Automation or Augmentation?
Task Characteristics, Managers' Objectives, and AI Adoption in Science

Maximilian Koehler
ESMT Berlin

Henry Sauermann
ESMT Berlin

EARLY DRAFT – PLEASE DO NOT DISTRIBUTE

ABSTRACT

Artificial intelligence (AI) is diffusing rapidly throughout the economy, raising questions regarding its impact on human workers. Current theories predict that the degree to which humans are replaced (automation) or supported (augmentation) will depend on characteristics of the task such as its degree of routinization and manual vs. cognitive nature. Yet, existing empirical evidence tends to come from aggregate occupation data or individual case studies, with limited research using larger samples of projects or tasks. Moreover, whether and how AI is implemented may not only depend on task characteristics but also on managers' objectives, which may go beyond narrow efficiency to include benefits resulting from human employment itself ("employment goals"). We provide novel empirical evidence using data from more than 1,200 research projects that involve crowd members in different tasks such as data collection, data analysis, or creative problem solving. We confirm that the use of AI is associated with task characteristics, although the patterns are more nuanced than those shown in work on prior automation technologies such as computers and robots. Moreover, we find that managers pursuing employment goals are less likely to use AI for automation, while they are just as likely as others to use AI to augment the work of human workers. We contribute to the literature on AI adoption in organizations by providing rare evidence from project-level data, and by highlighting that the path towards "good AI" or "bad AI" is, partly, a matter of human choice.

We thank Scistarter.org for providing us with access to project-level data and ESMT Berlin for financial support. We are grateful for helpful feedback from colleagues and seminar participants. All errors are our own.

1 Introduction

Artificial intelligence applications are diffusing rapidly in many sectors, including manufacturing, services, education, research, and government (Agrawal et al., 2018; Bianchini et al., 2022; Margetts & Dorobantu, 2019; Mollick & Mollick, 2023; Raisch & Krakowski, 2021). Echoing concerns about the impact of prior waves of new technologies such as robots and computers, this has triggered discussions about whether AI will diminish the role of humans by automating tasks, or whether AI will augment humans by supporting them in their work and making them more effective. Although initial writings often claimed one side of the debate or the other, more nuanced discussions recognize that the effects of AI will differ across jobs and occupations. The underlying premise is that the strengths and weaknesses of AI – and those of humans – depend on the nature of the task (Brynjolfsson & Mitchell, 2017; Felten et al., 2021). Building on prior work on automation, scholars have focused on the routinization and the manual vs. cognitive nature of tasks, but have also explored other aspects such as the role of tacit knowledge and the relevance of social skills or creativity (Autor, 2015; Felten et al., 2021; Grennan & Michaely, 2020). Given that the capabilities and constraints of AI differ from those of prior technologies, however, results from prior work on automation may not generalize to AI. The current empirical evidence on the role of AI in particular remains limited (Aghion et al., 2018; Brynjolfsson & Mitchell, 2017; Tschang & Almirall, 2021).

Perhaps more importantly, discussions focusing on task characteristics and technical capabilities of AI ignore the fact that adoption decisions in organizations are – at least for now – made by humans such as CEOs, managers or project organizers. These decision makers may consider not only potential cost savings or quality improvements from employing AI but may also pursue goals that require the continued employment of humans ("employment goals"). Among others, this may relate to political pressures to preserve employment (Gu et al., 2020) and to managers' individual prosocial preferences that make them reluctant to fire workers (Guenzel et al., 2023; Keum & Meier, 2023). Although such employment goals have been shown to matter in other contexts, there is little evidence on their role in shaping the adoption of AI.

Our paper makes two key contributions. First, we provide novel empirical insights into the relationships between task characteristics and the automation vs. augmentation of human workers. One stream of prior research has examined such relationships using large scale macro data on employment in different occupations or by skill levels to study temporal shifts that may reflect the influence of AI – but also many other factors (Aghion et al., 2018; Frank et al., 2019).

Another stream provides rich insights from specific tasks such as playing chess, examining patents, responding to IT requests, or interpreting medical images (Allen & Choudhury, 2022; Krakowski et al., 2023; Lebovitz et al., 2022). Each of these studies helps understand the nuanced interplay between humans and machines for a particular task but does not speak to systematic patterns across tasks. We provide empirical insights from a context that offers unique insights into the internal operations of a wide range of projects, allowing us to measure specific task characteristics as well as whether AI is used to automate, to augment, or not used at all. Our data come from more than 1,200 crowd science projects that each involve humans in a narrow range of tasks but collectively span a broad range of research tasks that differ with respect to their degree of routinization as well as cognitive vs. physical nature (Barbosu & Gans, 2022; Franzoni & Sauermann, 2014; Haklay et al., 2021). As predicted, we find that automation is highest for projects with highly routinized tasks, while augmentation is more common in projects with intermediate degrees of routinization. Similarly, automation increases with the degree to which tasks are cognitive (vs. physical), whereas augmentation tends to be more common when tasks require both cognitive and physical activities.

Second, we provide suggestive evidence that whether and how AI is adopted also depends on the goals of managers. In our particular context, project organizers may not only pursue productivity (project's ability to generate research outputs more rapidly, at lower costs, and higher quality) but also goals that require the active involvement of human workers ("employment goals"). This involvement may be required, for example, to enable participants' learning about science, allow the general public to shape the direction of research, and reduce public skepticism against science (Sauermann et al., 2020; Trouille et al., 2019). Our results show that organizers who pursue employment goals are generally less likely to use AI, which primarily reflects lower levels of automation, not lower levels of augmentation. Although the particular reasons for why active involvement of humans is seen as important may be unique to our context, "employment goals" that go beyond productivity are clearly important in the broader economy (Gu et al., 2020; Guenzel et al., 2023; Keum & Meier, 2023). As such, our results provide initial evidence that the diffusion of AI does not have to be "automatic" but is also shaped by the goals of organizations and human decision makers. In other words, whether society moves towards "bad AI" that automates away human work or "good AI" that provides opportunities for humans to do more meaningful work is at least partly a matter of human choice (Acemoglu, 2023).

In the following section 2, we develop hypotheses regarding the role of the two task characteristics that are most salient in the prior literature: the degree of routinization and the cognitive vs. physical nature of tasks (Autor et al., 2003; Felten et al., 2021). We also draw on prior literature on the role of managers' prosocial motives to theorize about the role of managers' goals in shaping decisions about whether and how to adopt AI. In section 3, we provide background on our empirical context and discuss measures. Section 4 discusses the results, including hypotheses tests and supplementary analyses. Section 5 concludes by summarizing results, acknowledging limitations, and discussing our contributions to the literature.

2 Conceptual background

2.1 AI, augmentation, and automation

Artificial intelligence or "AI" refers to machines that simulate or mimic cognitive human functions, typically based on algorithmic predictions (Agrawal et al., 2018; Brynjolfsson et al., 2018; Raisch & Krakowski, 2021). AI can be used for descriptive purposes, such as finding hidden patterns in data, but also to forecast future outcomes or generate artefacts such as text or images (e.g., DALL-E or ChatGPT). While traditional rule-based algorithms are static and difficult to scale to more complex problems, AI algorithms can learn from data, adapt to new information, and also handle more complex problems. In this paper, we focus on AI rather than rule-based algorithms but we abstract from particular technical specifications or implementations (e.g., supervised vs. unsupervised vs. reinforcement learning).

When implemented in organizations, AI can have different impacts on humans and their work. Following the prior literature, we broadly distinguish "automation" and "augmentation" (Choudhary et al., 2023; Krakowski et al., 2023; Lebovitz et al., 2022). *Automation* is said to occur when AI takes over complete tasks from humans. Once sufficient training data is available, for example, AI can automatically classify images of galaxies in astrophysics research, taking over a task that used to consume considerable time on the part of PhD students and Postdocs (Lintott, 2019). Note that automation of a particular task does not mean that human workers have to lose their jobs because jobs typically involve bundles of tasks. As such, workers may shift their attention to other tasks within a given job, and job profiles may also expand to include new tasks (Acemoglu & Restrepo, 2018; Brynjolfsson et al., 2018).

Augmentation occurs when AI and humans collaborate on a particular task. This often means that AI systems support workers by preparing decision input or making recommendations, but workers stay in the loop and make critical decisions. In radiology, for example, AI systems may recommend diagnoses to physicians who then make final judgments (Lebovitz et al., 2022).

Although augmentation may reduce the human time required to perform a particular task, AI does not take over the whole task but rather complements the human worker in performing that task.

The boundary between automation and augmentation can obviously be fuzzy and the role of humans and AI in performing a particular task may change over time (Raisch & Krakowski, 2021). Perhaps more importantly, drawing this distinction requires us to focus on a task at a particular level of aggregation because shifting to a more disaggregated task definition implies a shift from augmentation (e.g., human and AI collaborate on task T) to automation (e.g., AI automates subtask T1 and the human is left with subtask T2). Following the standard task model, the following conceptual discussion simply assumes a well-defined task at a particular level of aggregation (Felten et al., 2021; Restrepo, 2023). We will discuss relevant tasks in our empirical contexts more concretely in section 3.

2.2 The role of task characteristics

A growing literature discusses the antecedents and consequences of automation, primarily through "older" technologies such as robots and computers but now also artificial intelligence. This work focuses on how task characteristics shape the ability of technologies to perform these tasks and technologies' resulting ability to replace or support human workers.[1] In this paper, we focus on two task characteristics that are particularly prominent in the prior literature and that are also important in our empirical context: Task routinization and the physical vs. cognitive nature of tasks.

2.2.1 Routinization

Routinized tasks involve the execution of repetitive operations in a largely stable environment, where the same stimulus requires the same response over time. Such tasks can typically be codified (Autor, 2015; Autor et al., 2003; Dixon et al., 2021; Pentland & Rueter, 1994).[2] Older automation technologies such as robots and computers are more likely to be used

[1] Although our discussion will focus on capabilities of AI versus humans and resulting differences in task performance, managers will also consider relative costs (Arntz et al., 2016; Autor et al., 2003). To simplify the discussion, we will assume that if a particular task characteristic makes the task easier for AI to perform, then this can translate into higher performance, lower costs to develop the AI, or both, either way increasing a relative advantage over humans.

[2] The construct of task routinization discussed in the prior literature entails several different aspects such as repetitiveness and codifiability. Even though there may be advantages to disentangling these aspects and developing distinct predictions, we follow the prior literature by using routinization as a more aggregated construct. This enables us to link to the prior literature but also simplifies our conceptual discussion and empirical analysis.

for routinized than non-routinized tasks because these technologies require explicit instructions and face difficulties adjusting to changing environments. Authors writing about AI, however, suggest that codification may be less important for intelligent algorithms who may be able to replicate human performance without receiving explicit instructions (Aghion et al., 2018; Brynjolfsson et al., 2018). For example, an AI can identify patterns in data on past human decisions and can then make similar decisions without explicit instructions on how to evaluate decision options. AI – especially when embedded in broader systems that collect real time environmental information – may also be better able to respond to changing environments. Even though AI may now be able to also perform less routinized tasks, however, we expect that its relative advantages over humans are greater in routinized tasks for the reasons expressed in the older literature on automation, but also because the repetitive nature of routinized tasks is likely to yield greater volumes of data and opportunities for AI to learn and improve its performance. As such, we predict the following pattern across projects involving different types of tasks:

Hypothesis 1a: Projects involving high task routinization are more likely to use automation than projects involving low or medium task routinization.

We suggest that task routinization also improves opportunities to use AI to *augment* human workers. Consider again the example of an AI that provides decision support – this AI likely makes better recommendations for tasks that are routinized and for which larger volumes of data are available (e.g., image classification by radiologists). However, the degree of routinization required for augmentation is likely lower than that required for automation given that humans remain in the loop and can compensate for any deficiencies of AI (e.g., the radiologist making the final diagnosis). Assuming that automation promises greater efficiency than augmentation due to lower labor costs or higher speed, efficiency-oriented organizations should automate when possible and should augment only if tasks are not sufficiently routinized to be automated but sufficiently routinized for AI to support humans (Huang & Rust, 2018).[3]

Hypothesis 1b: Projects involving medium task routinization are more likely to use augmentation than projects involving low or high task routinization.

[3] Although we focus on cross-sectional implications, Huang and Rust (2018) highlight dynamic implications: AI adoption should "progress" over time from augmentation to automation as AI capabilities develop or as tasks become more amenable to the use of AI (e.g., due to increasing routinization and digitalization).

2.2.2 Physical vs. cognitive nature of tasks

Physical tasks involve physical movements and the use of physical skills, e.g., moving boxes from one room to another or lifting a patient. Cognitive tasks require mental effort and information processing; examples include pattern recognition, problem-solving, and decision-making.

While some tasks are primarily cognitive or physical, others entail important physical and cognitive elements. For example, the task to collect data on animals in nature requires someone to move to different geographical locations to encounter animals (primarily physical) but also to identify the species of the animals that are observed (primarily cognitive). We will call tasks that have important physical and cognitive elements "hybrid" tasks.[4]

Software-based AI is unable to perform activities that require mobility or physical manipulation (Tschang & Almirall, 2021), suggesting that AI should not be able to automate physical tasks. In contrast, AI is strong in processing information. As such, AI should be more likely to be used for automation in cognitive than in physical tasks (Autor, 2015; Felten et al., 2021). Although AI alone cannot perform hybrid tasks due to its inability to perform physical aspects, it may support humans in performing some of the cognitive aspects (augmentation). Of course, AI could also be used to augment rather than replace humans in cognitive tasks, especially if such tasks involve cognitive activities that current AI still struggles with (Kittur et al., 2019; Ponti et al., 2022). Given the increasing information processing capabilities of AI, however, augmentation is likely more prevalent in hybrid tasks, where humans remain needed for physical aspects.

Hypothesis 2a: Projects involving cognitive tasks are more likely to use automation than projects involving physical or hybrid tasks.

Hypothesis 2b: Projects involving hybrid tasks are more likely to use augmentation than projects involving physical or cognitive tasks.

---

[4] Of course, virtually all tasks involve both physical and cognitive elements. Thinking of a continuum, we define hybrid tasks as those where both elements are similarly important, while physical tasks are those where the physical element dominates and cognitive tasks are those where the cognitive element dominates. As the above example illustrates, the classification of tasks requires a fixed level of analysis (see also Autor & Handel, 2013; Haeussler & Sauermann, 2020): If we disaggregate the hybrid task of "data collection in nature" into "walking to find an animal" and "identifying the species of an animal", then the former task would by physical while the latter would be cognitive.

2.3 The role of managers' goals

Our discussion in section 2.2 followed the implicit assumption in much of the literature that adoption is driven by efficiency considerations: New technologies are used to automate or augment human workers when it is economically efficient to do so (Aghion et al., 2018; Bessen, 2016; Brynjolfsson & Mitchell, 2017; Furman & Teodoridis, 2020). Although this logic should be useful to understand overall diffusion patterns across organizations, whether and how AI is adopted in a particular organization is typically up to human decision makers such as managers or project organizers. These decision makers, in turn, may pursue goals that go beyond economic efficiency, introducing heterogeneity in decisions whether and how to adopt AI that is not explained by task characteristics alone. In particular, managers may not only pursue greater productivity and efficiency but also goals that relate directly to the employment of human workers. [5] For example, research has shown that managers take into account the pressure from unions and policy makers, leading them to employ more employees than technically necessary (Gu et al., 2020). Similarly, managers with prosocial motives are more reluctant to fire workers, especially if organizations have slack resources or if corporate oversight is weak (Keum & Meier, 2023). Most relevant for our study, recent work shows that managers with prosocial motives invest less in automation, arguably in order to preserve their employees' jobs (Keum, 2023).[6]

In our empirical context, the active involvement of human crowd members allows organizers to accomplish several goals that go beyond the production of scientific output. For example, some organizers involve crowds not only to produce research but also to enable participants to learn about science, to allow citizens to shape the direction of research, to strengthen awareness of environmental and social challenges, or to counter rising skepticism against science (Haklay

---

[5] There is a more general literature on technology adoption that highlights efficiency considerations as well as various other social, organizational, and behavioral influences (e.g., Frambach & Schillewaert, 2002; Greenwood et al., 2019; Plouffe et al., 2001). There is potential for greater integration of this literature in discussions of AI, but as far as we can tell, even that literature pays little attention to goals associated with human employment per se. As highlighted by agency theory, managers may also have other relevant goals, such as income maximization at the expense of both the organization and workers. We abstract from such other goals.

[6] Managers in extremely competitive industries may not have discretion to pursue goals other than efficiency. In many industries, however, firms do have slack resources or are shielded from competition, allowing for at least some degree of managerial choice (Gu et al., 2020; Keum, 2023). More generally, heterogeneity in management practices is pervasive in many industries and countries (Bloom et al., 2012) and firms differ in how they resolve trade-offs between competing goals such as efficiency, social responsibility, and sustainability (Ethiraj & Levinthal, 2009; Obloj & Sengul, 2020; Stevens et al., 2015).

et al., 2021; Sauermann et al., 2020; Van Brussel & Huyse, 2018).[7] We call organizers that pursue goals that require the active involvement of humans as pursuing "employment goals".

There is no inherent trade-off between the production of research and employment goals; indeed, a larger workforce typically increases both scientific output and learning opportunities for workers (Sauermann et al., 2020; Shibayama et al., 2015). However, trade-offs arise when a technology such as AI arrives that can automate the work of humans. We argue that organizers pursuing only efficiency goals should automate where possible to increase the speed and quality of production while reducing labor costs.[8] In contrast, organizers that also pursue employment goals should automate less to preserve the benefits of employment.[9] However, such organizers could still use AI to augment human workers because augmentation keeps humans in the loop. Indeed, using AI to augment human workers may even enable projects to better achieve employment goals if AI support lowers participation barriers for workers (e.g., by reducing skill requirements) (Furman & Teodoridis, 2020).

Hypothesis 3a: Organizers with employment goals are less likely to use automation than organizers without employment goals.

Hypothesis 3b: Organizers with employment goals are more likely to use augmentation than organizers without employment goals.

3 Empirical context, data, and measures

3.1 Empirical context

Our empirical context is crowd science (CS): Research projects that recruit participants through an open call (Franzen et al., 2021; Franzoni & Sauermann, 2014; Haklay et al., 2021). CS projects are typically initiated and led by professional scientists, including at top tier institutions such as Cornell, Oxford, and Stanford University. As such, crowd science is not an alternative to "traditional" science but rather a novel approach that enables scientists to increase

---

[7] Such goals are not unique to the context of crowd science. For example, the education of future generations of scientists through active participation in research is a major function of academic research labs (Shibayama et al., 2015; Stephan, 2012).

[8] In our empirical context, direct labor costs are zero because participants are not compensated financially. However, involving participants is still costly in terms of recruiting and managing and many projects face labor constraints in terms of how many people are willing to participate (Sauermann & Franzoni, 2015). As such, automation can increase the speed of production while decreasing costs related to participant management (Trouille et al., 2019).

[9] For simplicity, we assume that all organizers pursue productivity goals and we focus on heterogeneity with respect to employment goals.

the productivity and impact of research.[10] Projects involve crowd members primarily in data collection and processing but also a range of other tasks such as identifying novel research questions and solving complex problems. Participants include other scientists but also non-professional "citizens" who contribute time and effort as well as knowledge they may have gained through experience with particular problems (e.g., as patients).

Crowd science projects are a suitable research context for several reasons. First, individual CS projects involve crowd members for narrowly defined and quite specific tasks rather than broader bundles of tasks (as do typical jobs), allowing us to more clearly characterize task characteristics and explore their relationship with AI. Second, projects vary considerably with respect to the nature of tasks as well as organizer objectives, allowing us to exploit cross-project variation to test our hypotheses. Finally, crowd science projects tend to be very transparent regarding their internal organization, allowing insights that would be difficult to gain in traditional academic or commercial projects. Of course, this novel context requires us to consider generalizability (see section 5).

3.2 Data collection

The website SciStarter.org is the largest public catalogue of crowd science projects, serving as a platform for projects that seek contributors and individuals who are willing to participate in research projects. We obtained from SciStarter leadership a list with all 1,528 projects currently running on the platform (as of Spring 2023).[11] A co-author and a research assistant ("coders") then collected detailed data on each of these projects using different sources such as the project profile on SciStarter.org, projects' main websites, blogs and reports, as well as research articles published by the projects. In some cases, the coders also signed up for projects to perform tasks as participants. Both coders investigated the full set of projects, combining information and agreeing on final assessments via email as well as in scheduled Zoom sessions. They performed the information search and coding using detailed coding guidelines that were developed and pre-tested by the authors. This process yielded measures for the main variables, i.e., use of AI, task characteristics, as well as employment goals. The coders also coded several

[10] In some fields, CS projects are now at the core of research activity. For example, eBird has contributed roughly half of the biodiversity data in the Global Biodiversity Information Facility (https://ebird.org/news/2021-year-in-review). Zooniverse projects have yielded many publications in fields such as space, physics, medicine, and social sciences (https://www.zooniverse.org/about/publications).

[11] Although SciStarter.org is the most comprehensive sample frame available, it has better coverage in the U.S. than in other regions. As such, the projects we analyze may not necessarily be representative of the global population of CS projects.

control variables. Some additional control variables come from system information graciously provided by SciStarter leadership.

Of the 1,528 projects, 230 turned out to be inactive at the time of the data collection and were dropped from the sample. We dropped an additional 31 projects that were not relevant (e.g., they were not primarily research-oriented). Our final sample consists of 1,267 projects.

3.3 Measures

3.3.1 Dependent variables

The coders searched all available sources for evidence that a project uses AI either to automate or to augment human participants performing project tasks.

Automation. Coders assessed whether a project uses AI as a "worker" who performs a particular task without the direct involvement of a human. In the project Galaxy Zoo, for example, an AI automatically classifies images of galaxies based on training data generated by project participants. In the project Grammar Maven, linguists use both AI and human contributors to independently identify grammatical regularities in written language.[12] By definition, all projects in our sample involve human crowd members in some capacity. As such, automation typically involves that AI takes over tasks from humans once sufficient training data has been generated or that AI and humans work in parallel on the same task without direct collaboration. In principle, the task could be performed by humans alone, although this would typically take longer time and require a larger number of crowd members (Trouille et al., 2019). In our sample, 18% of projects use AI for automation (dummy coded =1).

Augmentation. Coders assessed whether a project uses AI to support crowd members in performing a particular task. In the project eBird, for example, crowd contributors have access to the AI tool Merlin, which helps them to classify birds and record observations made in nature.[13] Although AI could not perform tasks such as recording bird observations made in nature autonomously, it can increase the effectiveness and efficiency of participants who perform this task. In our sample, 7% of projects use AI for augmentation (dummy coded = 1). This includes a small share of projects (2.6% of total) that use AI for both augmentation and automation.[14]

---

[12] https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/; https://scistarter.org/grammar-maven

[13] https://scistarter.org/ebird

[14] Our unit of analysis is a project. The majority of projects involve crowd members in only one particular task, creating a direct correspondence between tasks and projects. If a project involves multiple tasks, it may be coded as using both automation and augmentation, and our coding of task characteristics (see next section) reflects an assessment of the average across tasks.

3.3.2 Independent variables

Routinization. The coders rated task routinization on a 3-point scale (1= low, 2= medium, 3= high). In doing so, they considered the frequency and type of actions carried out by participants and the context in which these actions occur (Autor et al., 2003). Routinization is coded as high for repetitive tasks performed in a constant environment, such as classifying images online. Routinization is coded as low for tasks that are not repetitive and/or that have to be performed in very different conditions, such as collecting complex data at different archaeological sites. Routinization is coded as medium for tasks that are somewhat repetitive and may be performed under similar or only slightly different conditions. Examples include collecting repeated bird observations in a particular geographical area, or using online tools to generate RNA structures with different chemical properties in the project EteRNA.[15]

Physical vs. cognitive vs. hybrid. We proxy for the nature of tasks by coding the degree to which they are performed offline versus online. Tasks are classified as physical when they are performed primarily offline, without the use of computers or smartphones. Examples include the collection of historical artefacts or animals in physical space (e.g., mosquitoes), which are then sent to principal investigators via postal mail for identification and further processing. Tasks are assumed to be primarily cognitive if they are performed online, e.g., the classification of digital images of galaxies or the design of new RNA structures using computer-based interfaces. a typical example is the task to collect biodiversity data in nature and submit identified records to projects using online interfaces. Although there will be some cognitive aspects to offline tasks (e.g., deciding whether to catch a particular mosquito) and physical elements to online tasks (e.g., typing on a keyboard), offline vs. online performance should be a useful proxy for the relevance of physical vs. cognitive aspects.

Employment goals. To assess whether project organizers pursue non-productivity goals that require human engagement ("employment goals"), the coders reviewed sources such as goal statements on project websites, blogs, as well as reports and publications. In our particular context, a common employment goal is that projects seek to educate participants about science through active engagement in research activities. For example, the Butler County Stream Team project involves crowd members in collecting water samples, testing samples, and identifying benthic macroinvertebrates. The project also offers educational events, courses and workshops for participants to learn about water quality and relevant measures.[16] Other projects seek to

---

[15] https://scistarter.org/eterna

[16] https://scistarter.org/butler-county-stream-team

sensitize participants to environmental problems and to encourage their appreciation of nature through activities such as measuring air quality or collecting data on biodiversity. Some projects also explicitly seek to "democratize" science by enabling participants to participate in decision making and to take ownership of results. When assessing the presence of such non-productivity employment goals, the coders considered explicit goal statements but also indirect evidence such as published reports about knowledge increases among participants or acknowledgements of participants in scientific publications.

3.3.3 Control variables

Primary research field. There may be important differences across fields in the costs and benefits of implementing AI, but also in the goals of research projects and in the ways in which organizers work with crowd members. To control for such differences, we code projects' primary research field based on project profiles on SciStarter as well as project webpages.

Task scope. The benefits as well as costs of adoption AI may depend on the scope of projects in terms of the number of stages of the research process (and, thus, tasks) that crowd members participate in. In a first step, we coded whether projects cover each of eight possible tasks (Franzoni et al., 2022): Defining the research question, acquiring funding, developing materials and methods, collecting data, processing or analyzing data, solving problems, writing the paper, diffusing results. Given that most projects cover only a single task, we then created a dummy variable that distinguishes projects covering one task from those covering more than one task.

Project size. Project size may be an important control because larger projects may have more resources to invest in AI technology and also generate more data that AI can learn from. We collected information on project size in terms of the number of crowd members involved, typically from information provided on project webpages or blogs. To simplify the analysis, we code three categorical variables indicating that projects are small (<1,000 participants), medium (1,000-9,999), or large (10,000+). We code a separate dummy indicator for projects where this variable is missing (based on our review of project pages, these appear to be mostly very small projects). We recognize that the number of participants is potentially endogenous in that projects with greater use of automation may involve fewer people. Although other size measures (such as funding amounts) or not available, we do run robustness checks that exclude the size measure.

Geographic scope. Projects differ in the geographic scope of their work, which typically has implications for the location of crowd members. Some projects, for example, seek to study phenomena such as air pollution or biodiversity change in a particular city or region. Others address very general research questions, collecting data and involving crowd members at national or global level. Although we see geographic scope primarily as an alternative proxy for project size, it may also capture the diversity of data (e.g., the global project eBird deals with a much broader range of species than a local bird observation project in the state of Florida), which may have implications for the benefits and costs of using AI. Information on project scope is taken from systems data provided by SciStarter; we code a dummy variable that takes on the value of 0 for local/regional projects, and 1 for national/international projects. We create a separate dummy for projects where this variable is missing.

Project age. The use of AI may depend on project age, e.g., because older projects have accumulated more data that AI can learn from, but also because younger projects may find it easier to integrate new technologies into their workflows. Age data is taken from systems data provided by SciStarter (years since the first listing of a project).[17]

Platform affiliation. Many crowd science projects are independent, others reside on larger platforms that provide access to an established user base or technical infrastructure such as user management systems. Most relevant in our context, platforms may also provide access to shared AI tools, reducing projects' costs of implementing AI (Koehler & Sauermann, 2023). The largest platforms include Biocollect, iNaturalist, and Zooniverse.[18] We coded platform membership based on information provided on project websites and by searching for traces such as shared domain names or member logins.[19]

4 Results

4.1 Descriptives

Table 1 shows summary statistics for key variables; Table 2 shows correlations. Figure 1 visualizes the means of automation and augmentation for each category of our independent variables. Panel 1 shows that automation is highest for highly routinized tasks, consistent with Hypothesis 1a. Contrary to our expectations, however, it is higher for tasks with low

[17] SciStarter launched a new website and migrated data to a new system in 2018. As such, the maximum observed age is 6 years.

[18] Given that projects typically have a choice on whether and how to adopt AI tools offered by platforms (Koehler & Sauermann, 2023), we code AI adoption independently from platform membership as discussed above. The exception are projects on iNaturalist, which by design all use the same AI tool to support contributors with species identification (coded as using augmentation).

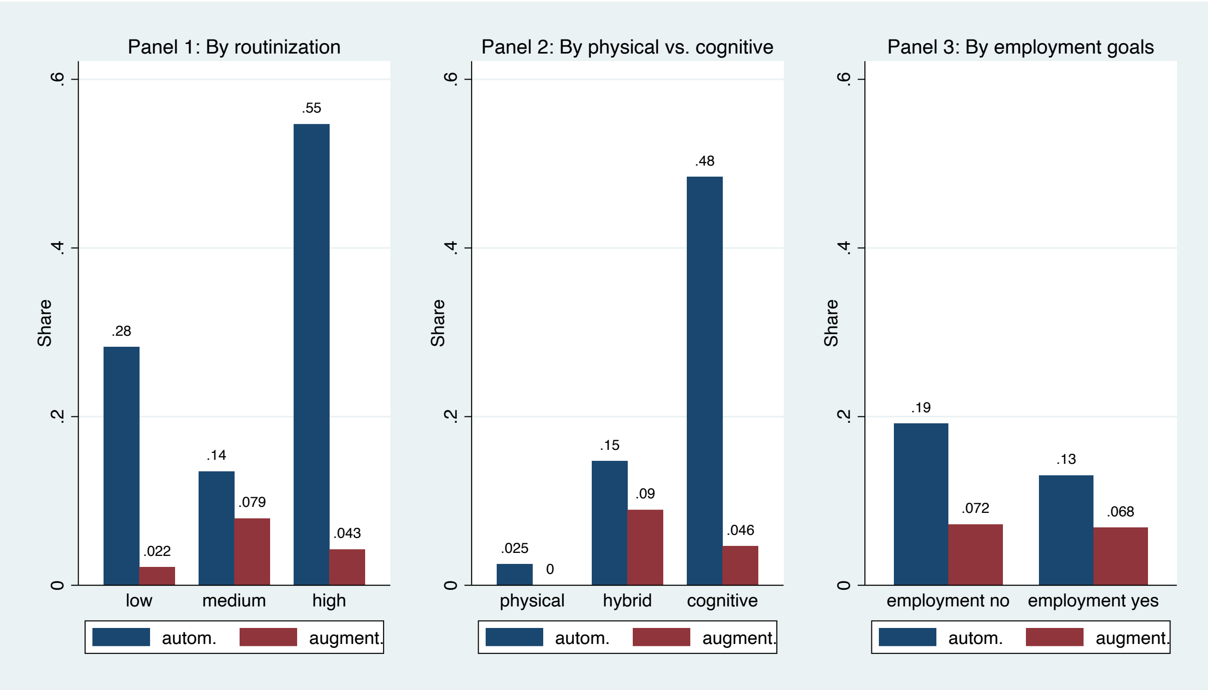
[19] A simple listing on SciStarter.org was not coded as platform affiliation. Projects using SciStarter's user management system (indicated on the profile page by "Scistarter hosted project") were coded as platform affiliated.

routinization than for tasks with medium routinization. We will verify this result in regression analyses below and explore potential underlying reasons. Panel 1 also shows that augmentation is highest in tasks with medium routinization, consistent with H1b.

Panel 2 shows that automation is highest for cognitive tasks, consistent with H2a. Augmentation is highest for hybrid tasks, consistent with H2b.

Panel 3 shows that automation is lower in projects with employment goals, consistent with H3a. Projects with employment goals seem to differ little from others with respect to augmentation, showing no support for H3b.

Figure 1: Share of projects using automation and augmentation, by task routinization, physical vs. cognitive nature of the task, and employment goals



4.2 Main regressions

We estimate a series of regressions of automation and augmentation, using linear probability models with robust standard errors (Angrist & Pischke, 2008). Model 1 in Table 3 shows that projects with high task routinization are more likely to use automation than projects with medium or low task routinization, supporting H1a. A typical example is the project Galaxy Zoo, which asks crowd members to perform highly routinized image classification tasks but then uses the training data to automate image classification using AI (Trouille et al., 2019).

Although the negative coefficient for low routinization is not significantly different from that for medium routinization ($F(1, 1252)=2.65$, $p=0.10$)), it is interesting that a notable share of low routinization projects use AI for automation (see also Figure 1, Panel A). A closer look at the particular projects in this category shows that several involve crowd members in experimentation or problem solving (low routinization tasks) and then use AI to analyze the problem solving strategies employed by many crowd members to recognize patterns and develop better algorithms for problem solving.[20] In other words, AI in these projects can learn from humans' problem solving behaviors even though problem solving is not routinized or codified. Once better problem solving algorithms are developed, they can take over work from human crowd members.[21]

Model 2 predicts the use of augmentation. Supporting H1b, the indicators for low routinization and high routinization are negative and significant, suggesting that medium routinization projects (reference group) use the highest levels of augmentation. A prominent example is the project eBirdBy analyzing data and imagery from NASA, students gain real-time insights into the development of storms, while AI predictions enable them to make informed assumptions about the future paths of storms..

Models 3 and 4 focus on the physical vs. cognitive nature of tasks as independent variables. As predicted (H2a), projects involving cognitive tasks are much more likely to use automation than projects involving physical or hybrid tasks. Typical examples include projects that ask crowd members to help with the processing or analysis of digital data, e.g., on platforms such as Zooniverse. Model 4 shows that the rate of augmentation is higher in projects involving hybrid tasks than in projects involving physical tasks and cognitive tasks (though the latter difference is not statistically significant). This is consistent with our conjecture that hybrid tasks are difficult to automate given the need for physical activities while still providing opportunities to employ AI for cognitive aspects (H2b). Typical examples are projects asking crowd members to collect biodiversity data and supporting them by providing automated species recognition (e.g., eBird, or projects on the platform iNaturalist).

Models 5 and 6 include the dummy variable indicating that a project pursues employment goals. As predicted, projects with employment goals are less likely to use automation than projects without (H3a). In contrast, projects with employment goals are just as likely to use augmentation as those without. Although the latter result fails to support H3b (predicting greater

[20] E.g., https://www.scienceathome.org/games/quantum-minds/about/

[21] The project Foldit initially involved crowd members in protein structure prediction, while also analyzing player strategies to improve algorithmic predictions. Given the improved capabilities of AI, this project now focuses on protein design. See https://fold.it/portal/node/2008706.

use of augmentation), it suggests that organizers may see a conflict between employment goals and automation but not between employment goals and augmentation. An example for the latter is the project WaterInsights, which asks crowd members to measure and report water quality in their geographic locations. In addition to collecting data for water monitoring, this project also aims to educate participants, as evidenced in the offering of special test kits and lesson plans for use in classrooms. Participants in this project can use an AI tool to interpret images of test strips, increasing the quality of data while also lowering entry barriers for participation and allowing participants to learn how to interpret test results.[22]

Models 7 and 8 include all independent variables at simultaneously. Although some of the coefficients change in magnitude, the findings for task routinization and task nature are very similar to those reported earlier. The positive coefficient for employment goals, however, becomes insignificant. This likely reflects positive correlations between employment goals and task characteristics (see supplementary analyses in section 4.3).

Finally, we highlight a few particularly interesting control variables. First, we find that large projects are more likely to use automation than smaller projects. Similarly, the rate of automation is higher among projects with international scope. These observations are consistent with the idea that large projects and those with greater geographic scope have greater resources to invest in AI, have more work to do that AI can help with, or have more data that AI can learn from. Platform-affiliated projects are not more likely to use automation but are more likely to use augmentation than stand-alone projects. The latter seems to reflect that some large platforms such as iNaturalist provide common tools that projects can use to support the work of their crowd members.

4.3 Supplementary analyses

Tables 4 reports a series of supplementary analyses. First, models 1-4 use as dependent variable a measure that collapses the measures of automation and augmentation into a single measure of whether or not projects use any AI (*anyai*). Although the coefficients of routinization and the nature of the task are similar in direction as those in the main regressions of automation, the negative coefficient for medium routinization in model 1 is notably smaller in magnitude than in the main regression of automation (Table 3, model 1). This reflects that medium routinization projects have the lowest likelihood of automation but the highest level of augmentation (see also Figure 1). Thus, although augmentation and automation both represent

---

[22] https://waterinsights.org/

AI adoption in a general sense, it is important to distinguish them because they are related in different ways to task characteristics (as well as organizer goals).

Second, we explore the possibility that employment goals play a greater role in shaping AI adoption if tasks are more amenable to the use of AI in terms of their degree of routinization or their cognitive vs. physical nature. Models 5-6 include interaction terms between employment goals and task characteristics; none of the interaction terms is statistically significant.

Third, we turn our attention to employment goals as a variable that has received little attention in prior work. To explore which projects are more likely to pursue such goals, model 7 regresses employment goals on task characteristics as well as control variables. We find that employment goals are more prevalent in projects that involve low routinization tasks, as well as physical and hybrid (vs. cognitive) tasks. The causality behind these correlations is not clear – it could be that project organizers with employment goals set up projects differently (e.g., because they fear that highly routinized tasks may provide fewer learning opportunities for participants) but also that organizers whose projects require tasks with particular characteristics find greater opportunities to pursue employment goals (e.g., tasks that require physical interaction between organizers and crowd members provide greater opportunities for exchange and learning). We also see that employment goals are more prevalent in some fields than others and are more common in projects that involve crowd members in multiple tasks. We find no significant relationships with project size, geographic scope, age, or platform affiliation. Although it is difficult to interpret these relationships, they do suggest interesting avenues for future research on why some organizers pursue employment goals while others do not, and how exactly these goals shape AI adoption, other organizational choices, and ultimately the scientific and broader impacts of projects.

Finally, we recognize that the decisions to use AI for automation and augmentation may not be independent. One way to account for this would be to estimate a multinomial logit model with no AI, augmentation, automation, as well as automation&augmentation as four outcome categories. Unfortunately, we cannot estimate this model reliably due to some small cell sizes (especially the lack of projects involving physical tasks and augmentation). An alternative is the bivariate probit model, which estimates two probit regressions simultaneously while allowing error terms to be correlated. By using probit as the foundation, this model also offers an alternative to linear probability models for estimating binary outcomes. Table 5 re-estimates our main regressions using bivariate probit. The results are very similar to those in Table 3, although the coefficients for routinization are somewhat weaker in regressions of augmentation (note that these models require us to use the same reference categories for categorical independent variables, which changes the interpretation relative to Table 3).[23]

5 Summary, limitations, and contributions

We used data from more than 1,200 crowd science projects to explore how the use of AI to automate or augment human workers relates to task characteristics as well as organizer goals. Consistent with prior theorizing, we find that the use of AI tends to be higher in projects involving more routinized tasks and in projects involving cognitive (vs. physical) activities. However, these patterns appear more nuanced than those suggested in prior literature on older technologies (Autor et al., 2003; Restrepo, 2023). In particular, although automation is highest in projects with high routinization, a considerable share of low routinization projects also take advantage of automation, e.g., by letting AI learn from the activities of many workers even if those activities are not repetitive at the level of the individual and are not codified. We also find that the patterns differ in that automation is used most frequently when tasks are primarily cognitive, while augmentation tends to be more common in hybrid tasks that combine both physical and cognitive aspects. Taken together, these results are consistent with the idea of a "progression" from no use of AI, to augmentation if AI offers some benefits but cannot fully replace human workers, to automation if tasks are both highly routinized and cognitive (Huang & Rust, 2018). Notwithstanding the importance of task characteristics, our results also highlight the potential role of a very different aspect: Whether AI is adopted and whether it is used to automate or augment workers is also related to the goals of organizers: Projects that pursue not only productivity but also goals that require human employment (e.g., education of workers) are less likely to use AI, which reflects lower levels of automation but not lower levels of augmentation.

Although our study provides novel project-level evidence on potential drivers of automation and augmentation, we recognize important limitations. First, our data come from a unique context – projects that use an open call to involve large numbers of contributors in research tasks. Although this context has advantages such as a narrow range of tasks performed by workers, considerable variation in task characteristics across projects, and high degrees of transparency for outside observers, future research is needed to replicate our results using data from other types of organizations. One particularly notable feature of our context is that labor

[23] We also re-estimated the main models excluding controls for size and geographic scope to account for the potential endogeneity of these variables. The main results are robust (available upon request).

costs are low (contributors do not get paid, although there are costs of engaging and managing them). This allowed us to focus our theorizing on capabilities of AI to deal with different types of tasks, while abstracting from the role of labor costs in shaping AI adoption (Autor & Handel, 2013; Bessen, 2016). Nevertheless, future research is needed to study the role of labor costs, as well as of the costs of developing AI.

A second limitation relates to the measurement of AI. Although our coders performed extensive searches to see whether and how projects are using AI, they had to rely on public information from sources such as websites, reports, and academic publications. Given that we coded AI use only if there was sufficient evidence, it is likely that we missed cases of augmentation or automation. Moreover, we cannot rule out that this measurement error is larger for some projects than others, although controls for aspects such as project size and age should capture some of such differences. Relatedly, all crowd science projects by definition involve human workers and we are not able to observe projects that have completely replaced humans with AI. Although such cases are likely rare[24], it is not clear whether and how our results would apply to this extreme end of the spectrum.

A third limitation is that our data are cross-sectional, yielding correlational but not causal evidence. We cannot rule out, for example, that task characteristics are partly endogenous, e.g., that organizers structure workflows differently if they seek to take advantage of AI technologies or if they pursue employment goals in addition to productivity (Agrawal et al., 2022; Koehler & Sauermann, 2023). Although this would still imply an important role of task characteristics, the observed patterns would have to be interpreted more carefully as the outcomes of multiple interrelated organizational decisions. Relatedly, we control for several potentially confounding factors such as field and project age, but we cannot rule out that other unobserved factors shape both AI adoption and our independent variables. Future research could complement our correlational evidence with a greater focus on identification, e.g., by tracking projects over time, exploiting shocks (such as the arrival of new pre-packaged AI technologies or legal worker protections), or using lab experiments to study adoption decisions under more controlled conditions.

Despite its limitations, our study contributes to a growing interdisciplinary body of research on the adoption of AI in organizations and its implications for human workers (Autor, 2015; Brynjolfsson & Mitchell, 2017; Furman & Seamans, 2019; Krakowski et al., 2023). Much of

[24] Outside of crowd science, there are a few reported examples of “robot scientists” or “self-driving labs”, but those appear to be of currently limited practical relevance (MacLeod et al., 2020; Seifrid et al., 2022; Sparkes et al., 2010).

the empirical evidence in this area comes from aggregate data at the level of jobs and occupations or from unique cases such as chess playing and medical imaging. We complement this work by analyzing data on a large number of projects that entail a very limited number of tasks each, allowing us to provide quite direct insights into the relationships between task characteristics and the use of AI for both augmentation and automation. Perhaps more importantly, the results suggest a need to broaden our view beyond task characteristics and efficiency considerations to account for the role of managerial goals related to the employment of human workers. Although such goals have been recognized as important in other contexts (Keum & Meier, 2023), they may have profound implications also for the diffusion of AI and its impact on human workers. In particular, while a narrowly efficiency-oriented model would suggest a progression from augmentation to automation as AI capabilities increase and organizations find ways to make tasks more accessible to AI, employment concerns may lead some organizations to remain at the stage of augmentation. Indeed, employment goals may lead organizations to fundamentally rethink work processes and roles to create opportunities for human-AI collaboration that may end up creating more value than seemingly superior automation (Acemoglu, 2023). This, in turn, would have implications not only for firms' performance and competitive positions but would also speak to broader concerns regarding the impact of AI on employment, inequality, and societal peace.

References

Acemoglu, D. 2023. Panel disscussion at AoM Annual Meeting.

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

Aghion, P., Jones, B. F., & Jones, C. I. 2018. Artificial intelligence and economic growth, ***The economics of artificial intelligence: An agenda***: 237-282: University of Chicago Press.

Agrawal, A., Gans, J., & Goldfarb, A. 2018. ***Prediction Machines: The simple economics of artificial intelligence***: Harvard Business Press.

Agrawal, A., Gans, J., & Goldfarb, A. 2022. ***Power and prediction: The disruptive economics of artificial intelligence***: Harvard Business Press.

Allen, R., & Choudhury, P. 2022. Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. ***Organization Science***, 33(1): 149-169.

Angrist, J. D., & Pischke, J.-S. 2008. ***Mostly harmless econometrics: An empiricist's companion***: Princeton University Press.

Arntz, M., Gregory, T., & Zierahn, U. 2016. The risk of automation for jobs in OECD countries: A comparative analysis.

Autor, D. H. 2015. Why are there still so many jobs? The history and future of workplace automation. ***Journal of economic perspectives***, 29(3): 3-30.

Autor, D. H., & Handel, M. J. 2013. Putting tasks to the test: Human capital, job tasks, and wages. ***Journal of Labor Economics***, 31(S1): S59-S96.

Autor, D. H., Levy, F., & Murnane, R. J. 2003. The skill content of recent technological change: An empirical exploration. ***The Quarterly Journal of Economics***, 118(4): 1279-1333.

Barbosu, S., & Gans, J. S. 2022. Storm crowds: Evidence from Zooniverse on crowd contribution design. ***Research Policy***, 51(1): 104414.

Bessen, J. E. 2016. How computer automation affects occupations: Technology, jobs, and skills. ***Boston Univ. school of law, law and economics research paper***(15-49).

Bianchini, S., Müller, M., & Pelletier, P. 2022. Artificial intelligence in science: An emerging general method of invention. ***Research Policy***, 51(10): 104604.

Bloom, N., Genakos, C., Sadun, R., & Van Reenen, J. 2012. Management practices across firms and countries. ***Academy of Management Perspectives***, 26(1): 12-33.

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

Brynjolfsson, E., Mitchell, T., & Rock, D. 2018. ***What can machines learn and what does it mean for occupations and the economy?*** Paper presented at the AEA papers and proceedings.

Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. 2023. Working Together, Forever? Project Evaluation, AI, and Managerial Redundancy.

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

Ethiraj, S. K., & Levinthal, D. 2009. Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. ***Organization Science***, 20(1): 4-21.

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

Frambach, R. T., & Schillewaert, N. 2002. Organizational innovation adoption: A multi-level framework of determinants and opportunities for future research. ***Journal of Business Research***, 55(2): 163-176.

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., et al. 2019. Toward understanding the impact of artificial intelligence on labor. ***Proceedings of the National***

Academy of Sciences, 116(14): 6531-6539.
Franzen, M., Kloetzer, L., Ponti, M., Trojan, J., & Vicens, J. 2021. Machine Learning in Citizen Science: Promises and Implications. In K. Vohland, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perello, M. Ponti, R. Samson, & K. Wagenknecht (Eds.), ***The Science of Citizen Science***: 183-198: Springer.
Franzoni, C., Poetz, M., & Sauermann, H. 2022. Crowds, citizens, and science: A multi-dimensional framework and agenda for future research. ***Industry and Innovation***, 29(2).
Franzoni, C., & Sauermann, H. 2014. Crowd Science: The organization of scientific research in open collaborative projects. ***Research Policy***, 43(1): 1-20.
Furman, J., & Seamans, R. 2019. AI and the Economy. ***Innovation policy and the economy***, 19(1): 161-191.
Furman, J. L., & Teodoridis, F. 2020. Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering. ***Organization Science***, 31(2): 330-354.
Greenwood, B. N., Agarwal, R., Agarwal, R., & Gopal, A. 2019. The role of individual and organizational expertise in the adoption of new practices. ***Organization Science***, 30(1): 191-213.
Grennan, J., & Michaely, R. 2020. Artificial intelligence and high-skilled work: Evidence from analysts. ***Swiss Finance Institute Research Paper***(20-84).
Gu, Z., Tang, S., & Wu, D. 2020. The political economy of labor employment decisions: Evidence from China. ***Management Science***, 66(10): 4703-4725.
Guenzel, M., Hamilton, C., & Malmendier, U. 2023. CEO Social Preferences and Layoffs.
Haeussler, C., & Sauermann, H. 2020. Division of labor in collaborative knowledge production: The role of team size and interdisciplinarity. ***Research Policy***, 49(6): 103987.
Haklay, M., Dörler, D., Heigl, F., Manzoni, M., Hecker, S., et al. 2021. What Is Citizen Science? The challenges of definition. In K. Vohland, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perello, M. Ponti, R. Samson, & K. Wagenknecht (Eds.), ***The Science of Citizen Science***: 13-33: Springer.
Huang, M.-H., & Rust, R. T. 2018. Artificial intelligence in service. ***Journal of service research***, 21(2): 155-172.
Keum, D. 2023. Managerial Prosocial Preferences and Automation Innovation. ***SSRN Working Paper #4503762***.
Keum, D. D., & Meier, S. 2023. License to Layoff? Unemployment Insurance and the Moral Cost of Layoffs. ***Organization Science***.
Kittur, A., Yu, L., Hope, T., Chan, J., Lifshitz-Assaf, H., et al. 2019. Scaling up analogical innovation with crowds and AI. ***Proceedings of the National Academy of Sciences***, 116(6): 1870-1877.
Koehler, M., & Sauermann, H. 2023. Crowds and machines: Artificial Intelligence as a manager in research projects, ***Working Paper***.
Krakowski, S., Luger, J., & Raisch, S. 2023. Artificial intelligence and the changing sources of competitive advantage. ***Strategic Management Journal***, 44(6): 1425-1452.
Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. 2022. To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. ***Organization Science***, 33(1): 126-148.
Lintott, C. 2019. ***The crowd and the cosmos: Adventures in the Zooniverse***: Oxford University Press.
MacLeod, B. P., Parlane, F. G., Morrissey, T. D., Häse, F., Roch, L. M., et al. 2020. Self-driving laboratory for accelerated discovery of thin-film materials. ***Science Advances***, 6(20): eaaz8867.

Margetts, H., & Dorobantu, C. 2019. Rethink government with AI. ***Nature***, 568(7751): 163-165.

Mollick, E., & Mollick, L. 2023. Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts, ***Working Paper***.

Obloj, T., & Sengul, M. 2020. What do multiple objectives really mean for performance? Empirical evidence from the French manufacturing sector. ***Strategic Management Journal***, 41(13): 2518-2547.

Pentland, B. T., & Rueter, H. H. 1994. Organizational routines as grammars of action. ***Administrative Science Quarterly***: 484-510.

Plouffe, C. R., Hulland, J. S., & Vandenbosch, M. 2001. Richness versus parsimony in modeling technology adoption decisions—understanding merchant adoption of a smart card-based payment system. ***Information Systems Research***, 12(2): 208-222.

Ponti, M., Skarpeti, A., & Kestemont, B. 2022. AI and Citizen Science for Serendipity. ***arXiv preprint arXiv:2205.06890***.

Raisch, S., & Krakowski, S. 2021. Artificial intelligence and management: The automation–augmentation paradox. ***Academy of Management Review***, 46(1): 192-210.

Restrepo, P. 2023. Automation: Theory, Evidence, and Outlook, ***NBER Working Paper #31910***.

Sauermann, H., & Franzoni, C. 2015. Crowd science user contribution patterns and their implications. ***Proceedings of the National Academy of Sciences***, 112(3): 679-684.

Sauermann, H., Vohland, K., Antoniou, V., Balaz, B., Goebel, C., et al. 2020. Citizen science and sustainability transitions. ***Research Policy***, 49(5): 103978.

Seifrid, M., Pollice, R., Aguilar-Granda, A., Morgan Chan, Z., Hotta, K., et al. 2022. Autonomous chemical experiments: Challenges and perspectives on establishing a self-driving lab. ***Accounts of Chemical Research***, 55(17): 2454-2466.

Shibayama, S., Baba, Y., & Walsh, J. P. 2015. Organizational design of university laboratories: Task allocation and lab performance in Japanese bioscience laboratories. ***Research Policy***, 44(3): 610-622.

Sparkes, A., Aubrey, W., Byrne, E., Clare, A., Khan, M. N., et al. 2010. Towards Robot Scientists for autonomous scientific discovery. ***Automated Experimentation***, 2(1): A1.

Stephan, P. 2012. ***How Economics Shapes Science***: Harvard University Press.

Stevens, R., Moray, N., Bruneel, J., & Clarysse, B. 2015. Attention allocation to multiple goals: The case of for-profit social enterprises. ***Strategic Management Journal***, 36(7): 1006-1016.

Trouille, L., Lintott, C. J., & Fortson, L. F. 2019. Citizen science frontiers: Efficiency, engagement, and serendipitous discovery with human–machine systems. ***Proceedings of the National Academy of Sciences***, 116(6): 1902-1909.

Tschang, F. T., & Almirall, E. 2021. Artificial intelligence as augmenting automation: Implications for employment. ***Academy of Management Perspectives***, 35(4): 642-659.

Van Brussel, S., & Huyse, H. 2018. Citizen science on speed? Realising the triple objective of scientific rigour, policy influence and deep citizen engagement in a large-scale citizen science project on ambient air quality in Antwerp. ***Journal of Environmental Planning and Management***, 62(3): 1-18.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min	Max
automation	0.18		0	1
augmentation	0.07		0	1
routinization: low	0.07		0	1
routinization: medium	0.84		0	1
routinization: high	0.09		0	1
task nature: physical	0.12		0	1
task nature: hybrid	0.72		0	1
task nature: cognitive	0.15		0	1
employment goals	0.13		0	1
field: biology	0.45		0	1
field: earth/environment	0.37		0	1
field: physical sciences	0.05		0	1
field: social/humanities	0.06		0	1
field: other	0.07		0	1
task scope >1	0.10		0	1
size: small	0.18		0	1
size: medium	0.08		0	1
size: large	0.07		0	1
size: missing	0.67		0	1
geo scope: local/regional	0.61		0	1
geo scope: national/international	0.26		0	1
geo scope: missing	0.13		0	1
age (in years)	4.95	1.521	1	6
platform affiliated	0.33		0	1

Table 2: Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1 automation	1																						
2 augmentation	0.1283*	1																					
3 routinization: low	0.0713*	-0.0543	1																				
4 routinization: medium	-0.2831*	0.0660*	-0.6296*	1																			
5 routinization: high	0.2990*	-0.0359	-0.0893*	-0.7177*	1																		
6 task nature: physical	-0.1545*	-0.1050*	0.0417	0.0519	-0.1039*	1																	
7 task nature: hybrid	-0.1513*	0.1111*	-0.1999*	0.4507*	-0.3986*	-0.6086*	1																
8 task nature: cognitive	0.3300*	-0.0419	0.2104*	-0.6083*	0.5912*	-0.1605*	-0.6856*	1															
9 employment goals	-0.0527	-0.0052	0.1581*	-0.0667*	-0.0562*	0.1429*	-0.0755*	-0.0372	1														
10 field: biology	-0.1017*	0.0864*	-0.1002*	0.1583*	-0.1131*	-0.0388	0.2102*	-0.2259*	-0.1354*	1													
11 field: earth/environment	-0.0889*	-0.0292	-0.1007*	0.1551*	-0.1085*	0.1071*	0.0548	-0.1665*	0.0713*	-0.6921*	1												
12 field: physical sciences	0.2238*	-0.018	0.1258*	-0.2550*	0.2141*	-0.0607*	-0.1473*	0.2389*	-0.0281	-0.1999*	-0.1691*	1											
13 field: social/humanities	0.0377	-0.0338	0.1668*	-0.2196*	0.1320*	-0.0084	-0.2336*	0.2983*	0.1270*	-0.2332*	-0.1974*	-0.0570*	1										
14 field: other	0.1441*	-0.0655*	0.1224*	-0.1811*	0.1224*	-0.0680*	-0.1687*	0.2722*	0.0315	-0.2516*	-0.2129*	-0.0615*	-0.0717*	1									
15 task scope >1	0.0044	0.0802*	0.2003*	-0.0853*	-0.0702*	0.2081*	-0.1039*	-0.0616*	0.1962*	-0.0271	0.0114	0.0011	0.0226	0.0089	1								
16 size: small	-0.0642*	0.1845*	-0.0255	0.0826*	-0.0830*	-0.0051	0.1201*	-0.1447*	-0.0021	0.032	0.0628*	-0.0431	-0.0763*	-0.0724*	0.0873*	1							
17 size: medium	0.1123*	0.1244*	-0.0084	-0.1097*	0.1482*	-0.0631*	-0.0334	0.0994*	-0.0136	0.022	-0.0442	0.0137	0.0391	-0.0075	0.0877*	-0.1404*	1						
18 size: large	0.2048*	0.04	0.0272	-0.1625*	0.1839*	-0.0872*	-0.1320*	0.2442*	-0.0063	-0.0513	-0.0692*	0.1258*	0.0788*	0.0517	0.0079	-0.1293*	-0.0850*	1					
19 size: missing	-0.1271*	-0.2445*	0.0106	0.0875*	-0.1216*	0.0893*	-0.0046	-0.0762*	0.0132	-0.0106	0.0135	-0.0425	-0.0048	0.0344	-0.1266*	-0.6540*	-0.4298*	-0.3960*	1				
20 geo scope: local/regional	-0.0717*	-0.0597*	-0.0258	0.1069*	-0.1139*	0.0862*	0.0569*	-0.1499*	0.0038	0.1309*	-0.0252	-0.023	-0.1084*	-0.0847*	0.0028	-0.2042*	0.0042	-0.0008	0.1630*	1			
21 geo scope: national/int'l	0.1609*	0.0631*	0.0471	-0.1839*	0.1935*	-0.0897*	-0.1300*	0.2441*	0.0468	-0.1399*	-0.0372	0.0808*	0.1723*	0.1117*	0.0216	0.1619*	0.0315	0.054	-0.1791*	-0.7469*	1		
22 geo scope: missing	-0.1077*	0.004	-0.0246	0.0866*	-0.0890*	-0.0077	0.0885*	-0.1030*	-0.0673*	-0.0068	0.0861*	-0.0731*	-0.0690*	-0.0235	-0.0327	0.0850*	-0.0477	-0.0702*	-0.0019	-0.4773*	-0.2278*	1	
23 age (in years)	-0.0554*	-0.0801*	-0.0456	0.1489*	-0.1500*	0.0554*	0.1111*	-0.1890*	-0.0119	0.1138*	0.0058	-0.0475	-0.1769*	-0.0258	-0.0245	-0.1769*	-0.0021	0.0044	0.1417*	0.8305*	-0.6660*	-0.3361*	1
24 platform affiliated	-0.0006	0.3146*	-0.0912*	0.024	0.051	-0.1713*	0.1428*	-0.0205	-0.0536	0.0262	0.0351	0.0217	-0.0753*	-0.0634*	-0.0308	0.2436*	0.0849*	-0.0134	-0.2396*	-0.1400*	0.0542	0.1333*	-0.0875*

Note: *=Significant at 5%

Table 3: Main regressions

	1	2	3	4	5	6	7	8
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	autom.	augm.	autom.	augm.	autom.	augm.	autom.	augm.
routinization: low	-0.195**	-0.043*					-0.134*	-0.047*
	[0.065]	[0.022]					[0.068]	[0.021]
routinization: medium	-0.274**	reference					-0.181**	reference
	[0.050]						[0.060]	
routinization: high	reference	-0.071**					reference	-0.077*
		[0.027]						[0.033]
task nature: physical			-0.344**	-0.053**			-0.253**	-0.052**
			[0.046]	[0.010]			[0.054]	[0.011]
task nature: hybrid			-0.236**	reference			-0.145**	reference
			[0.044]				[0.052]	
task nature: cognitive			reference	-0.033			reference	0.008
				[0.022]				[0.028]
employment goals					-0.072*	0.005	-0.045	0.011
					[0.032]	[0.022]	[0.031]	[0.023]
field: biology	reference	reference	reference	reference	reference	reference	reference	reference
field: earth/environment	0.003	-0.034*	0.005	-0.030*	0.008	-0.035*	0.011	-0.032*
	[0.021]	[0.015]	[0.021]	[0.015]	[0.021]	[0.015]	[0.021]	[0.015]
field: physical sciences	0.281**	-0.040	0.267**	-0.056	0.366**	-0.066*	0.248**	-0.044
	[0.062]	[0.033]	[0.066]	[0.032]	[0.064]	[0.032]	[0.063]	[0.033]
field: social/humanities	-0.013	-0.038	-0.062	-0.040	0.050	-0.055	-0.047	-0.042
	[0.053]	[0.028]	[0.052]	[0.026]	[0.052]	[0.029]	[0.053]	[0.028]
field: other	0.152**	-0.045*	0.100	-0.052**	0.208**	-0.062**	0.107	-0.051*
	[0.054]	[0.021]	[0.054]	[0.019]	[0.053]	[0.020]	[0.055]	[0.020]
task scope >1	-0.003	0.057*	0.034	0.065*	0.002	0.055*	0.043	0.068*
	[0.033]	[0.028]	[0.035]	[0.027]	[0.036]	[0.028]	[0.035]	[0.029]
size: small	reference	reference	reference	reference	reference	reference	reference	reference
size: medium	0.133**	0.034	0.127**	0.024	0.179**	0.022	0.111*	0.029
	[0.048]	[0.043]	[0.048]	[0.043]	[0.049]	[0.042]	[0.048]	[0.043]
size: large	0.219**	0.003	0.187**	-0.005	0.270**	-0.010	0.179**	-0.003
	[0.059]	[0.040]	[0.058]	[0.041]	[0.058]	[0.039]	[0.058]	[0.041]
size: missing	0.009	-0.084**	0.001	-0.084**	0.019	-0.087**	0.002	-0.084**
	[0.027]	[0.024]	[0.027]	[0.024]	[0.027]	[0.024]	[0.027]	[0.024]
geo scope: local/regional	reference	reference	reference	reference	reference	reference	reference	reference
geo scope: national/int'l	0.118*	-0.033	0.108*	-0.037	0.122*	-0.034	0.108*	-0.037
	[0.047]	[0.029]	[0.047]	[0.029]	[0.049]	[0.029]	[0.046]	[0.029]
geo scope: missing	-0.006	-0.072*	-0.011	-0.070*	-0.033	-0.065*	-0.006	-0.072*
	[0.042]	[0.030]	[0.042]	[0.030]	[0.043]	[0.030]	[0.042]	[0.030]
age	0.018	-0.021*	0.018	-0.020*	0.011	-0.019*	0.020	-0.021*
	[0.013]	[0.009]	[0.013]	[0.009]	[0.014]	[0.009]	[0.013]	[0.009]
platform affiliated	-0.006	0.154**	-0.015	0.147**	-0.000	0.153**	-0.019	0.148**
	[0.022]	[0.018]	[0.023]	[0.018]	[0.023]	[0.018]	[0.023]	[0.018]
Constant	0.252*	0.221**	0.241*	0.223**	0.027	0.207**	0.319**	0.230**
	[0.098]	[0.061]	[0.094]	[0.061]	[0.088]	[0.061]	[0.098]	[0.062]
Observations	1,267	1,267	1,267	1,267	1,267	1,267	1,267	1,267
R-squared	0.171	0.153	0.179	0.152	0.140	0.147	0.191	0.157
df	14	14	14	14	13	13	17	17

Note: Robust standard errors in brackets. *=significant at 5%; **=significant at 1%

Table 4: Supplementary analyses

	1	2	3	4	5	6	7
	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	anyai	anyai	anyai	anyai	autom.	augm.	employment
routinization: low	-0.163*			-0.091			0.145**
	[0.067]			[0.069]			[0.052]
routinization: medium	-0.202**			-0.091			-0.010
	[0.050]			[0.061]			[0.038]
routinization: high	reference			reference			reference
task nature: physical		-0.360**		-0.314**			0.192**
		[0.048]		[0.056]			[0.048]
task nature: hybrid		-0.218**		-0.174**			0.104**
		[0.045]		[0.053]			[0.037]
task nature: cognitive		reference		reference			reference
employment goals			-0.072*	-0.041	0.066	0.046	
			[0.035]	[0.035]	[0.220]	[0.122]	
routinization: low					-0.088	0.032	
					[0.075]	[0.032]	
routinization: medium					-0.177**	0.076*	
					[0.064]	[0.036]	
routinization: high					reference	reference	
task nature: physical					-0.102**	-0.039**	
					[0.023]	[0.011]	
task nature: hybrid					reference	reference	
task nature: cognitive					0.149**	0.003	
					[0.056]	[0.030]	
employmentXrout_low					-0.223	-0.029	
					[0.229]	[0.136]	
employmentXrout_med					-0.076	-0.033	
					[0.219]	[0.122]	
employmentXnat_phys					-0.024	-0.056	
					[0.043]	[0.032]	
employmentXnat_cog					-0.118	0.052	
					[0.138]	[0.079]	
field: biology	reference	reference	reference	reference	reference	reference	reference
field: earth/environment	-0.016	-0.011	-0.011	-0.007	0.012	-0.031*	0.077**
field: physical sciences	0.255**	0.220**	0.314**	0.215**	0.254**	-0.044	0.019
field: social/humanities	-0.013	-0.071	0.033	-0.056	-0.041	-0.045	0.213**
field: other	0.141*	0.078	0.181**	0.086	0.112*	-0.053*	0.104*
task scope >1	0.049	0.092*	0.054	0.104**	0.046	0.072*	0.157**
size: small	reference	reference	reference	reference	reference	reference	reference
size: medium	0.177**	0.159**	0.211**	0.150**	0.115*	0.029	-0.005
size: large	0.195**	0.151*	0.232**	0.146*	0.176**	-0.002	0.019
size: missing	-0.052	-0.061	-0.044	-0.060	0.003	-0.085**	0.014
geo scope: local/regional	reference	reference	reference	reference	reference	reference	reference
geo scope: national/int'l	0.077	0.064	0.081	0.064	0.106*	-0.040	0.018
geo scope: missing	-0.063	-0.064	-0.083	-0.062	-0.009	-0.074*	-0.061
age	-0.002	-0.001	-0.008	-0.000	0.019	-0.022*	-0.001
platform affiliated	0.106**	0.092**	0.111**	0.089**	-0.018	0.148**	-0.001
Constant	0.359**	0.400**	0.194*	0.442**	0.171	0.155*	-0.041
	[0.106]	[0.102]	[0.096]	[0.106]	[0.105]	[0.064]	[0.074]
Observations	1,267	1,267	1,267	1,267	1,267	1,267	1,267
R-squared	0.164	0.186	0.151	0.189	0.194	0.159	0.099
df	14	14	13	17	21	21	16

Note: Robust standard errors in brackets. *=significant at 5%; **=significant at 1%

Table 5: Supplementary analyses (bivariate probit)

VARIABLES	1a biprobit autom.	1b biprobit augm.	2a biprobit autom.	2b biprobit augm.	3a biprobit autom.	3b biprobit augm.	4a biprobit autom.	4b biprobit augm.
routinization: low	-0.520**	0.318					-0.349	0.409
	[0.197]	[0.500]					[0.208]	[0.526]
routinization: medium	-0.827**	0.691*					-0.545**	0.854
	[0.142]	[0.343]					[0.179]	[0.479]
routinization: high	reference	reference					reference	reference
task nature: physical			-1.580**	-5.992**			-1.302**	-6.327**
			[0.260]	[0.490]			[0.277]	[0.521]
task nature: hybrid			-0.722**	0.280			-0.435**	-0.207
			[0.130]	[0.273]			[0.160]	[0.369]
task nature: cognitive			reference	reference			reference	reference
employment goals					-0.323*	0.135	-0.226	0.284
					[0.153]	[0.205]	[0.159]	[0.213]
field: biology	reference	reference	reference	reference	reference	reference	reference	reference
field: earth/environment	0.028	-0.325**	0.026	-0.271*	0.045	-0.320*	0.055	-0.318*
	[0.101]	[0.124]	[0.103]	[0.126]	[0.101]	[0.127]	[0.105]	[0.131]
field: physical sciences	0.874**	-0.302	0.813**	-0.473	1.112**	-0.579	0.779**	-0.301
	[0.185]	[0.365]	[0.194]	[0.354]	[0.182]	[0.358]	[0.188]	[0.353]
field: social/humanities	0.004	-0.188	-0.200	-0.133	0.222	-0.410	-0.126	-0.251
	[0.196]	[0.391]	[0.203]	[0.353]	[0.184]	[0.389]	[0.207]	[0.380]
field: other	0.535**	-0.528	0.345*	-0.566	0.704**	-0.638	0.385*	-0.648
	[0.165]	[0.474]	[0.172]	[0.464]	[0.159]	[0.494]	[0.174]	[0.431]
task scope >1	-0.017	0.373*	0.155	0.501**	0.006	0.369*	0.172	0.541**
	[0.140]	[0.165]	[0.150]	[0.162]	[0.145]	[0.160]	[0.149]	[0.175]
size: small	reference	reference	reference	reference	reference	reference	reference	reference
size: medium	0.499**	0.254	0.454**	0.154	0.636**	0.161	0.405*	0.211
	[0.172]	[0.199]	[0.174]	[0.206]	[0.171]	[0.190]	[0.175]	[0.206]
size: large	0.726**	0.202	0.592**	0.086	0.868**	0.092	0.583**	0.086
	[0.189]	[0.233]	[0.190]	[0.258]	[0.186]	[0.229]	[0.192]	[0.252]
size: missing	0.008	-0.582**	-0.032	-0.615**	0.053	-0.614**	-0.029	-0.617**
	[0.132]	[0.149]	[0.138]	[0.153]	[0.135]	[0.148]	[0.138]	[0.156]
geo scope: local/regional	reference	reference	reference	reference	reference	reference	reference	reference
geo scope: national/int'l	0.464**	-0.250	0.420*	-0.387	0.475**	-0.322	0.436**	-0.322
	[0.169]	[0.269]	[0.169]	[0.273]	[0.167]	[0.272]	[0.168]	[0.269]
geo scope: missing	-0.172	-0.499	-0.207	-0.612*	-0.266	-0.482	-0.180	-0.542*
	[0.203]	[0.260]	[0.209]	[0.267]	[0.204]	[0.259]	[0.209]	[0.263]
age	0.084	-0.167*	0.078	-0.194*	0.050	-0.170*	0.090	-0.178*
	[0.051]	[0.077]	[0.052]	[0.081]	[0.050]	[0.078]	[0.052]	[0.079]
platform affiliated	-0.033	1.241**	-0.084	1.177**	-0.010	1.222**	-0.099	1.212**
	[0.101]	[0.143]	[0.101]	[0.148]	[0.100]	[0.141]	[0.103]	[0.152]
Constant	-0.975**	-1.476**	-0.920**	-0.805	-1.584**	-0.770	-0.735*	-1.306*
	[0.351]	[0.514]	[0.352]	[0.567]	[0.330]	[0.488]	[0.357]	[0.535]
athrho	0.536**		0.451**		0.449**		0.519**	
	[0.105]		[0.108]		[0.102]		[0.105]	
Observations	1,267	1,267	1,267	1,267	1,267	1,267	1,267	1,267
df	28	28	28	28	26	26	34	34

Note: Robust standard errors in brackets. *=significant at 5%; **=significant at 1%