynchronous information-sharing skills decreasing by 21.6%. The monitoring skills component of reward distribution experiences a decline of 22.8%, whereas demand for operational exception management skills decreases by 18.9%. Table D3 in Online Appendix D replicates the analyses using a median-based (rather than mean-based) dichotomization of GenAI exposure with consistent results.

Figure 1 plots coefficients with 95% confidence intervals from a dynamic DiD specification. The specification is consistent with our DiD approach, incorporating the same control variables, fixed effects, and clustered standard errors, but this model includes month-year dummies rather than relying on a single post-ChatGPT indicator using November 2022[10] as the reference period (i.e., $t - 1$). Thus, we estimated the following:

$$y_{jit} = \alpha + \sum_{z \neq -1} \beta_z 1[t_{it} = z] \; + \; \theta X_{it} + \delta_i + \delta_t + \epsilon_{jit},$$

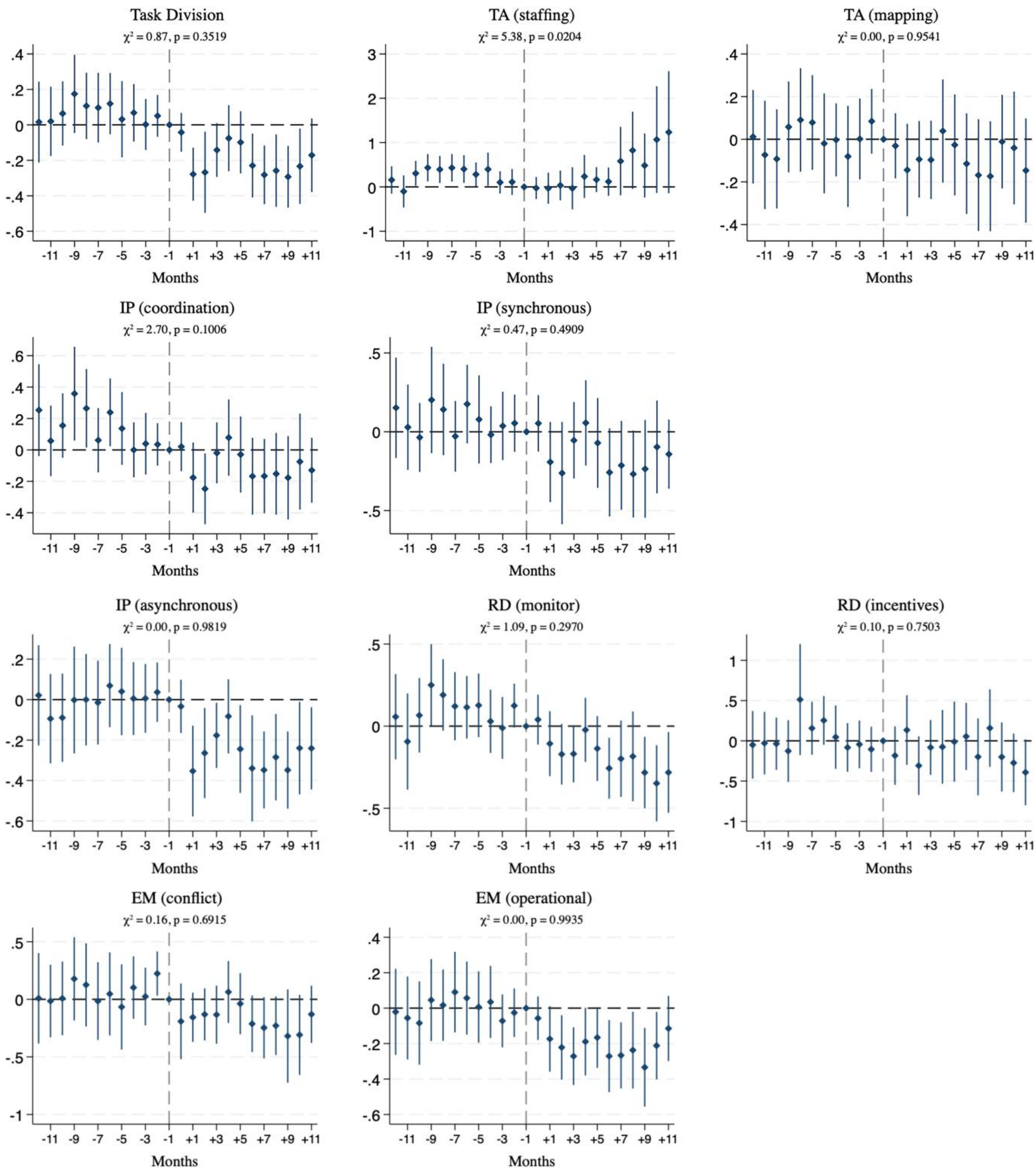
where z represents the months relative to the year-month when ChatGPT is introduced. The plotted estimates support the validity of the parallel-trends assumption for most of our dependent variables in the pretreatment period. The only exception is task allocation (staffing), where the pre-ChatGPT distribution (from December 2021 to November 2022) reveals a 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 pretreatment coefficients. We fail to reject the null hypothesis for all outcomes of interest, indicating no statistically significant differences in pretreatment trends except for task allocation (staffing), which is significant at 5%.

Continuous Industry-Level GenAI Exposure

We next relax the binary treatment classification by interacting the post indicator with the continuous industry-level GenAI exposure measured as of November 2022. Panel A of Table 4 reports these estimates. The findings mirror those reported in Table 3 except for information provision (synchronous) skills, where the effect differs.

The results show that firms with higher GenAI exposure experience a statistically significant reduction in

Figure 1. (Color online) Dynamic Effects



Notes. These graphs show the effect of ChatGPT on problems of organizing. The vertical axes show Poisson coefficients (and 95% confidence intervals). The horizontal axes show time relative to treatment. The graph also contains the result of the χ^2 test for the joint significance of the pre-treatment coefficients. This test is used to formally assess the parallel-trends assumption, which is crucial for a valid difference-in-differences analysis. A nonsignificant *p*-value suggests that the assumption holds.

Table 4. Demand for Skills with Continuous Industry-Level GenAI Exposure (Panel A) Post-GPT-3.5 and Post-GPT-4 (Panel B)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Task	Task allocation		Information provision			Reward distribution		Exception management	
Variable	Division	(Staffing)	(Mapping)	(Coord)	(Sync)	(Async)	(Monitor)	(Incentive)	(Conflict)	(Operational)
					Panel A					
PostGPT × GenAI Exposure	−0.349**	0.502	−0.118	−0.311*	−0.241	−0.343**	−0.321*	−0.186	−0.321	−0.293*
(November 2022)	(0.107)	(0.520)	(0.124)	(0.132)	(0.132)	(0.125)	(0.126)	(0.148)	(0.169)	(0.118)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.001	5.303*	4.177*	5.145*	3.927	4.043	3.410	−0.725	0.503	5.004*
	(1.927)	(2.292)	(1.859)	(2.425)	(2.473)	(2.144)	(2.218)	(3.437)	(3.077)	(2.328)
N	38,551	30,302	33,997	37,368	38,336	38,204	35,373	26,344	33,501	37,905
Pseudo-R^2	0.929	0.820	0.865	0.919	0.921	0.919	0.902	0.685	0.882	0.918
					Panel B					
GPT-3.5 × Treated	−0.262**	−0.265	−0.096	−0.272*	−0.203	−0.220*	−0.164	−0.145	−0.213	−0.150
	(0.092)	(0.164)	(0.094)	(0.106)	(0.113)	(0.108)	(0.090)	(0.133)	(0.135)	(0.100)
GPT-4 × Treated	−0.260***	0.299	−0.083	−0.225**	−0.204*	−0.250**	−0.291***	−0.153	−0.219	−0.230**
	(0.069)	(0.362)	(0.086)	(0.084)	(0.086)	(0.079)	(0.082)	(0.113)	(0.134)	(0.073)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.781	5.344*	4.168*	4.961*	3.752	3.766	3.236	−0.901	0.404	4.830*
	(2.012)	(2.314)	(1.896)	(2.531)	(2.574)	(2.242)	(2.300)	(3.409)	(3.103)	(2.408)
N	38,551	30,302	33,997	37,368	38,336	38,204	35,373	26,344	33,501	37,905
Pseudo-R^2	0.929	0.820	0.865	0.920	0.921	0.919	0.902	0.685	0.882	0.918

Notes. Hiring demand data were tracked over a ±12-month window from the ChatGPT release. All specifications include firm and year-month fixed effects. Standard errors are robust to heteroskedasticity and clustered at the firm level. Refer to Online Appendix D for the full tables with controls. Async, asynchronous; Coord, coordination; Sync, synchronous.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

demand for task division skills in job postings, with a coefficient estimate of −0.349 ($p < 0.01$). Similarly, there is a substantial and statistically significant decline in demand for skills associated with information provision (coordination; $\beta = -0.311$, $p < 0.05$) and information provision (asynchronous; $\beta = -0.343$, $p < 0.01$). The demand for information provision (synchronous) skills remains statistically unchanged. For reward distribution, we find a significant decline in demand for monitoring skills ($\beta = -0.321$, $p < 0.05$), whereas the demand for skills focused on the design of incentives shows no statistically significant variation. For exception management, demand for operational skills decreases significantly ($\beta = -0.293$, $p < 0.05$), whereas conflict-resolution skills remain unaffected. Finally, we find no significant changes in demand for task allocation skills whether related to staffing or mapping. Table D5 in Online Appendix D replicates these analyses using a monthly continuous GenAI exposure measure with consistent results.

Distinguishing the Effect of GPT-3.5 vs. GPT-4

Our study rests on the premise that the introduction of ChatGPT may have helped managers form expectations regarding the scope and applicability of AI technologies. Accordingly, we posit that managers update their hiring demand 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 as a legacy model for over two years. Consistent with our theory, subsequent capability improvements to GPT-4 should strengthen the effects.

To test this, we split our treatment into GPT-3.5 (November 30, 2022) and GPT-4 (March 14, 2023), with a GPT-3.5 dummy that takes the value of one between December 2022 and February 2023 and zero otherwise and a GPT-4 dummy that takes the value of one from March 2023 onward and zero otherwise. Panel B of Table 4 replicates the DiD specification in panel B of Table 3 while separately estimating the effects of GPT-3.5 and GPT-4. Tables D7 and D8 in Online Appendix D further replicate this analysis using, respectively, the continuous GenAI exposure measure as of November 2022 and the time-varying monthly exposure measure. Results in panel B of Table 4 show 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 an increase in the significance level for task division and information provision (coordination). Tables D7 and D8 in Online Appendix D further provide support for our claim, with most coefficients exhibiting increases in magnitude, suggesting that the release of GPT-4 intensified the observed effects of GenAI exposure on firms' hiring demand. This increase in the effect when GPT-4 became available is consistent with a change in expectations modifying hiring demand.

Robustness

We conduct 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 pseudointervention dates set at 6, 9, and 12 months prior to the actual launch of ChatGPT. For each pseudointervention, we constructed samples spanning ±12 months around the assigned date and replicated our DiD analysis. The resulting estimates are reported in Tables E1–E3 in Online Appendix E. We find no significant effects for the pseudointerventions at 9 and 12 months. For the 6-month pseudointervention, the only significant coefficient is smaller in magnitude and exhibits lower statistical significance. These estimates provide evidence that our main results are unlikely to be driven by random fluctuations, which reinforces our causal interpretation.

Second, we perform two sensitivity analyses. First, we follow Rambachan and Roth (2023) to assess the robustness of our results to violations in the parallel-trends assumption. Figure F1 in Online Appendix F reports the results of this analysis at the monthly level, showing how our estimates change when we allow for bounded deviations from strict parallel trends. Second, Tables F1 and F2 in Online Appendix F replicate our main analysis using the extended two-way fixed effects (ETWFE) estimator proposed by Wooldridge (2023). In both cases, task division, information provision (asynchronous), reward distribution (monitoring), and exception management (operational) remain as the most robust outcomes, consistent with our main analysis.

Third, we replicate our analysis using a keyword-based approach to provide additional reassurance that our main results are not an artifact of the filtering approach via GPT-4o agents. To do so, we began by compiling a list of seed terms from existing definitions of the constructs of interest (e.g., Puranam et al. 2014; Puranam 2018, 2024). Using GPT-4o agents, we then expanded each set to 20 terms by providing the agents with both the seed words and our skill taxonomy (see Table B2 in Online Appendix B for the list of terms used). Adopting a dictionary-based approach, we operationalize the dependent variables as the total number of job postings for firm *f* in year-month *t* that contain at least one of the identified keywords. Tables G1–G4 in Online Appendix G present the regression results. Results remain consistent.

Fourth, 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 two-digit NAICS codes 51 (Information) and 54 (Professional, Scientific, and Technical Services), and we re-estimated our models. Tables H1 and H2 in Online Appendix H replicate the main analyses. 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 H1 in Online Appendix H, is not statistically significant, but it is statistically significant in Table H2 in Online Appendix H.

Fifth, we used an alternative method for classifying treated and control firms. We identified treated industries based on the share of job postings in nonroutine cognitive occupations following the classification approach of Jaimovich and Siu (2012). Industries where more than 50% of job postings in the year preceding the shock were for nonroutine cognitive roles were assigned to the treatment group. This method resulted in 1,052 companies classified as treated and 768 companies classified as controls (see Online Appendix I for additional details on the classification process). We then replicated our main analyses in Tables I1 and I2 in Online Appendix I. The findings remain largely consistent with our main results apart from exception management (conflict), 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 the GenAI exposure measure proposed by Felten et al. (2023). Although classification based on nonroutine cognitive occupations captures a relevant dimension of GenAI 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 nonroutine 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.

Sixth, we replicated our analyses using coarsened exact matching (CEM). 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 397 treated and 397 control units, totaling 794 companies. The CEM-based estimates are reported in Tables J1 and J2 in Online Appendix J. The results remain consistent with our baseline analysis, with two exceptions in Table J1 in Online Appendix J: information provision (synchronous) and reward distribution (monitoring).

Seventh, we tested the following additional specifications as reported in Online Appendix K. (1) We replicated the estimates by taking a seemingly unrelated regression approach to account for potential correlation among the unobserved components across dependent variables. Through this specification, we also tested alternative clustering for standard errors at the firm and skill levels and at the firm-skill level. The results in Table K1 in Online Appendix K largely confirm our main results. (2) Tables K2 and K3 in Online Appendix K 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 subdimensions. (3) We replicate our analysis using a simple ordinary least squares (OLS) specification with the canonical log transformation, $\ln(y + 1)$, of the dependent variables (Tables K4 and K5 in Online Appendix K). Results are broadly consistent with our main findings but should be interpreted with caution. Even after applying the log transformation, the distribution of the dependent variables remains more consistent with a Poisson process, suggesting that the PPML estimator remains the more appropriate choice for this analysis. (4) In Tables K6 and K7 in Online Appendix K, we replicated our main findings using the traditional DiD specification described in Angrist and Pischke (2009) estimated with both Poisson and negative binomial estimators (see Online Appendix K). Results remain consistent. (5) In Tables K8 and K9 in Online Appendix K, we relax the assumption of additive and separable effects of the launch of GPT-3.5 and GPT-4 running separate models for each shock. Estimates remain consistent.

Supplementary Analysis

To provide additional evidence on the skill deprioritization patterns, we conducted two extensions. First, Online Appendix L1 examines changes in skill demand relative to total demand. Following recent methodological recommendations, we estimate models of skill demand controlling for total demand rather than relying on ratio specifications that are less suited for statistical inference (Certo et al. 2020, Wulff et al. 2023). Tables L1 and L2 in Online Appendix L show that even after accounting for firm-level variation in total demand, the negative effects for coordination and task division skills remain consistent. Second, in Online Appendix L2, we analyze managerial hiring demand relative to nonmanagerial hiring demand as a proxy for managerial intensity and span of control. Tables L3 and L4 in Online Appendix L indicate that firms exposed to GenAI post fewer managerial openings (i.e., reduce managerial hiring demand) relative to nonmanagerial openings, consistent with an expansion in managerial span of control. Together with evidence of rising demand for GenAI-related skills (Online Appendix C), these findings suggest a selective deprioritization of coordination-intensive skills, with broader implications for firms' organizational structures.

Furthermore, one might question whether LLM-based technologies, such as ChatGPT, can plausibly shape firms' hiring priors in ways that lead to skill deprioritization. We address this with two falsification tests. First, we examine the launch of GitHub Copilot (June 21, 2022) using a DiD specification (Online Appendix L3). Given Copilot's narrower scope of applicability and technical user base, we would not expect comparable effects, and indeed, Tables L5 and L6 in Online Appendix L confirm this expectation. Second, we replicate our analysis using the broader AI exposure measure from Felten et al. (2021) (Online Appendix L4). Although we observe consistent effects on task division, reward distribution (monitoring), and exception management (operational), Tables L7 and L8 in Online Appendix L show that key domains where LLMs matter most, such as writing-related tasks, are not captured by this broader specification. At the same time, it shows a distinctive effect on task allocation (mapping), which opens up interesting future research on the non-LLM-based GenAI role in assigning individuals and resources to tasks. On the one hand, the latter reinforces the value of using the Felten et al. (2023) exposure to fully capture LLMs' impact on firms' priors. On the other hand, it also offers some initial evidence on how our findings may generalize to other types of GenAI applications. Overall, these falsification tests corroborate our interpretation that LLM-based technologies had an influence on firms' hiring priors, leading to skill deprioritization.

Discussion

This study examines how GenAI affects the set of skills that organizations seek in external hiring to address the general problems of organizing. Based on our theoretical framework, we derive three primary predictions. First, GenAI enables individuals to solve problems more quickly and draw on broader knowledge, reducing the resources required for traditional task division and scheduling. Second, information gathering and interpretation become more efficient, supporting a shift from organizational-push to individual-pull mechanisms of information access. Finally, monitoring and performance evaluation become increasingly automated, reducing the human resources traditionally allocated to these oversight functions. These predictions follow from conceptualizing organizations as dynamic queuing systems with increasingly capable problem-solvers enabled by GenAI (Glynn et al. 2020). We do not assume perfect optimization, but we recognize that organizations can incrementally adapt their work processes in response to technological capabilities, approximating more efficient organizational designs. We also do not assume immediate recognition of the potential GenAI contribution to productivity; instead, we focus on initial organizational responses to its availability, which we interpret as revealing how decision makers currently assess its scope.

GenAI Efficiencies

Our results show that the largest declines in skill demand appear in monitoring (reward distribution) and task division tasks. The deprioritization of monitoring skills likely stems from GenAI's suitability for standardized evaluation tasks: processing structured performance data, consistently applying evaluation criteria, and generating feedback. These capabilities plausibly substitute for managerial monitoring. 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 GenAI substitution and potentially higher returns from automation than other organizing functions. Additionally, monitoring tasks, at least from an organization's perspective, often involve structured and predictable decisions based on predefined criteria, making them more amenable to algorithmic approaches than tasks that require nuanced contextual judgment or social calibration.

Although it is plausible that the increase in automatic surveillance and assessment of work performance reduces the human role in monitoring for distributing rewards, it is also possible that GenAI may be used 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 decline in demand for coordination skills supports this interpretation, suggesting that GenAI may be changing how organizations manage task interdependencies. Interdependence of tasks arises from division of tasks into specialized subtasks and assigning expert problem-solvers to each. Yet, GenAI-enabled organizations may better integrate subtasks because GenAI enhances problem-solvers' capabilities, thereby reducing interdependence and consequently, the need to share information among problem-solvers. Reduced demand for coordination and monitoring skills thus aligns with a perspective in which GenAI enables reorganization in ways that economize on task interdependence.

Interestingly, although monitoring showed a strong decline, demand for incentives skills remained largely unchanged. This asymmetry suggests that organizations may be willing to automate performance tracking and assessments while continuing to rely on human judgment for designing and implementing reward systems, which still require contextual understanding and legitimacy that current GenAI systems lack. More broadly, the significant impact on task division and asynchronous information provision alongside insignificant effects on task allocation suggests a selective set of changes in organizational skill demands. GenAI currently appears to impact formal and documentation-heavy processes rather than real-time interactions and staffing decisions where the judgment is more interpersonal and context sensitive.

Potential GenAI Negative Externalities

Although our findings highlight the immediate efficiency gains that drive skill deprioritization, it is important to consider the potential negative externalities and longer-term challenges that may arise from these organizational adaptations. Because we do not directly measure such outcomes, we treat the following as plausible longer-run considerations rather than tested claims. First, the shift away from specialized roles toward broader, GenAI-assisted responsibilities creates a risk of knowledge hollowing and weaker error detection. As firms deprioritize skills in areas such as task division and monitoring, they may unintentionally erode the deep expertise or tacit knowledge that comes from specialization. When human problem-solvers assume broader roles, they may become less likely to spot subtle errors or find novel solutions to the problems that they face. Overreliance on GenAI for task decomposition or analysis could reduce an organization's capability to perform these functions independently, creating a strategic vulnerability.

Second, greater reliance on GenAI for information processing may erode critical skills and increase conflict. The shift to automated information summary and quick digestion of complex data necessarily removes nuances and context-specific information. This may have negative consequences for decision making and may paradoxically increase conflict because of misunderstandings (Koçak et al. 2023, Acemoglu and Wolitzky 2024). Our finding that demand for conflict-resolution skills did not decline even as demand for other management skills did may be an early indicator of this underlying tension.

Relatedly, LLM-enabled asynchronous information provision may lead to information overload while creating dependencies on tools whose limitations are not fully understood. The configuration of these systems (e.g., what data they access and how they present information) can subtly shape organizational priorities, potentially marginalizing certain viewpoints in ways that affect deliberation on complex tasks. Further, LLM-mediated monitoring may induce gaming behaviors as employees optimize for measurable metrics, potentially eroding trust, autonomy, and intrinsic motivation. Greater reliance on GenAI for operational exceptions may reduce opportunities for learning by doing and weaken employees' problem-solving capabilities. When task divisions become less explicit (as GenAI makes handoffs more fluid), organizational processes may become more fragile, breaking down when the technology fails or when problems fall outside its scope. These second-order risks qualify the near-term efficiency gains that we document.

Finally, the deprioritization of human monitoring skills signals a clear move toward new forms of algorithmic management, which introduces its own set of challenges. As GenAI systems take on monitoring and assessment, the design of what is measured and how it is measured becomes critically important, potentially shifting workers' focus from substantive problem-solving to blind execution of quantifiable metrics. Research has already highlighted employee resistance to the control and surveillance inherent in algorithmic management systems (Kellogg et al. 2020). Although we do not observe an immediate spike in conflict-resolution hiring demand, these pressures could accumulate as AI integration deepens, representing a significant challenge for future organizational design.

Contributions

Our findings contribute in three ways. First, they show that organizations engage in selective skill deprioritization rather than uniform capability enhancement. This extends prior 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 in which GenAI provides the strongest complementarities or substitution effects. The significant reductions 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 also aligns with recent work on how organizations adapt differentially to GenAI capabilities (Shrestha et al. 2019, Waardenburg et al. 2021).

Second, we advance queuing theory applications in organizational contexts (Cohen et al. 1972, Glynn et al. 2020) by demonstrating how GenAI-enhanced problem-solving capabilities differently shape organizing functions. The substantial decline in demand for 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 decline in demand for monitoring skills aligns with predictions that GenAI enables more efficient quality control processes with fewer human resources, linking to emerging work on algorithmic management (Kellogg et al. 2020, Bailey et al. 2022).

Lastly, our study contributes to the emerging literature on human-AI collaboration (Kellogg et al. 2024, Choudhary et al. 2025, Dell'Acqua et al. 2025). Although existing research offers microlevel insights into how human and AI agents may collaborate by complementing or substituting each other, it says less about the meso-level and macro-level consequences of novel divisions of labor. By examining meso-level organizational adaptations to human-AI ensembles, our study provides insights into the changing nature of work and how organizations are responding with new hiring demand practices, with implications that also extend to the broader labor market.

Mechanisms and Development

Our analysis of organizational responses to GPT-3.5 versus GPT-4 provides additional insights into the mechanisms driving skill deprioritization. The larger effects following GPT-4's introduction period, particularly for skills demand related to information provision (asynchronous), reward distribution (monitoring), and exception management (operational), suggest that capability improvement rather than mere awareness drives organizational adaptation. This natural experiment helps distinguish between two potential mechanisms: (1) changes in managerial beliefs about GenAI's potential, which would be primarily triggered by the initial ChatGPT-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 GenAI demonstrates more advanced capabilities, indicating that skill deprioritization is driven by substantive capability enhancement rather than by hype or novelty effects.

At the same time, it also suggests that the skill deprioritization that we currently observe may be the first step in a process of skill reprioritization. Two processes are occurring in parallel, and both involve gains in GenAI skills among employees. Lower-level workers may become less dependent on middle managers as they draw on GenAI to obtain the needed information for independent decision making, thereby developing skills in GenAI-augmented information and decision making. Fewer middle managers overseeing an organization of the same scale handle fewer coordination and exception management decisions per employee, but they likely handle the same number or even more in total. This also suggests that GenAI-augmented information collection and decision making will likely become a valued capability at multiple hierarchical levels, especially because middle managers also sometimes need to check whether the independent decisions taken at the lower levels are really optimal for the firm.

Overall, our study documents the immediate effects of firms' responses to the inflow of GenAI capabilities through the selective deskilling of their labor force. This is a crucial first step, and a key next step in future research is to identify the processes that underlie each of these effects as well as the second-order effects, including how GenAI tools fundamentally shape reorganization when nonhuman problem-solvers handle greater parts of their operations. Although GenAI is poised to enhance productivity and create new roles, it will necessitate a significant adaptation of the workforce and likely raise strategic considerations about skills relevant to the organization.

In the medium term, the restructuring of organizing functions driven by GenAI may intensify intraorganizational conflict, even if our data currently show no significant short-run effect on conflict-resolution roles. This aligns with the Kellogg et al. (2020) argument regarding employee resistance to algorithmic management. Pressures may accumulate as GenAI-based surveillance becomes more salient and as reduced interdependence shifts accountability toward individuals—conditions that employees are likely to resist. Organizations could respond by temporarily increasing demand for conflict-resolution skills, restructuring internal workflows, or adopting other strategies to proactively reduce human-AI frictions; this deserves further investigation.

In the longer term, GenAI may enable the emergence of novel organizational forms that blend human and GenAI 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 because of coordination costs. Increasing spans of control could reduce managerial intensity and layers, reallocating humans toward strategic and political tasks while operational decisions become more automated. This raises a strategic question about 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 response could be for organizations to develop new approaches to knowledge management that explicitly account for human-AI collaborative learning, in which AI enhances rather than substitutes for organizational learning processes. It could also involve deliberately maintaining human expertise in strategically critical domains while delegating other domains coupled with processes designed to enable human-AI learning rather than substitution.

Limitations

Although 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 other national contexts. Publicly listed companies in the United States are typically large, with more formalized hiring processes and better access to cutting-edge technologies compared with smaller or non-U.S. firms. Extending the analysis to different types of organizations and geographical and regulatory regimes is an important next step to assess the broader applicability of our results.

Second, our skills classification method combines topic modeling and manual labeling. Although 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. Future work should test whether alternative approaches to skill classification produce similar 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 GenAI capabilities, estimated effects could be attenuated. This means that our approach may be conservative as the effects would likely be even stronger without spillovers.

Fourth, as we noted, our short-run estimates reflect initial inferences about how GenAI capabilities may impact organizing behavior. Because the inferences are based on descriptions of what GenAI can do and initial trials, they are most likely incomplete. Changes in the nature and extent of organizational changes are likely to occur over time as GenAI capabilities become better understood. The reduction in demand for certain types of skills and organizational roles is a comparatively direct response; deeper reorganization, including the integration of more autonomous GenAI decision making and automated information distribution, may unfold over time. We, therefore, interpret initial responses as a lower bound on longer-run structural changes.

Finally, the theoretical framework that we employ along with the recent evidence from different economic settings suggests that the skill deprioritization that we identify is part of a broader trend. The AI exposure measure from Felten et al. (2021), which underpins our identification strategy in the Supplementary Analysis section (Online Appendix L4), is itself built on a framework assessing occupational exposure not just to language models but to a variety of AI applications, including image generation and code generation. This provides a theoretical basis for expecting similar transformative effects across different domains of white-collar work. Indeed, recent empirical work confirms this crossdomain impact. A study by Hui et al. (2024) on a large online freelance marketplace found that the introduction of both text-based (ChatGPT) and image-based (e.g., DALL-E 2) GenAI models led to significant reductions in employment and earnings for freelancers in affected occupations. This finding of adverse effects in a flexible gig economy, which mirrors the decline in hiring demand that we observe in established firms, suggests that the phenomenon is robust across different organizational forms. Notably, Hui et al. (2024) also find suggestive evidence that high-quality, top-rated freelancers were disproportionately affected, which resonates with our concept of skill deprioritization among knowledge workers.

Conclusion

Our study opens several exciting avenues for future research. Our findings raise important questions about the broader structural consequences of GenAI adoption. For instance, work on flat organizations (Lee and Edmondson 2017, Foss and Klein 2022) and multiple organizational goals (Gaba and Greve 2019) may offer alternative explanations for why certain skill domains, such as task allocation or conflict management, remain stable even as others decline. More broadly, the organizational forms enabled by GenAI and the constraints that limit their emergence should become a central topic in management research. We hope that this study encourages future work to further explore how GenAI is reshaping the organizational architecture of firms.

Acknowledgments

The authors thank the department editor, the associate editor, and the anonymous reviewers for their constructive guidance during the review process. The authors also thank Piyush Gulati, Milan Miric, Phanish Puranam, seminar participants at INSEAD and London Business School, and participants at the 2025 Strategic Management Society Annual Conference for their valuable comments and suggestions. All errors are the authors' own.

Endnotes

[1] By deprioritization, we mean that the skill is less used by firms, not that it is seen as normatively less valuable. Deprioritization of skills in the workplace often occurs because workplace use of skills and societal valuations of skills are frequently decoupled. One of the authors learned to operate a lathe in middle school, indicating some societal valuation of this skill, but has never used one since.

[2] ChatGPT represents an availability shock to GenAI and AI capabilities by reducing barriers across multiple dimensions. It lowers knowledge barriers by enabling nontechnical audiences to leverage complex GenAI and AI capabilities, it lowers prototyping costs by fostering rapid experimentation, and it may accelerate skill acquisition and shorten the learning curve—all of the above, even from a smartphone.

[3] See https://www.atlassian.com/platform/artificial-intelligence.

[4] See https://thejiraguy.com/2023/12/11/atlassian-intelligence-is-here/.

[5] See https://www.techcircle.in/2023/03/16/11-companies-using-gpt-4-in-consumer-products.

[6] See https://fedscoop.com/state-department-encouraging-workers-to-use-chatgpt.

[7] See https://www.shrm.org/topics-tools/news/technology/how-hr-using-generative-ai-performance-management.

[8] To illustrate how we arrived at the 10 subcategories, consider the "information provision" category. Puranam (2018, p. 11) defines this as ensuring that agents have information "to execute their own actions and coordinate actions with others." This definition alone (which contains "and") suggests a split. Our qualitative reading of thousands of job postings then revealed a third dimension related to the mode of communication—whether it was synchronous (e.g., presenting) or asynchronous (e.g., report writing). For this reason, we split "information provision" into three distinct subcategories: coordination, synchronous communication, and asynchronous communication. We applied a similar logic, which is grounded in theory and our qualitative data analysis, to the other core categories.

[9] A *bad control* is a control variable that is itself affected by the treatment. In DiD designs, controlling for variables that may be endogenous to the treatment can lead to biased estimates (Angrist and Pischke 2009).

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

References

Acemoglu D, Wolitzky A (2024) Employment and community: Socioeconomic cooperation and its breakdown. NBER Working Paper No. 32773, National Bureau of Economic Research, Cambridge, MA.

Adan I, Resing J (2015) Queueing systems. Working paper, Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, Netherlands.

Aghion P, Bloom N, Lucking B, Sadun R, Van Reenen J (2021) Turbulence, firm decentralization, and growth in bad times. *Amer. Econom. J. Appl. Econom.* 13(1):133–169.

Angrist JD, Pischke JS (2009) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

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.

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.

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

Bick A, Blandin A, Deming DJ (2026) The rapid adoption of generative AI. *Management Sci.*, ePub ahead of print January 20, https://doi.org/10.1287/mnsc.2025.02523.

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).

Burton RM, Obel B (2004) What is an organizational design? Burton RM, Obel B, eds. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit*, Information and Organization Design Series, vol. 4 (Springer, Boston), 43–85.

Certo ST, Busenbark JR, Kalm M, LePine JA (2020) Divided we fall: How ratios undermine research in strategic management. *Organ. Res. Methods* 23(2):211–237.

Chen J, Roth J (2024) Logs with zeros? Some problems and solutions. *Quart. J. Econom.* 139(2):891–936.

Choi S, Kang H, Kim N, Kim J (2025) How does artificial intelligence improve human decision-making? Evidence from the AI-powered Go program. *Strategic Management J.* 46(6):1523–1554.

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.

David PA (1990) The dynamo and the computer: An historical perspective on the modern productivity paradox. *Amer. Econom. Rev.* 80(2):355–361.

Dell'Acqua F, Kogut B, Perkowski P (2025) Super Mario meets AI: Experimental effects of automation and skills on team performance and coordination. *Rev. Econom. Statist.* 107(4):951–966.

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.* 70(10):8097–8108.

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 EW, 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. Preprint, submitted April 19, http://dx.doi.org/10.2139/ssrn.4414065.

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

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.

Gaba V, Greve HR (2019) Safe or profitable? The pursuit of conflicting goals. *Organ. Sci*. 30(4):647–667.

Garicano L (2000) Hierarchies and the organization of knowledge in production. *J. Political Econom.* 108(5):874–904.

Garicano L, Rossi-Hansberg E (2006) Organization and inequality in a knowledge economy. *Quart. J. Econom.* 121(4):1383–1435.

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 No. 18334, National Bureau of Economic Research, Cambridge, MA.

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, Valentine MA, Christin A (2020) Algorithms at work: The new contested terrain of control. *Acad. Management Ann.* 14(1):366–410.

Kellogg KC, Lifshitz-Assaf H, Randazzo S, Mollick ER, Dell'Acqua F, McFowland E III, 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, Harvard Business School, Boston.

Keren M, Levhari D (1979) The optimum span of control in a pure hierarchy. *Management Sci.* 25(11):1162–1172.

Keren M, Levhari D (1989) Decentralization, aggregation, control loss and costs in a hierarchical model of the firm. *J. Econom. Behav. Organ.* 11(2):213–236.

Koçak Ö, Puranam P, Yegin A (2023) Decoding cultural conflicts. *Frontiers Psych.* 14:1166023.

Law KK, Shen M (2025) How does artificial intelligence shape audit firms? *Management Sci.* 71(5):3641–3666.

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.

Rosenberg N, Trajtenberg M (2004) A general purpose technology at work: The Corliss steam engine in the late 19th century US. *J. Econom. Hist.* 64(1):61–99.

Santos Silva JMC, Tenreyro S (2006) The log of gravity. *Rev. Econom. Statist.* 88(4):641–658.

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.

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* (University of California Press, Berkeley).

Wooldridge JM (2023) Simple approaches to nonlinear difference-in-differences with panel data. *Econometrics J.* 26(3):C31–C66.

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

Wulff JN, Sajons GB, Pogrebna G, Lonati S, Bastardoz N, Banks GC, Antonakis J (2023) Common methodological mistakes. *Leadership Quart.* 34(1):101677.

Yilmaz ED, Naumovska I, Aggarwal VA (2023) AI-driven labor substitution: Evidence from Google Translate and ChatGPT. Preprint, submitted April 17, http://dx.doi.org/10.2139/ssrn.4400516.

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.