Labor as Capital: AI and the Ownership of Expertise

PRELIMINARY & INCOMPLETE

This draft: March 27, 2026. First Draft: August 12, 2025.

Zoë Cullen
Harvard

Danielle Li
MIT

Shengwu Li
Harvard

Abstract

Many AI tools are trained on knowledge generated by workers. When surveyed, workers report having substantial control over how much knowledge they share with employers and are willing to restrict it when they learn that that their expertise may be used to train models to perform similar work. Motivated by this evidence, we develop a formal model of "knowledge supply" to AI models and show that, given current institutions, workers may restrict knowledge supply to the detriment of overall productivity. Alternative policies can recover efficiency and raise both firm and worker welfare. Individual data ownership, despite being preferred by workers, eliminates knowledge withholding but creates negative externalities for workers: one worker's data strengthens the firm's bargaining position against others, potentially making all workers worse off. In contrast, collective data ownership achieves the first-best outcome, restoring knowledge supply while allowing workers to benefit from AI-driven productivity gains. These findings highlight the importance of labor agreements in shaping AI adoption in labor markets.

JEL Codes: C93, D83, J31, J41, J71.
Keywords: AI, innovation, contracts, labor markets, future of work

Cullen: zcullen@hbs.edu. Li: d_li@mit.edu. Li: shengwu_li@fas.harvard.edu.

This project was reviewed by the Institutional Review Board at MIT. Aaron Leong, Nicolas Miller, and Yuerong Zhuang provided excellent research assistance.

1 Introduction

In modern workplaces, surveillance is routine. Employees communicate on recorded platforms, project managers are tracked through task management tools, and call center agents, warehouse staff, and remote workers are all frequently monitored via audio, video, or screen recordings. As a result, work increasingly generates data about work: records of how exactly people do their jobs. Once used mainly for oversight and evaluation, these data can now be repurposed to train artificial intelligence (AI) systems to perform some of the very labor they document.

Consider a common application of generative AI: customer service chat assistants trained on recorded interactions between agents and customers. Firms often use recordings from top performers so that models can replicate their problem-solving approaches and communication style. Once developed, these models can be deployed across locations, scaling the expertise of top workers and raising the productivity of others (Brynjolfsson et al., 2025).

This reflects a broader shift in the nature of labor. Traditionally, expertise resides with workers and is difficult to transfer because it involves tacit knowledge developed through experience (Polanyi, 1967). Firms therefore "rent" this expertise by employing workers period by period. Surveillance-enabled AI changes this dynamic: by recording how workers perform their jobs, firms can more readily codify that expertise into AI systems that can then be owned and scaled.

This dynamic has a historical precedent in the scientific management movement of the early 20th century. Frederick Winslow Taylor, a pioneer of scientific management, observed that factory workers had substantial informal know-how. He argued that this "traditional knowledge may be said to be the principal asset or possession of every tradesman," and advocated that management should surveil skilled workers and codify their techniques (Taylor, 1911). Workers resisted this expropriation individually, by restricting output and concealing efficient methods, and through collective actions such as the Watertown Arsenal strike of 1911 (Montgomery, 1979; Aitken, 1960)

We refer to the recorded aspects of labor as workers' *knowledge supply.* To date, many AI systems have been trained on knowledge—documents, art, code—that workers have supplied without awareness or consent. Yet a substantial share of work-relevant expertise remains uncodified, residing with individual workers. Effectively deploying AI in the workplace will require data that captures this contextual expertise—knowledge specific to firms, clients, and moments in time.[1] As workers recognize the potential for AI to capture and replicate their skills, they may alter how they work or

[1] A variety of popular press articles cite the importance of tacit or context specific expertise. See, for instance, Murty and Kumar S (2026) and Hu (2025).

advocate for new governance regimes. This raises both a descriptive and normative question: how do current labor market policies impact workers' knowledge supply? And what policies should be adopted to maximize welfare?

We argue that existing labor arrangements neglect the role of knowledge supply, to the detriment of both productivity and worker welfare. Empirically, we conduct an online survey experiment and present evidence that 1) workers believe they possess substantial knowledge beyond their firms' documentation or training materials, which we refer to as uncodified knowledge; 2) that they perceive themselves as having the ability to influence how much knowledge they supply to their employers in the form of recordings and documentation; and 3) they prefer policies that give them the ability to control and sell their individual work-relevant data.

Theoretically, we develop a model of workers' *agency* in sharing knowledge to develop AI systems. In our model, firms can use knowledge supplied by workers to train AI systems that replicate their expertise. Anticipating that their knowledge contributions today improve the firm's outside option tomorrow, workers rationally withhold knowledge, reducing both current productivity and the quality of the AI systems firms can build. We then evaluate alternative data governance regimes. While workers' preferred policy, individual data ownership, restores knowledge sharing by allowing workers to sell their data, it creates a competition externality that drives down payments: because workers' data are substitutes, each worker's sale strengthens the firm's bargaining position against others. By contrast, collective bargaining over work data eliminates this externality and can achieve both efficient knowledge sharing and a more equitable division of the gains from AI.

The first part of our paper provides motivating empirical facts using surveys and survey experiments. In our primary study, we recruit 497 employed U.S. workers from the survey platform Prolific and ask them about their workplace knowledge and their preferences over how their work data is used. First, we document that workers report possessing substantial uncodified knowledge. Large majorities report having knowledge about customers, projects, and organizational processes, as well as judgment, problem solving, and communication skills that are not formally documented. Second, workers report considerable discretion over the knowledge data they generate about their work. Most say they can both withhold such data or affirmatively supply more—for example by changing the quality of their documentation, or how much they communicate on monitored platforms. Finally, we elicit preferences over three policy responses after describing how workplace data is used as training data: one that bans the use of surveillance data for AI training; one that grants individuals the right to sell their work data; and one that allows workers to collectively negotiate the use of data about their work for AI training purposes. While respondents support all three, individual data ownership emerges as the most popular option.

In our second study, we provide experimental evidence that workers withhold knowledge by exploiting the fact that the type of survey-response work performed by Prolific workers is also increasingly performed by AI models.[2] We randomly provide half of participants with truthful information about how worker data can be used to train AI systems. Treated participants watch a short video explaining that AI tools can be trained on records of worker behavior, including survey responses, while control participants view a similar video that omits any reference to worker data. We then offer to purchase access to respondents' past survey data and elicit their willingness to complete future surveys for pay. This design allows us to identify the causal effect of workers' awareness of how their data may be used for AI development on their willingness to supply it. We find that awareness decreases willingness to supply prior survey data, and increases reservation wages to participate in future surveys.

Taken together, our survey evidence shows that workers report possessing valuable uncodified knowledge and substantial discretion over how much data they share about their work. When workers learn that this data can be used to train AI systems to perform similar tasks, they become more likely to withhold it, suggesting that workers' knowledge supply is potentially elastic.

Motivated by this evidence, the second part of our paper develops a model of knowledge supply and AI development. Our model focuses on one important reason why workers may adjust their knowledge supply: career concerns. Specifically, when firms can use workers' knowledge to build technologies that perform the same work tomorrow, workers may have incentives to limit how much knowledge they reveal today.

In our model, one firm and multiple workers interact over two periods. In each period, the firm hires workers, who then decide how much knowledge to contribute. Workers differ both in their maximum potential knowledge contribution and in the cost they incur to withhold knowledge. The firm's output increases with workers' knowledge contributions and is additive across workers and across time. Workers are subject to workplace surveillance so that, through the act of working, they generate records of their processes and expertise. As a result, workers supply knowledge by default and must take costly actions to *withhold* knowledge by evading surveillance (we later extend the model to allow workers to incur costs to share additional knowledge).

[2] A growing number of companies are attempting to build AI models of human survey and marketing preferences. See, for example, https://hai.stanford.edu/news/ai-agents-simulate-1052-individuals-personalities-with-impressive-accuracy.

Knowledge supplied in period 1 becomes training data for an AI system used in period 2. Given a vector of period-1 knowledge contributions k_1, the firm trains an AI system of quality $\alpha(k_1)$, where α is a real-valued non-decreasing function. Even if a position goes unfilled, the firm can still generate output equal to $\alpha(k_1)$. AI in this model is therefore *expropriative*: the AI system improves the firm's outside option in period 2 by replicating workers' expertise. This can be interpreted either as literal AI automation or as the hiring of a low-skill worker augmented by the AI system. We assume that workers' knowledge contributions are substitutes in training the AI system, such that α exhibits decreasing differences. This captures the idea that once the firm has substantial knowledge data from some workers, the marginal contribution of additional workers' knowledge to AI capability becomes smaller.

Wages and employment in each period are determined by Nash-in-Nash bargaining between the firm and the workers. Two frictions are obstacles to efficiency. First, knowledge contributions are not contractible, so firms cannot simply offer to pay workers more in return for higher contributions. This reflects the tacit nature of much workplace expertise: firms cannot observe what individual workers know and therefore cannot write contracts that demand particular contributions. Second, period-1 contracts determine period-1 wages and employment but cannot commit parties to period-2 outcomes. As a result, workers face career concerns: their knowledge contributions today influence their bargaining position tomorrow.

In the baseline model without AI, workers have no reason to withhold knowledge. An equilibrium therefore exists in which all workers are employed in both periods, each contributes their full knowledge, and each receives a share of output proportional to their Nash bargaining weight.

Our next set of results consider the legal status quo case in which worker data, and the AI model they contribute to, is owned by the firm.[3] Suppose, first, that workers are unaware that their data can be used to train AI models, as is likely the case for many workers today. Workers will then naively continue to supply their full knowledge in period 1. While this (weakly) increases output in period 2 (because the AI system may be more productive than some workers), it also improves the firm's outside option, weakening workers' bargaining positions and lowering their wages in period 2. The gains from AI therefore accrue primarily to the firm, raising profits while reducing worker surplus relative to the no-AI benchmark.

However, workers are unlikely to remain naive about the potential uses of their work data. Once workers anticipate that their knowledge contributions may be used to train AI systems that

[3] For a legal perspective on ownership of labor data, see Viljoen et al. (2026).

compete with them in the future, they have incentives to withhold knowledge in period 1 in order to preserve their bargaining position in period 2. We show that, as long as withholding costs are not prohibitive, there exists an equilibrium in which workers choose to withhold knowledge, leading to a (weak) reduction in time-2 wages compared to the baseline no AI case. Relative to the naive case, knowledge withholding increases worker surplus, but is socially destructive, as it decreases overall (and firm) surplus. We further show that withholding is harder to sustain in equilibrium when workers are more substitutable. When colleagues can provide comparable data, an individual worker's withholding does less to protect their job from automation. Anticipating this, workers put less effort into withholding knowledge, and their period-2 wages fall as expropriation increases.

We use this model to shed light on several alternative policies: banning workplace surveillance, granting workers individual ownership over their work data (workers' preferred option, according to our survey evidence), and establishing collective ownership of work data among workers.

Banning surveillance cuts off the data necessary to train AI models and, in our setting, is akin to banning AI development as a whole. This policy restores the baseline no-AI equilibrium, improving workers' job security, but forgoing potential productivity gains from AI.

We model individual ownership as giving workers the right to bargain over whether their data is used at time-2, and at what price. In the event of disagreement, the firm would be unable to use that worker's data for AI model training. Individual ownership allows workers to profit from the value of their work data. As such, workers no longer have a motive to withhold knowledge, and there is an equilibrium with full knowledge contributions. In this way, individual ownership raises time-1 output while enabling the full productivity gains from AI. However, individual ownership does not guarantee that workers share in the gains from AI. Intuitively, individual ownership encourages each worker to contribute data, but their data has an externality: it improves the firm's bargaining position vis-a-vis other workers, lowering other workers' future wages. When workers' data are more substitutable, this effect drives down wages as the marginal value of their knowledge contributions is smaller. In the extreme case when workers are symmetric and their data are perfect substitutes, workers receive *none* of the output gains from AI, no matter their Nash bargaining weight.

The last policy we consider is one in which workers collectively own their data and bargain jointly with the firm over wages (in the event of disagreement, the firm cannot use any worker's data). Like individual ownership, collective ownership restores workers' knowledge contributions to the social first best. However, collective ownership also eliminates the competition externality present under individual ownership, because one worker's contribution no longer improves the firm's bargaining position against other workers. As a result, workers capture a larger share of the surplus generated by AI. Moreover, unlike individual ownership, collective ownership always leaves *both* workers and firms weakly better off relative to the no-AI benchmark. While our simple model abstracts from intra-union frictions, it nonetheless suggests that collective action could be a useful policy tool to ensure that workers share in the gains from AI.

Finally, motivated by our empirical evidence, we consider an extension in which workers can contribute *more* knowledge at a cost, in addition to withholding knowledge. When firms own worker data, workers have no incentives to make costly knowledge contributions, and indeed, continue strategically withholding data, as in the baseline firm-owned equilibrium. In contrast, we show that workers who individually own their data exert additional effort to codify their expertise so that they can be compensated for resulting AI improvements. However, the welfare effects of this additional effort are ambiguous: with costly contributions, it is possible that workers will exert more effort than is socially efficient, reducing total surplus compared to the no-AI case. By contrast, when workers collectively own their data, total surplus is always at least as high as in the no-AI case.

Overall, despite the popularity of individual ownership among our survey respondents, our model suggests two advantages of collective ownership. First, collective ownership always improves *worker* surplus compared to no AI, whereas individual ownership sometimes does not. Second, when workers can contribute additional knowledge at a cost, collective ownership always improves *total* surplus compared to no AI, whereas individual ownership sometimes does not.

These findings have important implications for governing the transition to AI-augmented production. While policy debates often focus on consumer data and privacy, workplace data raise distinct questions about the ownership of surveilled human capital. Greater worker awareness, absent changes in data rights, may reduce productivity as workers withhold valuable knowledge. Moreover, workers' preferred policy—individual data ownership—may not maximize their welfare because workers fail to internalize the negative impacts that their private data sales have on others. Our framework highlights policies that help workers internalize these externalities and better align private incentives with social welfare.

2 Background: Surveillance-enabled AI Development

While efforts to codify labor date back to at least the scientific management practices of the Industrial Revolution, these modern tools vastly expand its scope, granularity, and potential consequences.

Indeed, concerns over "AI expropriation"—whereby firms use worker data to train models that replicate their labor—are already surfacing in labor disputes across a range of industries (Glass, 2024). In this section, we provide background on the rise of workplace surveillance, the use of its data byproducts in AI development, and the institutional and legal frameworks shaping debates over the ownership and use of worker-generated data.

2.1 Modern Workplace Surveillance

Employee monitoring has become a pervasive feature of modern workplaces, with firms using "bossware" to track worker behavior. In offices, this includes keystroke logging, screenshots, and monitoring of emails and chats; in physical jobs, it extends to cameras, geolocation, and wearable biometric devices (U.S. Government Accountability Office, 2024). Survey evidence highlights the scale of monitoring. A 2024 OECD survey finds that about 90% of U.S. firms monitor workers in some way, while a worker survey reports that 68% of employees experience at least one form of electronic surveillance (Milanez et al., 2025; Hertel-Fernandez, 2024). Monitoring is especially common in large firms and expanded rapidly after the shift to remote work (Kantor and Sundaram, 2022; Turner, 2022).

Surveillance technologies often generate durable records of worker behavior. In the OECD survey, 75% of U.S. firms report collecting data on their workers through surveillance tools. Among firms that collect such data, most do not allow workers to opt out (Milanez et al., 2025). While existing policy concern has focused primarily on the direct consequences of monitoring—such as its effects on safety, anxiety, and privacy—there has been little attention to the fact that these same data can be repurposed to train workplace AI systems, including models that may ultimately perform or replace the very tasks being monitored.[4]

2.2 Worker Data and AI Development

Recent advances in AI, particularly large language models (LLMs), rely on both public and proprietary data. Foundational models (e.g., GPT) are trained on internet text, books, and code to develop general capabilities but often lack the domain-specific knowledge needed for organizational tasks.

Bridging this gap requires fine-tuning on proprietary, domain-specific data—such as customer support logs, call transcripts, and internal documentation—often generated by workers performing

[4]For example, U.S. Government Accountability Office (2024); Milanez et al. (2025); Hertel-Fernandez (2024) focus on documenting concerns related to autonomy, privacy, and work place safety, but no studies mention the use of surveillance data for AI training.

these tasks. These data capture not only formal procedures but also tacit knowledge: skills and heuristics observable in practice but difficult to codify (Polanyi, 1967). Surveillance technologies facilitate this process by recording how experienced workers perform their jobs.

The following examples illustrate how worker-generated data can be used to build AI systems:

> *Call Recordings and Customer Service AI*: Customer service models are trained on call transcripts and audio logs. In many cases, these transcripts are labeled not only with objective performance metrics such as call duration but also with specific indicators for whether the text was generated by a recognized top-performing agent. This type of labeling allows the fine-tuning process to capture and scale the skills of specific individuals (Brynjolfsson et al., 2025).
>
> *Screen Recording and Robotic Process Automation (RPA)*: Screen recordings and event logs capture employees' digital workflows, which RPA systems use to automate routine but highly context-specific tasks such as how to submit an expense report in a given firm (Rabbit Inc., 2024).
>
> *Clinical Notes and Medical AI*: Clinical notes written by providers are used to fine-tune medical AI systems. For example, NYUTron is trained on years of clinical documentation (Jiang et al., 2023). Such models can be used not only to provide clinical assistance, but also to automate healthcare tasks such as generating reports.

In addition to these current use cases, which primarily build on text or image based worker inputs, recent AI research is incorporating other types of worker input, such as video data. In robotics, for example, systems can now learn complex manual skills directly from human video demonstrations, leading to an industry of "arm farms" in which workers record themselves performing manual tasks in order to provide training data for robotic AI models (Christopher, 2025). More broadly, advances in AI may expand the kinds of worker data that can be used to train future models, making surveillance data that seems uninformative today potentially valuable in the future.

Finally, debate continues over whether AI will achieve human-level general intelligence (Grace et al., 2022; Allyn-Feuer and Sanders, 2023). While definitions of artificial general intelligence (AGI) emphasize broad reasoning and problem-solving capabilities (Bubeck et al., 2023), even highly capable models would still require context-specific training and examples to excel at actual workplace tasks (Mitchell, 2021). Although advanced systems may learn such knowledge autonomously, it is likely more efficient to provide models with examples from human experts who already possess this information. As a result, human-generated data and experience are likely to remain important inputs to AI systems (Ramani and Wang, 2023).

2.3 Institutional and Legal Context

The use of worker data in AI development raises important legal concerns. U.S. law currently grants employers broad control over workplace data. Consumer privacy laws often exclude employment data, and under the "work for hire" doctrine, materials produced on the job belong to the firm (Kim and Leavitt, 2026; U.S. Congress, 1909, 1976).

New AI capabilities challenge the sustainability of the current legal framework. For example, the work for hire doctrine was designed for tangible products like reports or designs, not the behavioral trace data that are now routinely captured in the modern workplace. Historically, such process-level data were neither valuable nor feasible to collect. Today, however, they may be important inputs for training AI systems that aim to replicate human expertise (Ajunwa, 2025).

These changes are sparking new legal and labor responses. For example, the Screen Actors Guild secured limits that bar studios from using an actor's video recordings to create an AI avatar without consent (SAG-AFTRA, 2023). More broadly, researchers and advocates have proposed new governance frameworks—such as collective data rights or worker data trusts—to ensure more equitable participation in the value created by workplace AI (Viljoen et al., 2026; Kim and Leavitt, 2026; Diamantis, 2023).

These developments underscore a core tension: AI systems rely on worker-generated data, yet little is known about workers' willingness to supply this data in a world where it can be used to alternatively scale or replace them. In the next section, we present empirical evidence on worker's job expertise: what knowledge they have, their ability to resist or aid its codification, and their preferences over alternative policies governing the data they generate at work.

3 Motivating Empirics

We conduct two related online surveys: the first provides broad descriptive evidence on worker's job expertise and the second provides experimental evidence on their willingness to share in a specific context.

3.1 Descriptive Survey

Our first survey focuses on the primary jobs of full-time employed workers in the U.S., recruited through the survey platform Prolific. The survey collects information on participants' work, with a focus on job-specific expertise that they possess, as well as their ability to influence how much of this knowledge can be recorded or documented by their employer.

The survey begins by collecting baseline information on respondents' jobs (occupation, industry, experience, firm characteristics) as well as on their familiarity with AI in the workplace (their use of AI tools, and their understanding of AI development). We ask respondents about the knowledge they use in their work and the records or traces of that knowledge that may be captured during employment. We also collect measures of how aware participants are of how the data generated about their work may be useful for the training of AI models.

We next ask workers a set of questions about their preferences and behavior under common understanding of how workplace data may be used in AI development. To ensure that respondents share a common understanding, all participants view a brief two-minute video highlighting how AI systems can learn from human-generated work data to replicate job tasks. The video provides several examples of how AI models can perform real workplace activities (customer service, software development, and administrative tasks) and shows how data generated by workers can be used to train these systems (call transcripts, written code, and screen-recording logs). The video therefore provides concrete examples of what we mean by labor data, giving respondents a common reference point when answering subsequent questions about the extent to which they generate or could provide such data to their employers.

After viewing the video, we collect responses about 1) the degree of expertise workers believe they have about different aspects of their jobs; 2) the extent to which their work is currently recorded by workplace surveillance systems versus embodied; and 3) their perceived ability to supply either more or less knowledge in the form of documentation or digital records for their employer.

3.1.1 Sample Descriptives

We recruited 497 currently employed U.S. workers through Prolific in the first quarter of 2026. The survey questions were designed around the person's *primary* job.

To provide a general sense of who the subjects are, we begin by presenting descriptive statistics of baseline characteristics, described in Table 1. By design, all were employed full-time at the time of the survey. Appendix Figure A1 illustrates the breakdown of respondent occupations, with the most common being those in management (17%), computer and mathematical (14%), business and financial operations (11%), education and library (9%) and office and administrative (8%) roles. Common job titles include manager, teacher, and software engineer. Just over 50% of workers spend at least 5 days in the office while 18% are fully remote. Most workers work on a salaried basis (57%), earning a median annual salary of $81,750. The remaining 47% are hourly workers, earning a median hourly rate of $22.60. Finally, the median subject is 36 years old. Fifty percent of subjects are male and 74% are White. We include a comparison between our sample and a cross-section of the U.S. full-time workforce in Appendix Table A1.

Our baseline questions assess three issues important for our study: whether workers are subject to workplace surveillance, whether workers possess specific knowledge that may be valuable to their firms, and whether workers can share more or less knowledge

3.1.2 Workplace Knowledge

We begin by establishing that workers believe they hold substantial workplace knowledge.

Panel A of Figure 1 reports workers' self-assessed expertise across several domains: customers or clients, colleagues and organizational processes, current projects or products, markets or external conditions, technical or professional skills, and judgment, problem-solving, or communication skills. For all categories except market trends, over 90% of workers report having "some" or "a lot" of personal expertise, and at least 50% report having "a lot."

Panel B asks how much of this expertise goes beyond what is captured in company documentation, training materials, or official records. The shares reporting uncodified knowledge are similar across domains, indicating that most workers believe some of their expertise is not formally recorded. Uncodified knowledge is especially common for skills that are tacit (over 60% of workers report having "a lot" of uncodified knowledge regarding judgment, communication, problem solving) or highly contextual (almost 50% of workers report having "a lot" of uncodified knowledge of specific people, processes, or client preferences within the firm). Appendix Figure A2 provides the breakdown by worker occupation.

In Figure 2, we construct an index of workers' uncodified knowledge (the share of work-relevant areas in which workers report having "a lot" of uncodified expertise) and correlate it with worker characteristics. Panels A and B show that workers with higher education and income report more uncodified knowledge. Panels C and D show that workers who are more unique (measured either by the number of colleagues performing the same role or by their self-reported uniqueness) also report possessing more uncodified knowledge.[5]

To move beyond general perceptions, we also ask about two specific forms of uncodified information: work-relevant content stored in personal communications (e.g., emails or chats on non-work accounts) and in personal AI prompt histories. Appendix Figure A4 shows that a majority of workers report having work-relevant information in both places: 55% in personal communications and 59% in AI prompt histories among AI users (43.3% in the full sample).

Taken together, these results suggest that many workers possess valuable but uncodified knowledge about how to perform their jobs effectively. The absence of firm-specific contextual expertise in AI training is often cited as a reason for slow adoption: without it, AI systems tend to perform less reliably in real organizational settings (Spirlet, 2026).

3.1.3 Workplace Surveillance

Next, we asked workers about the extent to which they are subject to various forms of workplace surveillance. One potential way firms can access workers' knowledge is through the use of various types of physical and digital monitoring, which can generate detailed records of workers' activities and interactions.

Figure 3 describes workers' beliefs about whether their employer monitors various aspects of their workplace activity. A large majority of respondents report that their employer monitors or collects data on multiple aspects of their work: 76% cite performance tracking, 74% mention communication monitoring, 62% report surveillance of output and deliverables, while 56% of workers report having their time tracked. Additionally, substantial minorities of workers indicate that they are subject to various types of recording: 48% to computer activity monitoring, 43% to video monitoring, 39% to location tracking, and 35% to audio recording. Fewer than 30% of workers report having received formal guidelines or policies about what data may be collected about their work or how it could be used.

3.1.4 Knowledge Supply

Our main descriptive findings relate to workers' ability to influence the amount of data they supply to their employers. To assess this, we describe both "passive" forms of data collection—where information is captured through monitoring of normal work behaviors (e.g., screen recordings or

[5] Results are robust to controlling for occupation fixed effects. See Appendix Figure A3 for binned scatterplots.

communications on monitored platforms)—and "active" forms—where workers are explicitly asked to demonstrate or document their knowledge (e.g., training videos or written documentation of workflows). We then separately ask whether workers could take actions to provide less or more information to their employer (*"Are there actions you could take to provide your employer with LESS (MORE) information about your work behaviors?"*). To separate ability from willingness, respondents are instructed to focus only on whether it is *possible* to provide less information, not whether they would *want* to do so.

Workers believe they have considerable discretion over the data they provide to their employers. Overall, 95% percent say they can take actions to withhold knowledge data from their employers. Panel A of Figure 4 shows the extent to which workers believe they have the ability to withhold knowledge from their employer across relevant domains of their job. 79% report that they have at least "some" actions they could communicate more off-the-record and 34% report they have "a lot" of such actions. 64% say they can take at least some actions to create less detailed or lower-quality documentation, and 24% say they can take a lot. Finally, roughly half of the respondents indicate some ability to avoid physical (56%) or digital (54%) surveillance and about one fifth (19% and 18% respectively) feel they have a lot of ability to do so.

Panel B shows the same results but for whether workers have actions they can take to provide their employers with more data. Again, workers report considerable flexibility with 97% report being able to supply additional knowledge data in some way. Relative to withholding, a greater share, approximately 40% of workers, report having "a lot" of actions they can take to affirmatively document more of their work expertise, such as by creating higher quality documentation (47%), communicating more on recorded platforms (42%), or generating additional recordings of their work, whether in-person (36%) or digitally (37%).

Appendix Figure A5 reports how knowledge supply actions varies across occupations. Workers in computer-related occupations, as well as those in business, financial, and management-related occupations report having the most options to either withhold or share their knowledge, with 25-30% of such workers saying that they have "a lot" of ways to withhold data and 35-45% of saying that they have "a lot" of ways to share more data. Appendix Figures A6 and A7 report binned scatterplots between an index of withholding and sharing actions respectively, and measures of work or worker characteristics, controlling for occupation fixed effects. In general, we find that the same workers who tend to have more uncodified knowlege—those who are more educated, have higher earnings, or who perceive themselves to be more unique—are also the ones who report having more ability to withhold or share it.

3.1.5 Policy Preferences

Finally, we survey workers on their preferences over several policies related to their labor data. We present experimental vignettes to our treated subjects, describing three alternatives (see Appendix B for the full wording of these policy alternatives). The first policy bans the use of passively-collected work data in AI development[6]; the second allows for individual control over one's own work data, including its sale [7]; and the third allows for collective worker control and sale of work-data [8]

To assess workers' preferences, we follow the methodology of Buckman et al. (2025).[9] Figure 5 displays the share of workers who indicate that they like each policy. While there is net positive support for all three options, individual ownership of work data emerges as the most popular choice. Approximately 64% of respondents favored a policy that would grant them the right to own and sell their individual work data for AI development. Figure 6 shows that most workers (71-75%) are not willing to take a pay cut in exchange for the implementation of a policy they support (nor do they need to be compensated if an unfavored policy were implemented). However, 8.5% of workers are willing to take at least a 10% pay cut in exchange for individual data ownership. In comparison, 12.4% and 11.5% of workers are willing to take 10% or greater pay cuts for collective data ownership or restrictions on AI development, respectively.

3.2 Knowledge Supply Experiment (Survey Work)

Finally, we provide some incentivized evidence on workers' willingness to *actually* shift their knowledge supply. To do so, we focus on a type of labor that all Prolific respondents engage in: paid

[6] "Under this policy: 1. Employer could not use work data to develop AI models, or sell work data to other firms to develop AI models'. 2. Those seeking to develop AI models would have to hire workers to specifically produce AI training data; they could not use data that was produced as the byproduct of every day work tasks."

[7] "Under this policy: 1. You decide if your data can be used for AI training. You could say no, and your employer couldn't include your data in their AI models. 2. Your employer can pay you to use your data, at a price that you both agree on."

[8] "Under this policy: 1. You and other workers in your role jointly decide whether your collective work data can be used for AI training. You could collectively say no, and your employer couldn't include any of your data in their AI models. 2. Your employer can pay you and your colleagues to use your collective data, at a price that you all agree on. You and your coworkers could then decide how to split the proceeds amongst yourselves individually."

[9] Specifically, for each policy, we ask: "How would you feel if this policy were in place where you work?" (Positive, I would like it / Neutral / Negative, I would dislike it). Depending on their response, participants were shown a contingent valuation item: "How big a pay cut would you accept in exchange for such a policy?" or "How much extra pay would you need in exchange for such a policy?" with response categories ranging from "'None" to "More than a 35% pay raise (or pay cut)."

survey response. This type of work is increasingly subject to automation by AI agents. "Synthetic panels," in particular, are an emerging application of AI tools in which AI systems seek to replicate human preferences in order to serve as a cheaper and more widely available tool for understanding preferences across different types of consumers (Henriques, 2025). Such models are often trained on large-scale human preference data, including the data collected through Prolific and other platforms. As such, survey takers' own survey responses can be used to train models that mirror their preferences or opinions.

We collect a separate sample of 971 participants and tailor our outcome measures to capture knowledge supply associated with survey taking itself. In contrast to our descriptive study where we provided information about the use of worker data in AI development to all participants, in this survey we randomize this information, with 50% receiving the same treatment video as in the descriptive survey, and 50% receiving a control video that covers the same topics (about how AI models are capable of performing human work), but without mentioning the use of worker data in AI training.[10] The aim of this experiment is therefore to assess the impact of a change in awareness about how labor data can be used on workers' willingness to supply it.

Following treatment, we assess workers' willingness to supply their labor data in various ways: allowing us access to records from their prior Prolific survey work, working to answer future surveys, and answering a short personality assessment in our present survey.

Panel A of Figure 7 shows that treated workers are substantially less willing to supply their previous Prolific survey data for a $10 payment.[11] We find that 41% of treated workers refuse compared to 25% of control group workers. Column 1 of Table 2 shows that this 16 percentage point (or 66% increase) increase in withholding is highly significant. This level of withholding is sufficient to potentially introduce substantial selection bias concerns for any synthetic sample, effectively lowering the productivity of the AI model. We note that this payment represents a large payment, given that the average payment for participation in our survey was $2.50 and that granting access to prior surveys does not require any additional survey work for the subject. Panel B shows that treated workers are also less willing to supply data in the future.[12] 28% of treated workers

[10] The treatment split deviates slightly from 50% due to the removal of unreliable observations during data cleaning, including respondents who completed the survey in an implausibly short time or reported difficulty understanding the questions.

[11] The specific question is "Are you willing to let us use your responses and metadata from your previous Prolific surveys for a bonus of $10? (For a share of willing participants, we will pay this bonus and seek the data to use from past survey collectors.)"

[12] The specific question is "In order to learn more about your specific background and preferences, we would like to administer a 1 hour follow-up survey to gather more examples about how you would answer questions. Are you

were not interested in participating in a follow-up survey at their current survey rate, compared to 17% of control group participants.

In sum, workers who are more aware of how their data may be used choose to forgo meaningful payments for both their past and future data. This withholding—whether by limiting access to past data or raising demands around future data supply (and to say nothing of potential selection biases in the data itself)—raises the costs of gathering data that accurately reflect consumer preferences, reducing AI productivity (in this case, the usefulness of a synthetic panel).

4 Theory

4.1 Overview

We develop a model of AI expropriation to understand how worker awareness of surveillance affects knowledge sharing and welfare outcomes.

Three key features differentiate our setting from a standard employment relationship:

1. **Two periods**: today's data train tomorrow's AI, creating dynamic effects of knowledge contribution.

2. **Non-contractible knowledge contributions**: the firm does not yet know what skilled workers do until it surveils them and codifies their tacit knowledge, so it cannot write contracts specifying exactly what knowledge workers should provide.

3. **Limited commitment**: workers' expertise walks with them in the status quo—firms cannot guarantee lifetime employment, and workers cannot commit themselves indefinitely. Everything therefore happens under the shadow of renegotiation.

Our model formalizes a dynamic hold-up problem created by surveillance-enabled AI. In period 1, workers reveal tacit know-how while doing their jobs. These records train an AI system that becomes the firm's period-2 outside option. When future bargaining happens under limited commitment, a stronger outside option allows the firm to push down wages. Anticipating this, workers strategically withhold knowledge today, even though withholding is privately costly and reduces current output.

interested in participating in this survey at your current hourly wage? If so, we include you in any follow up surveys using a separate link."

4.2 Model Setup

4.2.1 Players and Timing

The economy consists of one firm and a finite set of workers J. Time is discrete with two periods $t \in \{1, 2\}$, representing the present and the future.

Within each period

1. The firm bargains with each worker j over wage $w_t^j \geqslant 0$ and employment status $I_t^j \in \{0, 1\}$.

2. Each employed worker chooses a non-negative knowledge contribution $k_t^j \in K^j$, where K^j is a finite set.

3. Production occurs and wages are paid.

Between periods, everyone observes the vector of first-period contributions $k_1 \equiv (k_1^1, k_1^2, \ldots, k_1^{|J|})$.

We normalize each worker's output to be equal to their knowledge contribution k_t^j.[13]

We denote the vector of time-t contributions $k_t \equiv (k_t^j)_{j \in J}$, and similarly the vector of time-t wages $w_t \equiv (w_t^j)_{j \in J}$. Given some dataset $D \subseteq J$, we use k_t^D to denote the vector that is identical to k_t except that elements corresponding to $j \notin D$ are equal to 0.

We model AI quality using a function $\alpha : \mathbb{R}^J_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$. Given some time-1 contributions k_1 and some dataset D, the firm develops an AI system of quality $\alpha(k_1^D)$. Intuitively, this captures the productivity of an AI system that can augment unskilled workers or even replace workers entirely.

At time 2, for each worker j, the firm with dataset D gets output

$$\max \left\{ I_2^j k_2^j, \alpha(k_1^D) \right\}. \tag{1}$$

The no-AI case corresponds to $D = \varnothing$. The functional form of (1) means that the AI automates the role, instead of augmenting the worker, because the gains from hiring worker j arise only when worker j's output exceeds what the AI could separately achieve.

We assume that α is monotone increasing in knowledge inputs: that is, for all $\underline{k}_1 \leqslant \overline{k}_1$, we have $\alpha(\underline{k}_1) \leqslant \alpha(\overline{k}_1)$. We assume that knowledge contributions are substitute inputs; that is, for all $\underline{k}_1^j \leqslant \overline{k}_1^j$ and all $\underline{k}_1^{-j} \leqslant \overline{k}_1^{-j}$, we have

$$\alpha(\overline{k}_1^j, \underline{k}_1^{-j}) - \alpha(\underline{k}_1^j, \underline{k}_1^{-j}) \geqslant \alpha(\overline{k}_1^j, \overline{k}_1^{-j}) - \alpha(\underline{k}_1^j, \overline{k}_1^{-j}). \tag{2}$$

[13] Our other assumptions do not rely on the cardinal properties of k_t^j, so if output is some continuous increasing function of k_t^j, we could rescale it so that each k_t^j is equal to the resulting output level.

Here are some examples of α that are permitted by our assumptions:

1. Linear returns $\alpha(k_1) = \sum_j \beta_j k_1^j$ for constants $\beta_j \geqslant 0$,
2. Replicating the skills of the best worker $\alpha(k_1) = \beta \max_j k_1^j$ for constant $\beta \geqslant 0$,
3. Decreasing returns to the sum of contributions $\alpha(k_1) = h\big(\sum_j k_1^j\big)$ where $h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is increasing and concave.

4.2.2 Worker Characteristics

When worker j contributes knowledge k_t^j, they incur private cost $c^j(k_t^j)$. We assume that the cost function $c^j : K^j \to \mathbb{R}_{\geqslant 0}$ has a unique minimizer $\theta^j > 0$, which attains $c^j(\theta^j) = 0$. We interpret θ^j as the worker's skill; this is the level of knowledge that they would naturally reveal when monitored, absent deliberate efforts to withhold knowledge or to record it. In general, the worker may withhold knowledge at a cost ($k_t^j < \theta^j$) or contribute more at a cost ($k_t^j > \theta^j$).

To ease exposition, we assume for now that the worker may only withhold knowledge; that is, $\theta^j = \max K^j$. Under this assumption, the costs capture strategic behavior such as sandbagging, evading surveillance, contaminating data, or degrading documentation. We will later examine how the results change with costly contributions.

4.2.3 Payoffs

We assume that workers' payoffs are additive across time, linear in wage and withholding costs, and that their outside option yields zero payoff. Thus, worker j's utility is

$$U^j = I_1^j \left(w_1^j - c^j(k_1^j)\right) + \psi\, I_2^j \left(w_2^j - c^j(k_2^j)\right), \tag{3}$$

where $\psi > 0$ is the weight on the future. We allow for the case that $\psi > 1$; one can interpret this as meaning that the stakes from expropriation are so large as to outweigh time discounting.

We assume that the firm's payoff is additive across time and across workers, and linear in output and wages. That is, the firm's utility is

$$\Pi = \sum_{j \in J} \left[I_1^j \left(k_1^j - w_1^j\right) + \psi \left(\max\left\{I_2^j k_2^j, \alpha(k_1^D)\right\} - I_2^j w_2^j\right)\right]. \tag{4}$$

4.2.4 Solution concept

Observe that upon being hired at time 2, the unique best response for the worker is to set $k_2^j = \theta^j$, at cost $c^j(\theta^j) = 0$, and we will assume that they so do. For simplicity, we will break ties in firm-worker bargaining in favor of hiring.

An **assessment** in our model is a tuple consisting of:

1. Time-1 wages w_1.

2. Time-1 employment $\overline{J}_1$.

3. Time-1 knowledge contributions $k_1 \in \prod_{j \in J} K^j$.

4. Time-2 wages $\omega_2 : \prod_{j \in J} K^j \to \mathbb{R}^J_{\geqslant 0}$.

5. Time-2 employment $\mathcal{J}_2 : \prod_{j \in J} K^j \to 2^J$.

Note that we have specified contributions k_1^j for workers $j \notin \overline{J}_1$; these are the contributions those workers would make if (counterfactually) they were hired.

We assume that in each period, firms and workers engage in Nash-in-Nash bargaining, with exogenous bargaining weight $\gamma \in [0, 1]$ for each worker. In practice, many workplaces bargain bilaterally at the worker level (or via managers), while firm profits depend on the entire workforce. Nash-in-Nash is a tractable way to capture simultaneous bilateral bargaining while accounting for cross-worker spillovers from AI. The exogenous weight γ should be read as reduced-form bargaining strength, reflecting regulations or labor market conditions.

An assessment is an **equilibrium** if:

1. Time-1 wages w_1 and employment $\overline{J}_1$ are a Nash-in-Nash equilibrium, with the disagreement point for the worker j consisting of not being hired and not being paid.

2. For each worker $j \in J$, their contribution k_1^j maximizes their continuation payoff when the other workers contribute according to $k_1^{J_1}$.[14]

3. For each $k_1' \in \prod_{j \in J} K^j$, wages $\omega_2(k_1')$ and employment $\mathcal{J}_2(k_1')$ are a Nash-in-Nash equilibrium.

Our models differ in the firm's dataset in the event of agreement and disagreement; we describe each in the subsections that follow.

[14] Implicitly, this means that hired workers do not observe who else is hired when deciding their own contribution; so they best-respond to the *equilibrium* hired set $\overline{J}_1$. Notice also that workers not hired at time-1 have 'passive beliefs'. That is, if (off-path) they are hired, they believe the set of other workers hired is the same.

4.3 Results

4.3.1 No surveillance

Without surveillance, the firm's dataset is always $D = \varnothing$, and the disagreement point has worker j not hired and not paid.

Our model is designed to isolate the dynamic effects of expropriation by AI. Without surveillance, firms and workers are engaging in two identical and separable production stages. Knowledge contributed at time 1 has no effect on the worker's time-2 payoffs, so workers have no incentive to withhold knowledge. We state this formally in the following observation.

Observation 4.1. With no surveillance, every equilibrium has the following features:

1. full employment in both periods,
2. each worker contributes knowledge equal to their skill in both periods, $k_1^j = k_2^j = \theta^j$, and
3. wages $w_t^j = \gamma\theta^j$, where γ is the worker's Nash bargaining weight.

This equilibrium is a useful benchmark, because workers withhold no knowledge and incur no withholding costs. This equilibrium only falls short of the first-best because it achieves none of the potential productivity gains from AI.

Even under the alternative policies that follow, there will be equilibria with full employment in both periods, because the worker's outside option yields zero payoff. For ease of exposition, we will focus on these equilibria, so that the key policy-relevant comparisons we focus on are: total surplus, worker wages, and firm profits.

4.3.2 Firm-owned AI

With firm-owned AI, the firm's dataset is always $\overline{J}_1$, and the disagreement point has worker j not hired and not paid. Thus, even if no agreement is reached with worker j at time-2, the firm still produces $\alpha(k_1)$ using last period's data.

Suppose the workers naïvely contributed full knowledge at time-1. Holding fixed the workers' time-1 contributions k_1, firm-owned AI raises the output resulting from agreement between the firm and worker j at time-2, from θ^j to $\max\left\{\theta^j, \alpha\left(k_1^{\overline{J}_1}\right)\right\}$. But it also raises the firm's disagreement payoff, from 0 to $\alpha\left(k_1^{\overline{J}_1}\right)$. Thus, Nash-in-Nash bargaining results in a fall in the worker's wage, from

$\gamma\theta^j$ to $\gamma\left\lfloor\theta^j-\alpha\left(k_1^{\overline{J}_1},0\right)\right\rfloor_+$, where we use the notation $\lfloor x\rfloor_+$ to denote $\max\{x,0\}$. Thus, ignoring equilibrium effects, firm-owned AI increases time-2 output and the firm's time-2 profits.

Next we study what happens in equilibrium with firm-owned AI. The time-1 contribution game has a useful structure: Even though workers are harmed by raising k_1^j (it strengthens the AI they face tomorrow), the substitutes property implies the marginal harm from revealing more is smaller when others already revealed a lot. Thus, workers' time-1 contributions are strategic complements. Intuitively, if your colleagues have already "taught the AI most of what it needs," your own contribution does little incremental damage to your future wage but still saves you withholding effort $c^j(\cdot)$, so best responses are weakly increasing in others' contributions.

To ensure that our results are not vacuous, we start with a simple sufficient condition for the existence of equilibrium. We require that (essentially) withholding costs are not so severe that they outweigh a single-worker's time-1 output. Formally, let B^j be the best-response contributions of worker j when all other workers contribute 0, that is

$$B^j \equiv \arg\max_{k_1^j}\left\{w_1^j - c^j\left(k_1^j\right) + \psi\gamma\left\lfloor\theta^j-\alpha\left(k_1^j,0\right)\right\rfloor_+\right\}. \tag{5}$$

We assume that

$$\inf\left\{k_1^j - c^j\left(k_1^j\right) : k_1^j \in B^j\right\} \geqslant 0. \tag{6}$$

We now state a result that guarantees the existence of full-employment equilibrium, and implies a (weak) reduction in time-2 wages compared to the no-surveillance case.

Theorem 4.2. With firm-owned AI, under assumption (6), there exists an equilibrium with:

1. Full employment in both periods $J=\overline{J}_1=\mathcal{J}_2(\tilde{k}_1)$ for all $\tilde{k}_1$,

2. Time-2 wages $\omega_2^j(\tilde{k}_1)=\gamma\left\lfloor\theta_j-\alpha(\tilde{k}_1)\right\rfloor_+$ for all $\tilde{k}_1$ and all j.

Proof. We have restricted attention to equilibria with full time-2 knowledge contributions, and in every such equilibrium, Nash-in-Nash bargaining at time 2 implies that $\omega_2^j(\tilde{k}_1)=\gamma\left\lfloor\theta_j-\alpha(\tilde{k}_1)\right\rfloor_+$ for all $\tilde{k}_1$ and all j.

Given some time-1 hired set $\overline{J}_1$ and wages w_1, it follows that worker j chooses k_1^j to maximize the utility function

$$w_1^j - c^j(k_1^j) + \psi\gamma\left\lfloor\theta_j-\alpha\left(k_1^{\overline{J}_1\cup\{j\}}\right)\right\rfloor_+. \tag{7}$$

In order to show existence, we will establish that the simultaneous choice of k_1^j by the workers to maximize (7) is a supermodular game, in the sense of Milgrom and Roberts (1990). To do so, we first prove a technical lemma.

Lemma 4.3. Let X and Y be partially ordered sets. Suppose $f : X \times Y \to \mathbb{R}$ is monotone non-increasing[15] and has increasing differences. Suppose $g : \mathbb{R} \to \mathbb{R}$ is non-decreasing and convex. Then $g \cdot f : X \times Y \to \mathbb{R}$ has increasing differences.

We now prove Lemma 4.3. The function f has increasing differences, so for all $\underline{x} \leqslant \overline{x}$ and $\underline{y} \leqslant \overline{y}$, we have

$$f(\underline{x}, \overline{y}) - f(\overline{x}, \overline{y}) \leqslant f(\underline{x}, \underline{y}) - f(\overline{x}, \underline{y}). \tag{8}$$

By f monotone non-increasing, we have

$$\Phi \equiv f(\underline{x}, \overline{y}) - f(\overline{x}, \overline{y}) \geqslant 0, \tag{9}$$

$$f(\overline{x}, \overline{y}) \leqslant f(\overline{x}, \underline{y}). \tag{10}$$

It follows that

$$\begin{aligned} g(f(\underline{x}, \overline{y})) - g(f(\overline{x}, \overline{y})) &= g(f(\overline{x}, \overline{y}) + \Phi) - g(f(\overline{x}, \overline{y})) \\ &\leqslant g(f(\overline{x}, \underline{y}) + \Phi) - g(f(\overline{x}, \underline{y})) \text{ by (9), (10) and } g \text{ convex} \\ &\leqslant g(f(\underline{x}, \underline{y})) - g(f(\overline{x}, \underline{y})) \text{ by (8) and } g \text{ non-decreasing.} \end{aligned} \tag{11}$$

Thus, $g \cdot f$ has increasing differences. This completes the proof of Lemma 4.3.

Lemma 4.4. The simultaneous choice of k_1^j by the workers to maximize (7) is a supermodular game.

Most of the requirements for a supermodular game follow by inspection. The only non-trivial part is to show that for each worker j, the utility function

$$w_1^j - c^j(k_1^j) + \psi\gamma \max\left[\theta^j - \alpha\left(k_1^{\overline{J}_1 \cup \{j\}}\right)\right]_+ \tag{12}$$

[15] That is, for $\underline{x} \leqslant \overline{x}$ and $\underline{y} \leqslant \overline{y}$, we have $f(\underline{x}, \underline{y}) \geqslant f(\overline{x}, \overline{y})$.

has increasing differences in (k_1^j, k_1^{-j}). Observe that $\theta^j - \alpha\left(k_1^{\overline{J}_1 \cup \{j\}}\right)$ has increasing differences in (k_1^j, k_1^{-j}) by the assumption that worker contributions are substitutes, and it is monotone non-increasing. It follows that

$$\psi\gamma \max \left\lfloor \theta^j - \alpha\left(k_1^{\overline{J}_1 \cup \{j\}}\right) \right\rfloor_+ \tag{13}$$

has increasing differences by Lemma 4.3. Moreover, $w_1^j - c^j(k_1^j)$ has increasing differences trivially, because it does not depend on k_1^{-j}. The set of functions with increasing differences is closed under addition, so it follows that (12) has increasing differences in (k_1^j, k_1^{-j}). This completes the proof of Lemma 4.4.

Fixing time-1 wages w_1 and employment $\overline{J}_1$, there exists a contribution profile k_1 that satisfies the requirements of equilibrium, by Lemma 4.4 and Theorem 5 of Milgrom and Roberts (1990).

We now guess and verify that full employment at both periods, that is, $J = \overline{J}_1 = \mathcal{J}_2(\tilde{k}_1)$ for all $\tilde{k}_1$, is part of an equilibrium. This follows straightforwardly for time 2 because ω_2^j derived above gives the firm a non-negative payoff from hiring each worker.

We now consider time 1. Hiring worker j at wage w_1^j results in firm payoff

$$k_1^j - w_1^j + \sum_{l \neq j}(k_1^l - w_1^l) + \psi \sum_l \left(\alpha(k_1^J) + (1-\gamma)\left\lfloor \theta^l - \alpha(k_1^J) \right\rfloor_+ \right) \tag{14}$$

and worker payoff

$$w_1^j - c^j(k_1^j) + \psi\gamma \left\lfloor \theta^l - \alpha(k_1^J) \right\rfloor_+ . \tag{15}$$

Not hiring worker j results in firm payoff

$$\sum_{l \neq j}(k_1^l - w_1^l) + \psi \sum_l \left(\alpha(k_1^{J\setminus\{j\}}) + (1-\gamma)\left\lfloor \theta^l - \alpha(k_1^{J\setminus\{j\}}) \right\rfloor_+ \right), \tag{16}$$

and worker payoff

$$\psi\gamma \left\lfloor \theta^j - \alpha(k_1^{J\setminus\{j\}}) \right\rfloor_+ \tag{17}$$

Thus, compared to the disagreement point, hiring worker j at time 1 increases the pairwise surplus by

$$
\begin{aligned}
&k_1^j - c^j(k_1^j) + \psi \left(\max\left\{\theta^j, \alpha(k_1^J)\right\} - \max\left\{\theta^j, \alpha(k_1^{J\setminus\{j\}})\right\} \right) \\
&\qquad + \psi \sum_{l \neq j} \left(\alpha(k_1^J) + (1-\gamma) \left\lfloor \theta^l - \alpha(k_1^J) \right\rfloor_+ - \alpha(k_1^{J\setminus\{j\}}) - (1-\gamma) \left\lfloor \theta^l - \alpha(k_1^{J\setminus\{j\}}) \right\rfloor_+ \right). \quad (18)
\end{aligned}
$$

Next we show that

$$
k_1^j - c^j(k_1^j) \geqslant 0. \tag{19}
$$

By hypothesis, the contribution profile k_1 is a pure-strategy Nash equilibrium of the supermodular game considered in Lemma 4.4. It follows that

$$
k_1^j \in \arg\max_{\hat{k}_1^j} \left\{ w_1^j - c^j\left(\hat{k}_1^j\right) + \psi\gamma \left\lfloor \theta^j - \alpha\left(\hat{k}_1^j, k_1^{-j}\right) \right\rfloor_+ \right\}. \tag{20}
$$

Let $\underline{k}_1^j$ be the smallest element of the right-hand side of (5). By Topkis' theorem (Milgrom and Shannon, 1994), we have that $\underline{k}_1^j \leqslant k_1^j$. Then by (20) and α non-decreasing we have that $c^j(\underline{k}_1^j) \geqslant c^j(k_1^j)$. By the preceding inequalities and (6), we have

$$
0 \leqslant \inf\left\{\hat{k}_1^j - c^j\left(\hat{k}_1^j\right) : \hat{k}_1^j \in B^j\right\} \leqslant \underline{k}_1^j - c^j\left(\underline{k}_1^j\right) \leqslant k_1^j - c^j\left(k_1^j\right). \tag{21}
$$

We have proved (19). By (19) and α monotone non-decreasing, it follows that (18) is non-negative. We have proved that full employment at time 1 is consistent with Nash-in-Nash bargaining, which completes the proof. □

We henceforth restrict attention to equilibria of the form characterized in Theorem 4.2.

Observation 4.5. With firm-owned AI, each worker's time-2 wage is lower than the no-surveillance baseline, and strictly lower if the time-1 knowledge contributions lead to positive AI quality, that is if $\alpha(k_1) > 0$.

Observation 4.6. With firm-owned AI, workers withhold knowledge at time 1 compared to the no-surveillance baseline, that is for every worker j, we have $k_1^j \leqslant \theta^j$.

Proof. With no surveillance, each worker's time-1 knowledge contribution maximizes $k_1^j - c^j(k_1^j)$, which has a unique minimizer $k_1^j = \theta^j$. With firm-owned AI, each worker's time-1 knowledge contribution maximizes (12). The observation follows by Topkis's theorem. □

Observation 4.7. In equilibrium, knowledge withholding increases worker surplus and decreases firm surplus, compared to what would happen if each worker naïvely contributed $k_1^j = \theta^j$.

Proof. Let us define

$$u^j(k_1) \equiv w_1^j - c^j(k_1^j) + \psi\gamma \left\lfloor \theta_j - \alpha\left(k_1\right)\right\rfloor_+ . \tag{22}$$

Let k_1 denote an equilibrium time-1 contribution profile. We have

$$u^j(k_1^{-j}, k_1^{-j}) \geqslant u^j(\theta^j, k_1^{-j}) \geqslant u^j(\theta^j, \theta_1^{-j}), \tag{23}$$

where the first inequality is by worker best-response, and the second inequality is by u^j decreasing in the other workers' time-1 contributions. We have proved that knowledge withholding increases worker surplus. Since withholding is costly and reduces time-2 output, it reduces total surplus. It follows that it reduces firm surplus. □

Under what conditions do workers withhold more knowledge in equilibrium? We now state some comparative statics results, restricting attention to equilibria of the form characterized by Theorem 4.2.

Our assumptions allow for the possibility of multiple equilibria, which can complicate comparative statics. But the workers' time-1 contributions are strategic complements, by Lemma 4.4. Thus, holding fixed the primitives, there exists a highest equilibrium, in the sense that each worker contributes weakly more at time 1 than they do in any other equilibrium (Milgrom and Roberts, 1990, Theorem 5). And there also exists a lowest equilibrium, in the sense that each worker contributes weakly less than they do in any other equilibrium. We will keep track of the highest and lowest equilibria as the primitives change.

First, we show that workers withhold less (contribute more) when withholding has a higher marginal cost.

Theorem 4.8. Consider two profiles of cost functions, $(c^j)_{j\in J}$ and $(\tilde{c}^j)_{j\in J}$, where for all workers j and contributions $k_1^j \leqslant k_1^{j'}$, we have

$$c^j(k_1^j) - c^j(k_1^{j'}) \leqslant \tilde{c}^j(k_1^j) - \tilde{c}^j(k_1^{j'}). \tag{24}$$

Holding all other primitives fixed, let $\overline{k}_1$ and $\underline{k}_1$ denote the highest and lowest equilibrium time-1 contributions under $(c^j)_{j\in J}$, and similarly let $\overline{\tilde{k}}_1$ and $\underline{\tilde{k}}_1$ denote the highest and lowest equilibrium time-1 contributions under $(\tilde{c}^j)_{j\in J}$. We have $\overline{k}_1 \leqslant \overline{\tilde{k}}_1$ and $\underline{k}_1 \leqslant \underline{\tilde{k}}_1$.

Proof. By Lemma 4.4, the worker payoffs (7) from time-1 contributions induced by $(c^j)_{j\in J}$ and $(\tilde{c}^j)_{j\in J}$ define a pair of supermodular games indexed by parameter τ, with $\tau = 0$ corresponding to $(c^j)_{j\in J}$ and $\tau = 1$ corresponding to $(\tilde{c}^j)_{j\in J}$. Each worker j's utility has increasing differences in (k_1^j, τ) by (24). Thus, the result follows by Theorem 6 of Milgrom and Roberts (1990). $\square$

Next we show that equilibrium contributions rise when workers are better substitutes for each other. To state this formally, we slightly extend the model: Suppose that each worker occupies a distinct role, and let the AI quality in role j be denote $\alpha^j(k_1)$. All the results stated so far extend to this case, with essentially the same proofs.

Suppose we found a better way to use other workers' data to automate worker j's role. Intuitively, this would raise the AI quality in worker j's role for a given profile of contributions. And it would lower the marginal returns to AI quality from raising worker j's contribution. We now formally show that such a change increases equilibrium contributions.

Given two AI technologies, $\left(\alpha^j\right)_{j\in J}$ and $\left(\tilde{\alpha}^j\right)_{j\in J}$, we write that $\left(\tilde{\alpha}^j\right)_{j\in J}$ is **at least as good as** $\left(\alpha^j\right)_{j\in J}$ if for each role j, we have

$$\alpha^j(k_1) \leqslant \tilde{\alpha}^j(k_1) \text{ for all } k_1. \tag{25}$$

Given two AI technologies, $\left(\alpha^j\right)_{j\in J}$ and $\left(\tilde{\alpha}^j\right)_{j\in J}$, we write that $\left(\tilde{\alpha}^j\right)_{j\in J}$ has **stronger substitution than** $\left(\alpha^j\right)_{j\in J}$ if for each role j, we have

$$\alpha^j(\hat{k}_1^j, k_1^{-j}) - \alpha^j(k_1^j, k_1^{-j}) \geqslant \tilde{\alpha}^j(\hat{k}_1^j, k_1^{-j}) - \tilde{\alpha}^j(k_1^j, k_1^{-j}) \text{ for all } k_1^j \leqslant \hat{k}_1^j \text{ and all } k_1^{-j}. \tag{26}$$

Theorem 4.9. Suppose that $\left(\tilde{\alpha}^j\right)_{j\in J}$ is at least as good as $\left(\alpha^j\right)_{j\in J}$ and that $\left(\tilde{\alpha}^j\right)_{j\in J}$ has stronger substitution than $\left(\alpha^j\right)_{j\in J}$. Holding all other primitives fixed, let $\overline{k}_1$ and $\underline{k}_1$ denote the highest and lowest equilibrium time-1 contributions under $\left(\alpha^j\right)_{j\in J}$, and similarly let $\overline{\tilde{k}}_1$ and $\underline{\tilde{k}}_1$ denote the highest and lowest equilibrium time-1 contributions under $\left(\tilde{\alpha}^j\right)_{j\in J}$. We have $\overline{k}_1 \leqslant \overline{\tilde{k}}_1$ and $\underline{k}_1 \leqslant \underline{\tilde{k}}_1$.

Moreover, worker wages at time-2 are lower under technology α and contribution profiles $\overline{\tilde{k}}_1$ and $\underline{\tilde{k}}_1$, compared to technology $\tilde{\alpha}$ and contribution profiles $\overline{k}_1$ and $\underline{k}_1$ respectively.

The proof is in Appendix C.1.

4.4 Alternative Policies

4.4.1 Individual data ownership

Under individual-owned AI, the firm's dataset D is equal to the set of workers hired at *both* time-1 and time-2. Thus, the disagreement point for bargaining at time-2 between the firm and worker j involves the worker not being hired, not being paid, and their data being omitted from the dataset.

Since we are breaking ties in favor of employment and the pairwise surplus from employing an additional worker is at least zero, every equilibrium with individual-owned AI involves full employment at time 2. In such an equilibrium, the pairwise surplus from employing worker j at time-2 is

$$\lambda_j(k_1, \overline{J}_1) \equiv \max\left\{\theta^j, \alpha\left(k_1^{\overline{J}_1}\right)\right\} - \alpha\left(k_1^{\overline{J}_1\setminus\{j\}}\right) + \sum_{l\neq j}\left(\max\left\{\theta^l, \alpha\left(k_1^{\overline{J}_1}\right)\right\} - \max\left\{\theta^l, \alpha\left(k_1^{\overline{J}_1\setminus\{j\}}\right)\right\}\right) \tag{27}$$

Observe that $\lambda_j(k_1, \overline{J}_1)$ is non-decreasing in k_1^j, by α monotone non-decreasing. By Nash-in-Nash bargaining, worker j's equilibrium wage is $\gamma\lambda_j(k_1, \overline{J}_1)$. At time 1, worker j chooses k_1^j to maximize

$$w_1^j - c^j(k_1^j) + \psi\gamma\lambda_j(k_1, \overline{J}_1), \tag{28}$$

and by c^j non-increasing and λ_j non-decreasing in k_1^j, it follows that it is a best response to contribute $k_1^j = \theta^j$. Thus, the pairwise surplus from employing any worker j at time 1 is at least zero, so there is an equilibrium with full employment at time 1.

Observation 4.10. Under individual-owned AI, there is an equilibrium with full employment in both periods and with $k_1^j = k_2^j = \theta^j$.

Henceforth we restrict attention to the equilibrium described in Observation 4.10. Notice that, compared to firm-owned AI, workers contribute weakly more knowledge at time 1, which raises output directly at time 1 directly and at time 2 indirectly, by improving AI quality.

Observation 4.11. Total surplus is higher under individual ownership than under firm ownership.

However, individual ownership does not guarantee that workers are better off, compared to the no-AI baseline. The intuition for this is that Nash-in-Nash bargaining compensates workers based on the marginal contribution of their data, and under substitutes the total contribution exceeds the sum of the marginal contributions. That is,

$$\alpha(k_1^J) - \alpha(0) \geqslant \sum_j \left(\alpha(k_1^J) - \alpha(k_1^{J\setminus\{j\}})\right). \tag{29}$$

To see this, observe that we can number the workers arbitrarily from 1 to $|J|$, and rewrite the left-hand side of (29) as a telescoping sum

$$\alpha(k_1^J) - \alpha(0) = \sum_{l=1}^{|J|} \left(\alpha(k_1^{\{j:j\leqslant l\}}) - \alpha(k_1^{\{j:j\leqslant l-1\}})\right) \geqslant \sum_j \left(\alpha(k_1^J) - \alpha(k_1^{J\setminus\{j\}})\right), \tag{30}$$

where the inequality follows by the substitutes assumption.

Individual ownership generates a competition externality: Each worker accounts for the (non-negative) effect of their time-1 contribution on their time-2 wage, but does not account for how their time-1 contribution decreases other workers' time-2 wage. Thus, individual ownership does not ensure that workers receive a substantial share of the rents from AI. This problem is worse when workers' data are more substitutable for each other, as we now state formally.

Theorem 4.12. Suppose that $\left(\tilde{\alpha}^j\right)_{j\in J}$ has stronger substitution than $\left(\alpha^j\right)_{j\in J}$ and also that

$$\tilde{\alpha}^j(\theta) = \alpha^j(\theta) \text{ for all } j. \tag{31}$$

Then each worker's equilibrium payoff is lower under $\left(\tilde{\alpha}^j\right)_{j\in J}$ than under $\left(\alpha^j\right)_{j\in J}$.

The proof is in Appendix C.2.

In some cases, individual ownership can harm workers compared to the no-AI baseline, even when workers have full Nash bargaining power. We now state this formally.

Theorem 4.13. Suppose that:

1. There are at least three workers,
2. workers have identical maximum contributions, with $\theta^j = \theta^{j'}$ for all j and j',
3. contributions are perfect substitutes, that is $\alpha(k_1) = f(\max_j\{k_1^j\})$ for some increasing function f.

Then for any Nash bargaining parameter $\gamma \in (0,1]$, the full-contribution full-employment equilibrium under individual data ownership is strictly worse for each worker than the full-contribution full-employment equilibrium under no surveillance.

The proof is in Appendix C.3.

To summarize, individual data ownership restores efficiency, because it ensures that each worker has no incentive to withhold knowledge at time-1. But individual data ownership does not guarantee that workers share in the efficiency gains from AI. By Theorem 4.13, under some conditions workers can be strictly worse off under individual data ownership, compared to the no-surveillance case. Under those conditions, firms benefit from AI not only because it raises time-2 output, but also because it suppresses workers' time-2 wages.

4.4.2 Collective data ownership

So far we have considered workers bargaining individually with the firm, via Nash-in-Nash bargaining. Suppose instead that the workers bargain as a single union with the firm. The union has the same Nash bargaining weight γ, and a utility equal to the sum of the individual worker utilities. Suppose furthermore that in the event of disagreement at time 2, no workers are employed and the firms dataset is $D = \varnothing$. That is, we will call an assessment an **equilibrium with collective ownership** if:

1. Time-1 wages w_1 and employment $\overline{J}_1$ are a Nash bargaining solution between the firm and the union.

2. For each worker $j \in J$, their contribution k_1^j maximizes their continuation payoff when the other workers contribute according to $k_1^{J_1}$.

3. For each $k_1' \in \prod_{j \in J} K^j$, wages $\omega_2(k_1')$ and employment $\mathcal{J}_2(k_1')$ are a Nash bargaining solution between the firm and the union.

Observe that so long as worker j's wage $\omega_2^j(k_1)$ is non-decreasing in k_1^j, it is a best response for worker j to contribute $k_1^j = \theta^j$ at time 1. One example of such a scheme is to pay each worker an equal share $\gamma/|J|$ of output in each period.

Observation 4.14. There is an equilibrium with collective ownership with full employment in both periods and full knowledge contributions $k_t^j = \theta^j$ in both periods.

Under collective ownership, the time-2 disagreement point excludes *all* workers' data. This bundles individual data rights so that when one worker contributes, they do not inadvertently strengthen the firm's hand against others. This prevents competition externalities from arising, ensuring that workers share in the gains from AI. Next, we state that the benefits to workers of collective ownership are larger when their data are better substitutes for each other.

Observation 4.15. Suppose that $\left(\tilde{\alpha}^j\right)_{j\in J}$ has stronger substitution than $\left(\alpha^j\right)_{j\in J}$ and also that $\tilde{\alpha}^j(\theta) = \alpha^j(\theta)$ for all j. Then the gains to worker surplus of switching from individual ownership to collective ownership are larger under $\left(\tilde{\alpha}^j\right)_{j\in J}$ than under $\left(\alpha^j\right)_{j\in J}$.

Proof. Observe that under collective ownership, worker surplus under $\left(\alpha^j\right)_{j\in J}$ is

$$\gamma \sum_j \left(\theta^j + \psi \max\left\{\theta^j, \alpha^j(\theta)\right\}\right), \tag{32}$$

and worker surplus under $\left(\tilde{\alpha}^j\right)_{j\in J}$ is

$$\gamma \sum_j \left(\theta^j + \psi \max\left\{\theta^j, \tilde{\alpha}^j(\theta)\right\}\right). \tag{33}$$

By $\tilde{\alpha}^j(\theta) = \alpha^j(\theta)$, these are equal. The observation then follows by Theorem 4.12. □

4.5 Extension to costly contributions

So far we have assumed that the worker may only withhold knowledge at a cost; that is, $\theta^j = \max K^j$. We now consider the more general model, that permits $\theta^j < \max K^j$. Recall that θ^j is defined as the unique minimizer of the worker's cost function, which attains $c^j(\theta^k) = 0$. Moreover, this is the worker's equilibrium contribution in the no-surveillance baseline. Thus, by permitting $\theta^j < \max K^j$, we are positing that the worker may contribute strictly more knowledge than that baseline, at some personal cost.

Our analysis of the no-surveillance baseline and of firm-owned AI extends without modification to the case with costly contributions. The results are true as stated, and the proofs do not rely on the assumption that $\theta^j = \max K^j$.[16]

Under individual data ownership, allowing costly contributions changes the analysis technically and substantively. When $\theta^j = \max K^j$, the worker's time-1 best response is to contribute $k_1^j = \theta^j$

[16] That is, the following results extend with no modification: Observation 4.1, Theorem 4.2, Observation 4.5, Observation 4.6, Observation 4.7, Theorem 4.8, and Theorem 4.9.

regardless of the other workers' contributions, because their time-2 wage is weakly increasing in k_1^j and their time-1 costs are strictly decreasing in k_1^j. By contrast, if $\theta^j < \max K^j$, then the worker's time-1 best response may depend on the other workers' contributions.

Consequently, the technical change to the analysis is that there may not be an equilibrium in pure strategies for the workers.[17] Nonetheless, if we extend the definition of equilibrium to allow workers to mix over time-1 knowledge contributions, then an equilibrium exists. Since each worker's set of possible contributions K^j is finite, and the continuation payoffs resulting from each contribution are well-defined (as in the proof of Theorem 4.2), this follows by the existence of Nash equilibrium for finite games. Under a technical assumption, such an equilibrium also has full employment in both periods.

Theorem 4.16. Suppose that for every worker j, we have

$$\inf_{k_1^j \in K^j} \left\{ k_1^j - c^j(k_1^j) \right\} \geqslant 0. \tag{34}$$

Under individual-owned AI, there is an equilibrium (possibly with randomization in time-1 contributions) with full employment in both periods.

The proof is in Appendix C.4.

The substantive change to the analysis is that, with costly contributions, equilibrium contributions can be strictly higher than in the no-surveillance baseline. This is because workers are rewarded for their marginal contribution to AI quality, and can increase that contribution at a cost.

Recall that under firm ownership, equilibrium contributions never exceed θ^j (Observation 4.6). The conclusion that individual ownership raises contributions compared to firm ownership holds even under costly contributions, as we now state formally.

Theorem 4.17. Under any equilibrium with individual ownership, if worker j is hired and contributes k_1^j with positive probability, then $k_1^j \geqslant \theta^j$.

The proof is in Appendix C.5.

However, the welfare effects of individual ownership can change under costly contributions; Observation 4.11 does not follow without further assumptions. The key is that under individual ownership, workers may make costly contributions that are individually rational but not socially beneficial, as in the following example.

[17] This complication does not arise under firm ownership of data because worker contributions at time-1 are a supermodular game; see Lemma 4.4.

Example 1. There are two workers, with $K^1 = K^2 = \{5, 6\}$. The AI has quality 3 if one worker chooses contribution 6, quality 4 if both workers choose contribution 6, and 0 otherwise. Contribution 5 is free (so $\theta^1 = \theta^2 = 5$), and contribution 6 costs $\epsilon > 0$. In the unique equilibrium under firm ownership both workers contribute 5 at time 1. For ϵ small enough, there exists an equilibrium under individual ownership in which both workers contribute 6 at time 1. In the latter equilibrium, both workers pay positive costs but total output has not increased, so total surplus is strictly lower than under firm ownership, and also strictly lower than in the no-AI case.

Next we study an equilibrium with collective ownership, assuming for simplicity that the union splits total compensation in each period equally across workers. In general, the comparison of total surplus between individual ownership and collective ownership is ambiguous. Example 1 proves that individual ownership can sometimes result in strictly lower total surplus than the no-AI case. By contrast, collective ownership always raises total surplus compared to the no-AI case, as we now state formally.

Theorem 4.18. Under any equilibrium with collective ownership, expected total surplus is at least as high as in the no-AI case.

The proof is in Appendix C.6.

5 Conclusion

AI models can now perform a wide range of workplace tasks. But in many cases, model performance depends on access to detailed knowledge about how work is actually done, knowledge that has traditionally belonged to workers. While many foundational AI models were trained on historical records collected without workers' awareness or meaningful consent, further progress will depend on new data from workers who are increasingly aware that their actions, communications, and outputs can be used to train systems that replicate their work.

We provide evidence that workers' willingness to supply knowledge depends on their understanding of how their data will be used. This empirical fact motivates our theory, which considers one important reason why workers may want to adjust their knowledge supply: career concerns. We show that, under current institutional arrangements, concerns about expropriation give workers an incentive to hide knowledge. Such behavior can have negative consequences for all parties: workers can suffer lower productivity and wages in the present, firms can face lower profits from reduced ability to make use of labor expertise, and overall economic output can decline due to lower productivity gains from the adoption of effective AI systems. These frictions give workers, employers, and policymakers a shared interest in governance structures that mitigate fears of data-driven expropriation.

Our analysis also reveals a tension between workers' stated preferences and their longer-term welfare. Roughly 65% of respondents favor individual control over their work data, including the right to sell it for AI development. Our theory shows, however, that unilateral sales create a competition externality: a given worker's decision to sell their data improves the firm's outside option against all other workers. Because workers do not internalize this spillover, individual data sales can leave all workers worse off relative to a no-AI benchmark, even when they hold full bargaining power. By contrast, collective data ownership internalizes these externalities and increases worker surplus relative to no AI. Implementing such arrangements, however, raises practical challenges. Differences in the value of workers' contributions may generate intra-union tensions, and the legal basis for worker ownership is difficult to define when data are produced using firm-provided capital and intellectual property.

More broadly, our results suggest that the governance of workplace data will play a central role in shaping the trajectory of AI adoption and worker welfare. If current arrangements persist, rising worker awareness may slow the development of effective AI by inducing meaningful resistance, as evidenced by growing labor disputes around AI in the workplace, or by merely failing to motivate valuable knowledge contributions. Designing, testing, and refining alternative mechanisms, whether through public policies or private contracts, is an important direction for future research.

References

Aitken, H. G. J. (1960). *Taylorism at Watertown Arsenal: Scientific Management in Action, 1908–1915*. Cambridge, MA: Harvard University Press.

Ajunwa, I. (2025). Ai and captured capital. *Yale Law Journal Forum 134*. Published January 31, 2025. https://www.yalelawjournal.org/forum/ai-and-captured-capital.

Allyn-Feuer, A. and T. Sanders (2023). Transformative AGI by 2043 is <1% likely. *arXiv preprint arXiv:2306.02519*.

Brynjolfsson, E., D. Li, and L. Raymond (2025, April). Generative AI at Work. *The Quarterly Journal of Economics 140*(2), 889–942.

Bubeck, S., V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, H. Nori, H. Palangi, M. T. Ribeiro, and Y. Zhang (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. *arXiv preprint arXiv:2303.12712*.

Buckman, S. R., J. M. Barrero, N. Bloom, and S. J. Davis (2025, February). Measuring work from home. Working Paper 33508, National Bureau of Economic Research.

Christopher, N. (2025, November). Inside the race to train AI robots how to act human in the real world. Los Angeles Times. Available at https://www.latimes.com/business/story/2025-11-02/inside-californias-rush-to-gather-human-data-for-building-humanoid-robots. Accessed 2026-03-18.

Diamantis, M. E. (2023). Employed Algorithms: A Labor Model of Corporate Liability for AI. *Duke Law Journal 72*, 797–867.

Glass, A. (2024). Unions Give Workers a Voice Over How AI Affects Their Jobs.

Grace, K., Z. Stein-Perlman, and B. Weinstein-Raun (2022). 2022 Expert Survey on Progress in AI. AI Impacts report (Aug. 2022). Available at https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/.

Henriques, A. (2025). Synthetic panels in market research: What you need to know.

Hertel-Fernandez, A. (2024, October). Estimating the prevalence of automated management and surveillance technologies at work and their impact on workers' well-being. Technical report, Washington Center for Equitable Growth, Washington, DC.

Hu, K. (2025, December). Ai companies need business customers. why that could take a while. *Reuters*.

Jiang, L. Y., X. C. Liu, N. P. Nejatian, M. Nasir-Moin, D. Wang, A. Abidin, K. Eaton, H. A. Riina, I. Laufer, P. Punjabi, M. Miceli, N. C. Kim, C. Orillac, Z. Schnurman, C. Livia, H. Weiss, D. Kurland, S. Neifert, Y. Dastagirzada, D. Kondziolka, A. T. M. Cheung, G. Yang, M. Cao, M. Flores, A. B. Costa, Y. Aphinyanaphongs, K. Cho, and E. K. Oermann (2023). Health system-scale language models are all-purpose prediction engines. *Nature 619*(7969), 357–362.

Kantor, J. and A. Sundaram (2022, August). Workplace productivity: Are you being tracked? The New York Times (Interactive).

Kim, P. T. and R. Leavitt (2026). Data Rights for Workers. *Boston University Law Review*. Forthcoming; Wash. U. Legal Studies Research Paper 25-04-01.

Milanez, A., A. Lemmens, and C. Ruggiu (2025, February). Algorithmic management in the workplace: New evidence from an oecd employer survey. Technical Report 31, OECD, Paris.

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 1255–1277.

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica: Journal of the Econometric Society*, 157–180.

Mitchell, M. (2021). Why AI is harder than we think. *arXiv preprint arXiv:2104.12871*.

Montgomery, D. (1979). *Workers' Control in America: Studies in the History of Work, Technology, and Labor Struggles*. Cambridge: Cambridge University Press.

Murty, R. N. and R. Kumar S (2026, February). When every company can use the same ai models, context becomes a competitive advantage. *Harvard Business Review*.

Polanyi, M. (1967). *The tacit dimension* (Anchor Books ed. ed.). Doubleday Anchor book, A540. Garden City, N.Y.: Anchor Books.

Rabbit Inc. (2024). How Rabbit R1 works: Teach AI to do your tasks. https://www.rabbit.tech. Accessed: 2025-06-27.

Ramani, A. and Z. Wang (2023). Why transformative artificial intelligence is really, really hard to achieve.

SAG-AFTRA (2023). SAG-AFTRA: TV / Theatric Contracts 2023.

Spirlet, T. (2026, January). AI is turning expertise into a commodity. Box CEO says there's one way companies can stay ahead. Business Insider. Available at: https://www.businessinsider.com/ai-agents-expertise-box-ceo-context-gives-companies-competitive-advantage-2026-1. Accessed 2026-03-12.

Taylor, F. W. (1911). *The Principles of Scientific Management*. New York: Harper & Brothers.

Turner, J. (2022, June). The right way to monitor your employee productivity. Gartner.

U.S. Congress (1909). Copyright act of 1909, ch. 320, § 23, 35 stat. 1075, 1080 (repealed 1976). Public Law. Enacted March 4, 1909.

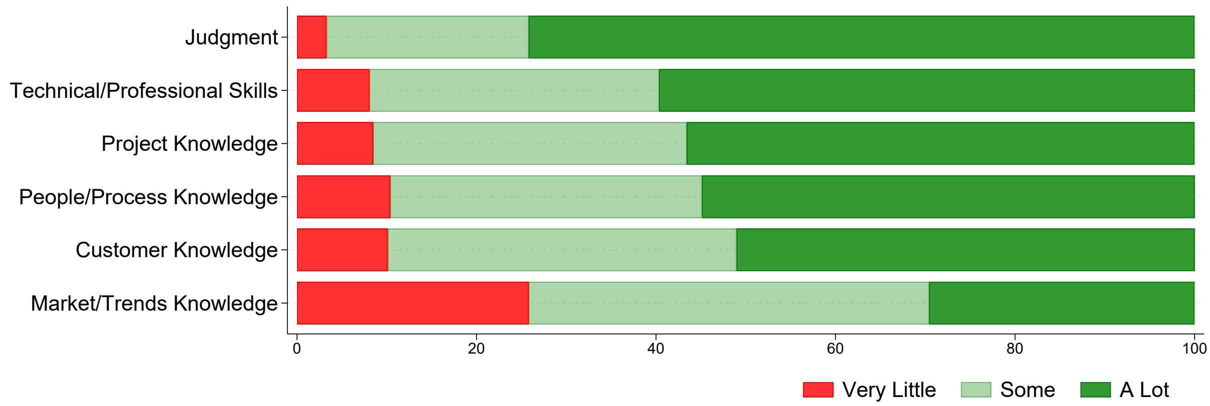
U.S. Congress (1976). Copyright act of 1976, pub. l. no. 94-553, § 101, 90 stat. 2542, 2544. Public Law. Codified at 17 U.S.C. § 101.

U.S. Government Accountability Office (2024, August). Digital surveillance of workers: Tools, uses, and stakeholder perspectives. GAO Report GAO-24-107639, U.S. Government Accountability Office, Washington, D.C.

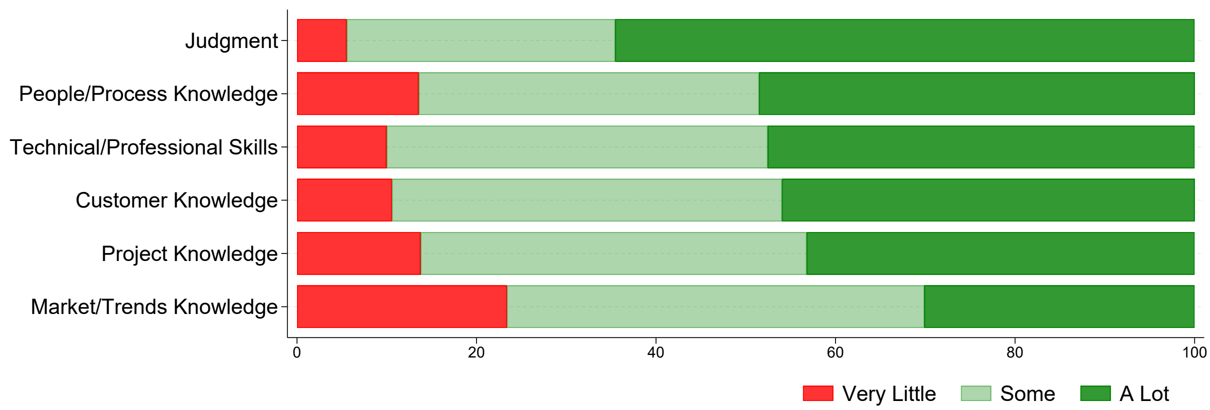
Viljoen, S., D. Li, and Z. Cullen (2026, March). Knowledge guilds: Sharing in the productivity gains from ai. Working paper.

Figure 1: Self Reported Worker Knowledge

A. Job Knowledge

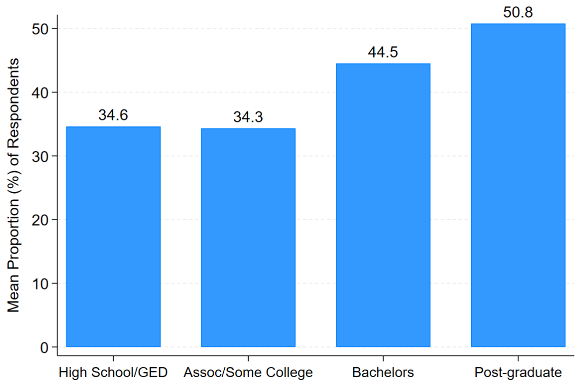


B. Job Knowledge Beyond Company Documentation ("Uncodified" Knowledge)

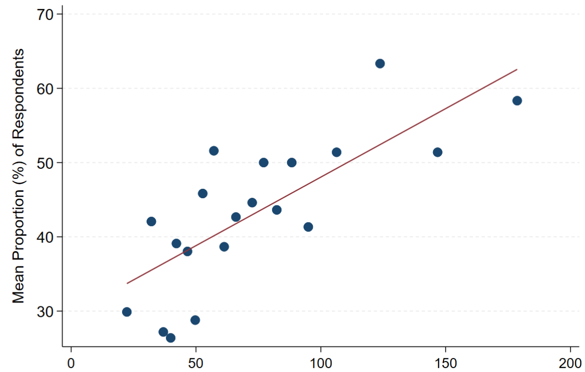


Notes: Panel A shows the extent of the respondents' personal knowledge across 6 knowledge categories. Respondents were asked "In your current role, how much job-relevant knowledge or skill do you personally possess in each of the areas below?" Panel B reports the extent to which knowledge in each category is uncodified. Specifically, respondents were asked "For each area, how much of your personal expertise goes beyond what's captured by your company's documentation (e.g. policy manuals, knowledge databases), training materials (e.g. videos, presentation slides), or other records (e.g. official email and chat records, employee AI prompt history)?" For each category, we exclude respondents who indicate that it is not a core part of their job.

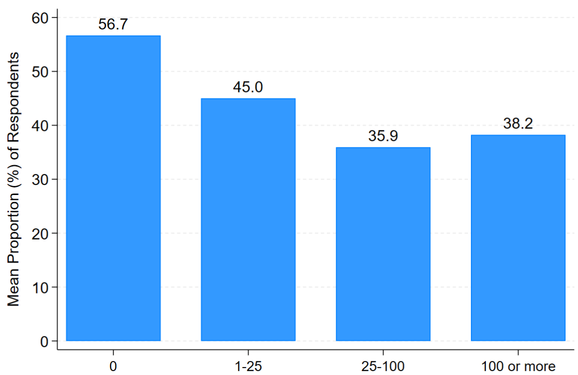
Figure 2: Characteristics of Workers with "A Lot" of Uncodified Knowledge



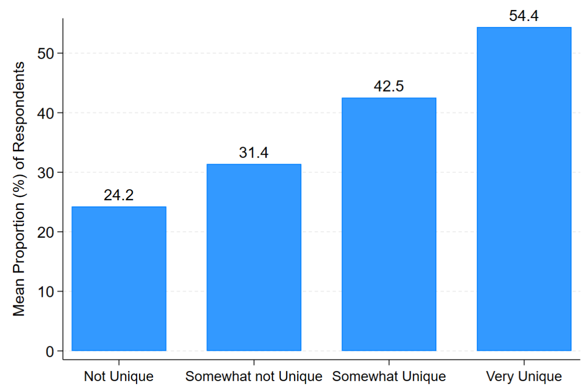
A. Education



B. Annual Income ($000s)



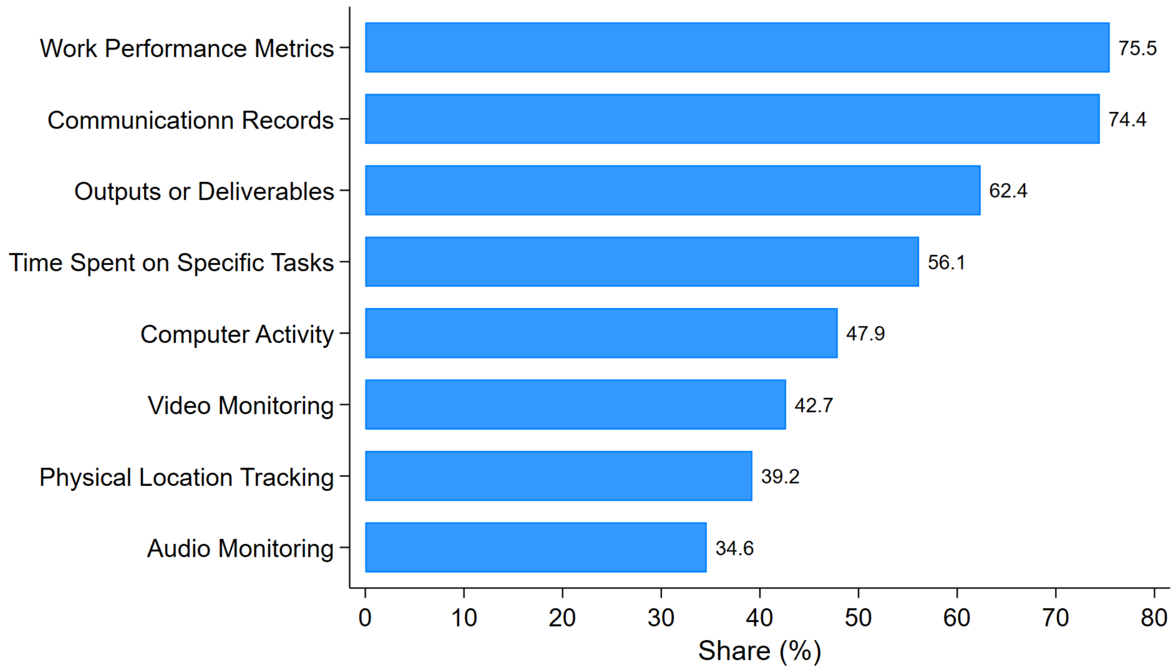
C. Cohort Size



D. Perceived Uniqueness

Notes: For six knowledge categories, respondents indicated whether their knowledge extended little, some, or a lot beyond recorded company documentation (knowledge databases, training materials, etc.). For each respondent, we compute the proportion of categories (out of six) marked "a lot." The figure reports the mean of this proportion across respondents, expressed as a percentage. Note that all hourly wages are converted to annual salaries. "Perceived Uniqueness" measures a respondent's self reported belief about whether their employer could find another person who would perform their job in the same way and with the same quality.

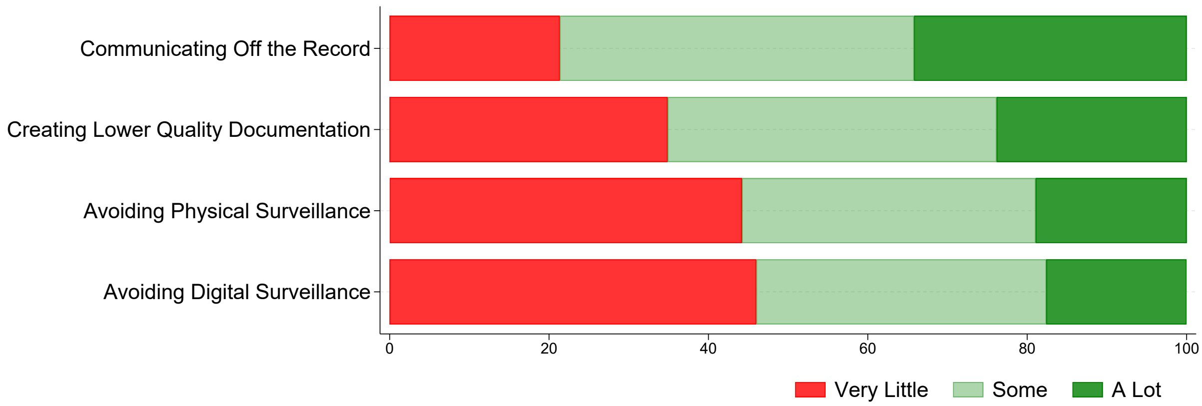
Figure 3: Modes of Workplace Surveillance



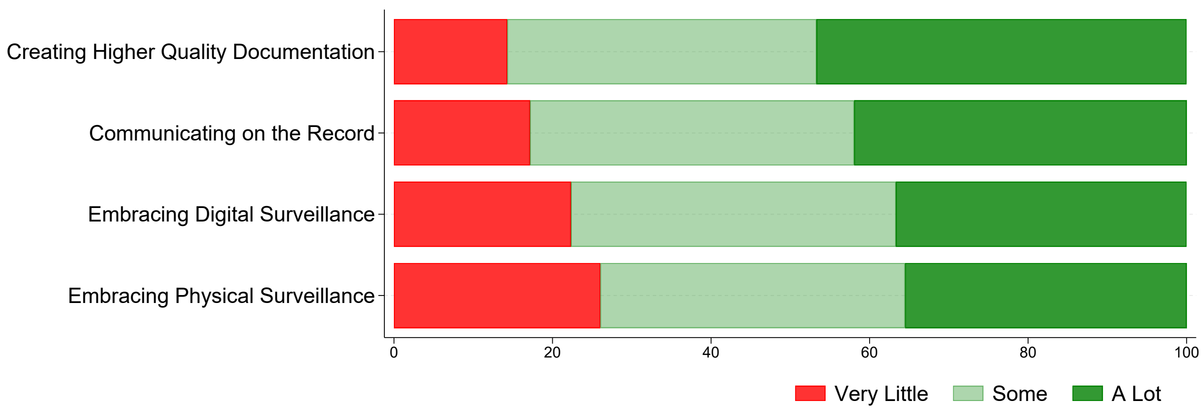
Notes: This figure shows the prevalence of various modes of workplace surveillance within our sample. Respondents were asked "Which of the following activities does your employer monitor or collect data on?" This figure reports the percentage of respondents who reported experiencing each mode of surveillance.

Figure 4: Self-Reported Ability to Withhold and Supply Knowledge

A. Actions to Withhold Knowledge

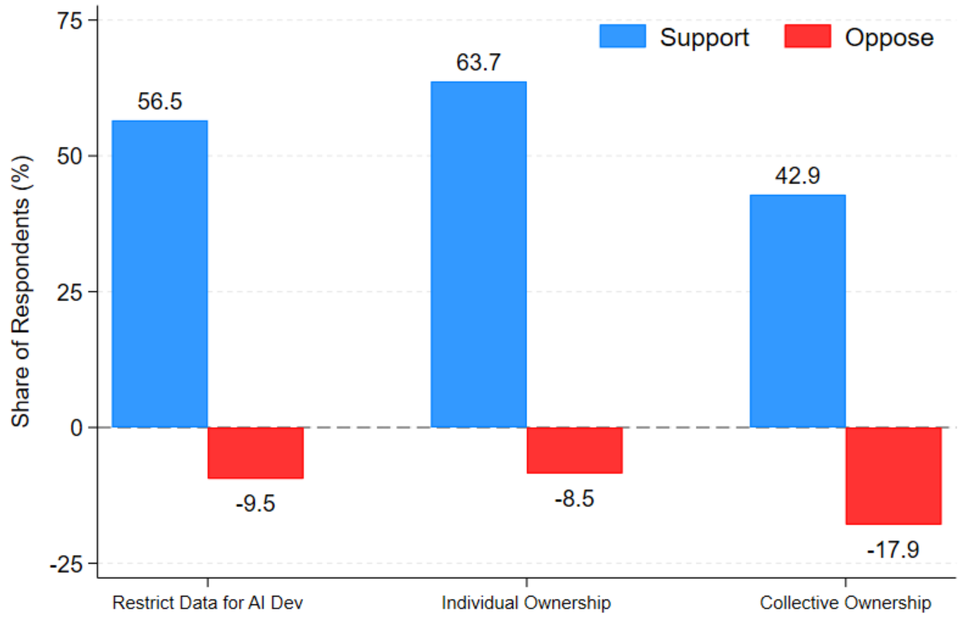


B. Actions to Supply Knowledge



Notes: Panel A shows the extent to which respondents have actions available to withhold knowledge from their employer across four action types, while Panel B shows the same for actions to supply knowledge. For both, participants were asked: "Are there actions that you could take to provide your employer with LESS (MORE) information about your work behaviors?". We emphasize that the question is about whether such actions are *possible*, not whether respondents would actually want to do so. For each action type, we exclude respondents who indicate that it is not relevant to their job.

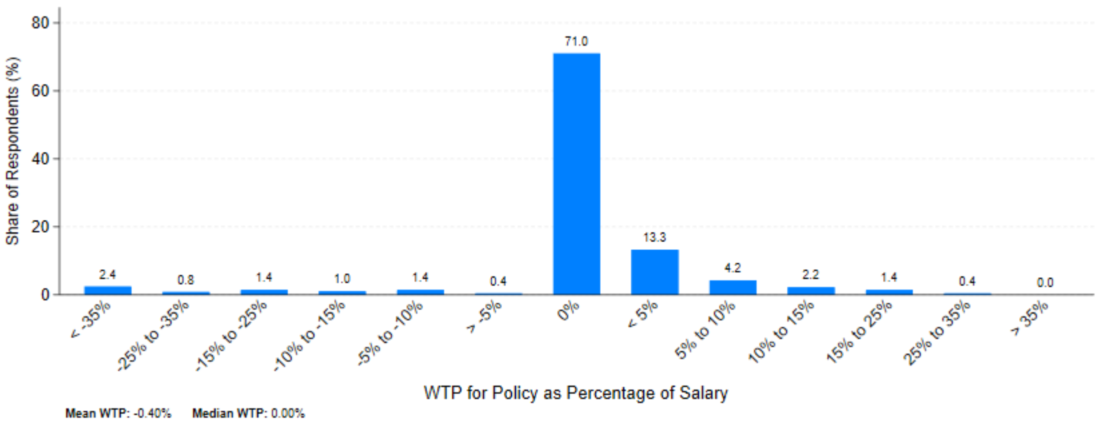
Figure 5: Policy Preferences



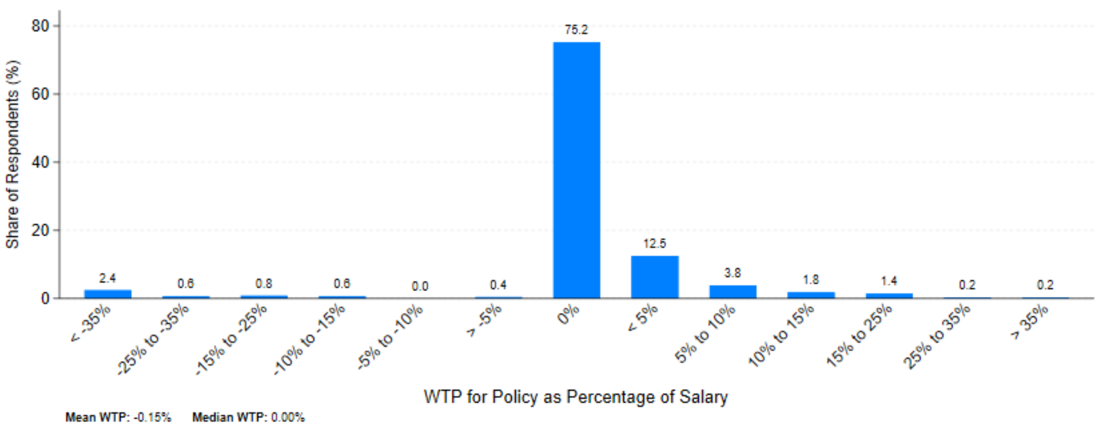
Notes: This figure presents the share of respondents who support vs. oppose a policy restricting data for AI development, a policy of individual data ownership, and a policy of collective data ownership. Those responding "Netural" are not shown.

Figure 6: Wilingness to Pay (WTP) for Policy Change

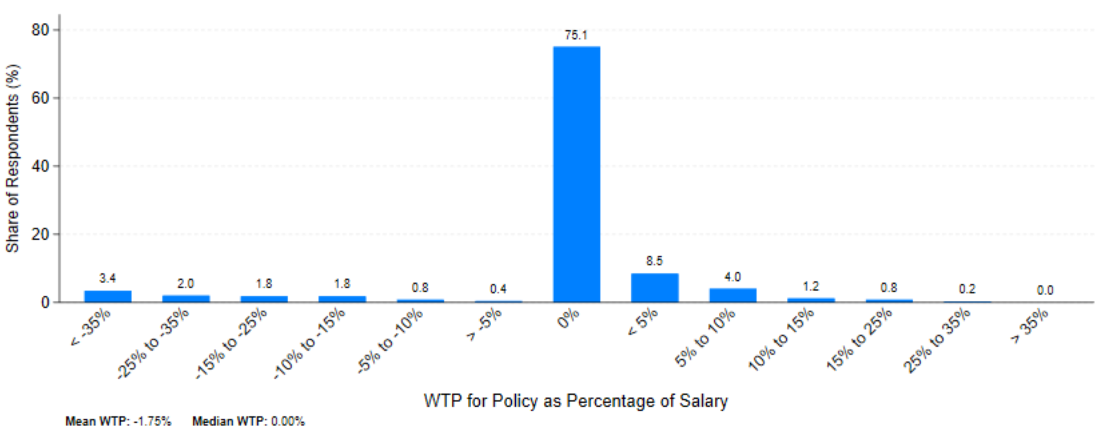
A. WTP for Restriction of Surveillance Data for Training AI Models



B. WTP for Individual Ownership of Work Data



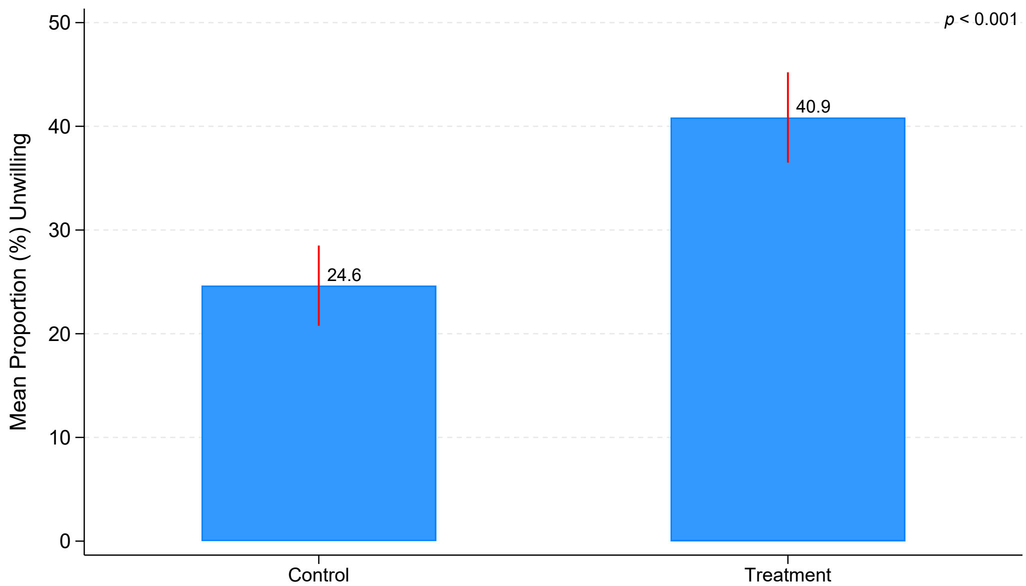
C. WTP for Collective Ownership of Work Data



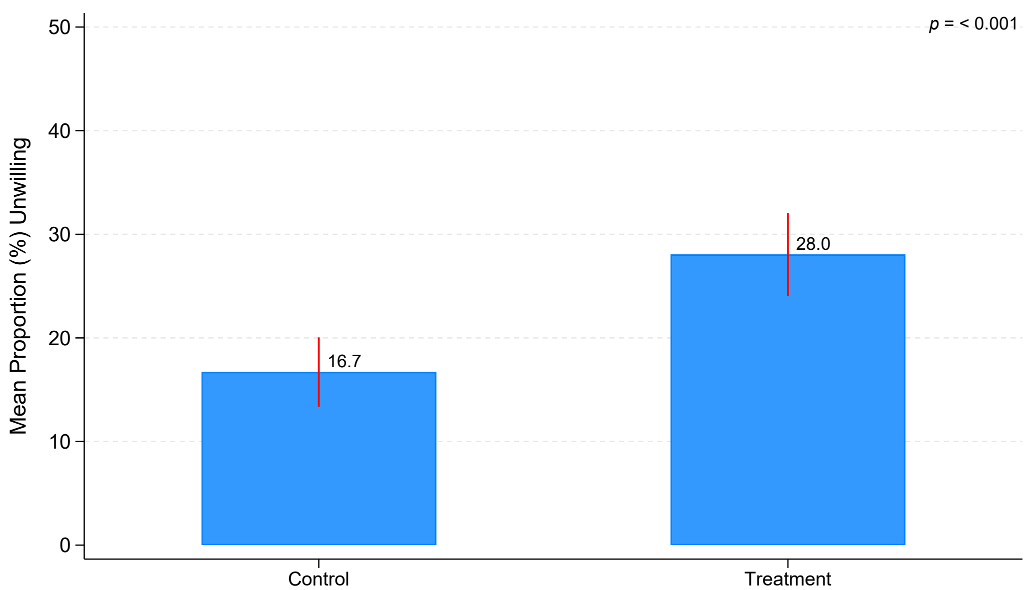
Notes: This figure presents the distribution of willingness to pay (WTP) for each of the three policies, expressed as a percentage of salary. A positive value reflects an acceptable salary reduction to enact the policy, whereas a negative value shows the necessary salary increase. Respondents who were neutral about a policy were coded as having 0% WTP.

Figure 7: Unwillingness to Share Survey Data

A. Unwillingness to Share Past Survey Metadata



B. Unwillingness to Participate in Future Survey



Notes: Panel A displays, by treatment status, the post-treatment share of respondents **unwilling** to provide past survey responses for up to $10. The specific question is "Are you willing to let us use your responses and metadata from your previous Prolific surveys for a bonus of $10? (For a share of willing participants, we will pay this bonus and seek the data to use from past survey collectors)". Panel B similarly reports the share **unwilling** to participate in a future follow-up survey at their current wage. The specific question is "In order to learn more about your specific background and preferences, we would like to administer a 1 hour follow-up survey to gather more examples about how you would answer questions. Are you interested in participating in this survey at your current hourly wage?" Confidence intervals are drawn at 95%. We report two-sided t-test p-values for the null hypothesis of equal means between treatment and control groups.

Table 1: Descriptive Survey Summary Statistics

	Mean	SD	Median	N
Age	37.77	10.61	36.00	493
Male	0.50			493
White	0.74			490
For-Profit Firm	0.73			497
Fully Remote	0.18			497
5+ Days in the Office	0.52			497
Unionized	0.14			497
Manager	0.47			497
Salaried	0.57			497
Annual Salary ($)	91,581.63	40,995.08	81,750.00	282
Hourly	0.43			497
Hourly Wage ($)	25.33	9.77	22.60	215

Notes: This table reports means, standard deviations, and medians for demographic and earnings characteristics. Standard deviations and medians are omitted for binary variables.

Table 2: Treatment Effect on Survey Data Sharing

	(1) Share Past Data	(2) Share Future Data
Treatment	-0.162***	-0.113***
	(0.030)	(0.026)
Control Mean	0.754	0.833
N	971	971

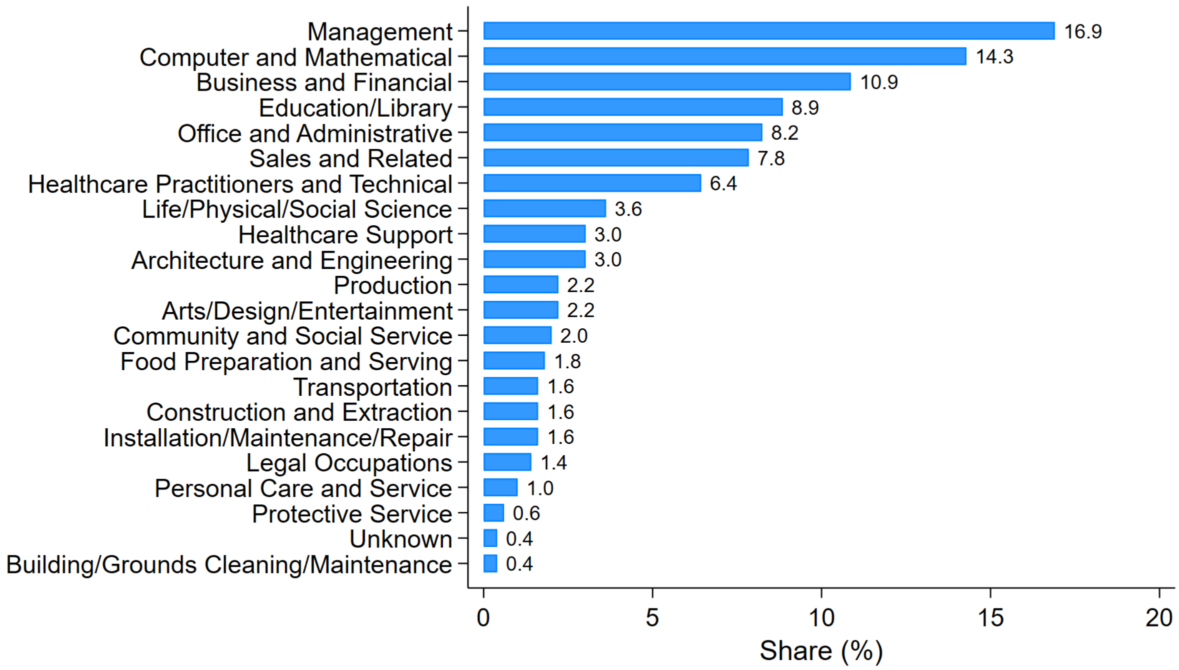
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table describes the effect of treatment on participants' willingness to supply their survey data. Column 1 estimates the impact on willingness to share past Prolific survey data for an additional $10, while Column 2 estimates the impact on interest in completing a future survey at the same survey pay rate.

A Appendix Tables and Figures

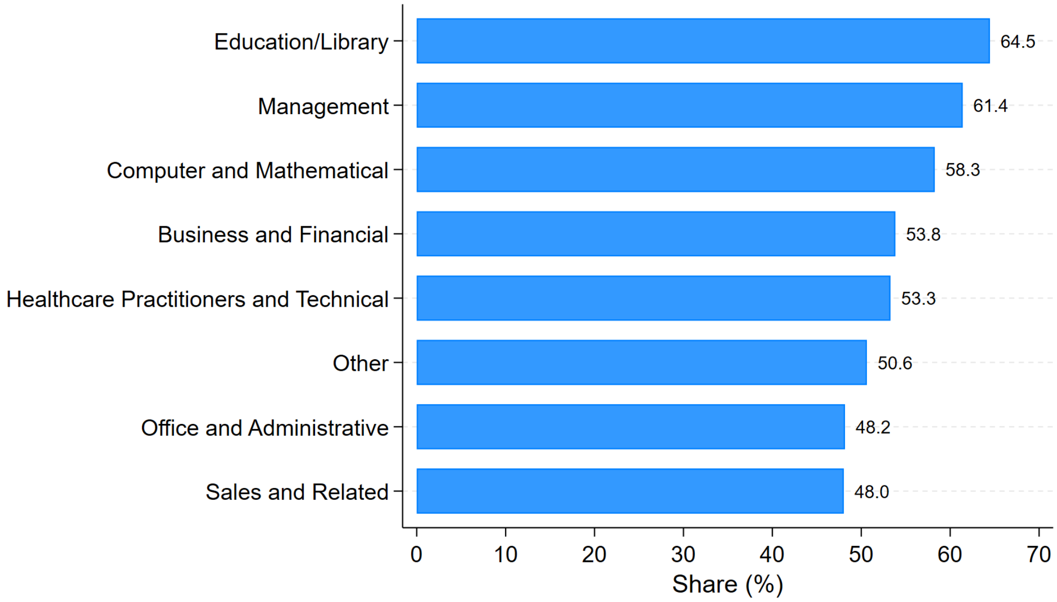
Figure A1: Respondent Occupations (SOC)



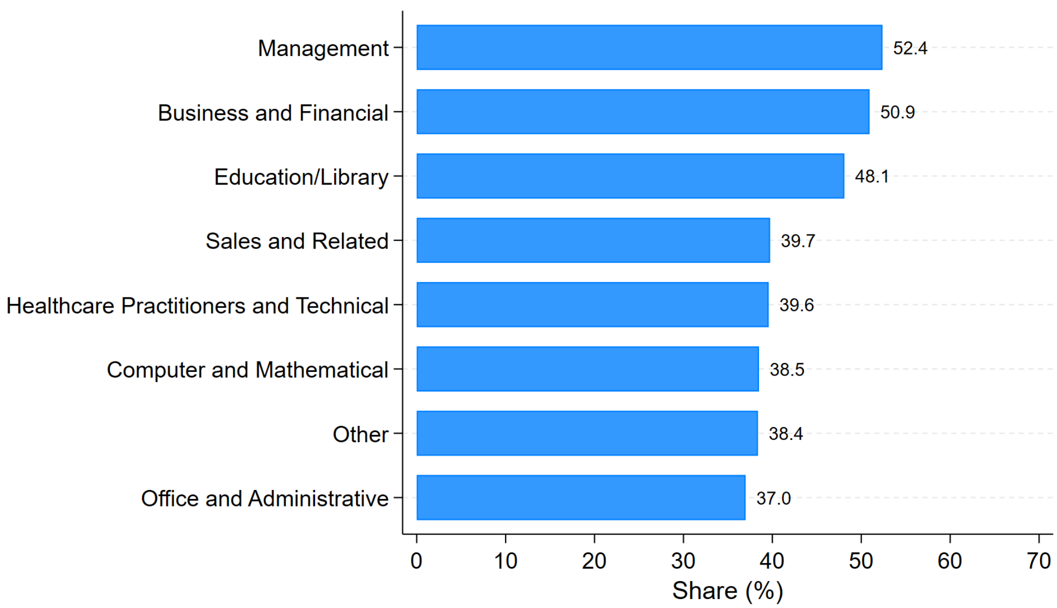
Notes: This figure shows the distribution of SOC major occupation groups within our sample.

Figure A2: Worker Excess Knowledge by Occupation

A. Mean Proportion with "A Lot" of Personal Knowledge

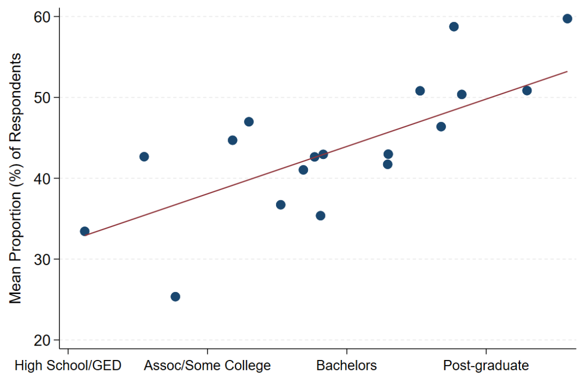


B. Mean Proportion with "A Lot" of Personal Knowledge Beyond Company Documentation

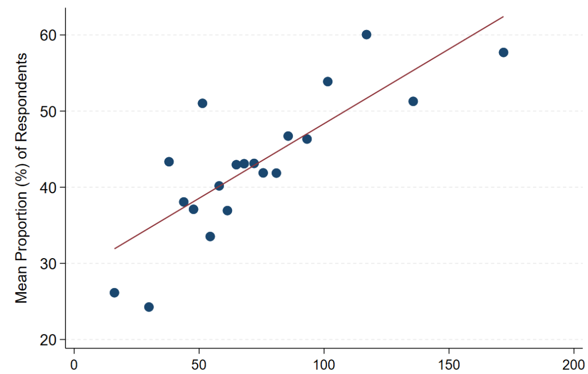


Notes: Occupation categories correspond to the 2018 Standard Occupational Classification (SOC) system. Occupations representing less than 5% of observations were consolidated into an "Other" category, including Life, Physical, and Social Science; Healthcare Support; and Arts, Design, Entertainment, Sports, and Media occupations.

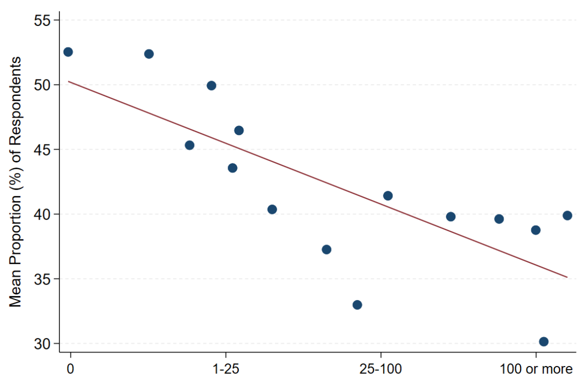
Figure A3: Characteristics of Workers with A Lot of Uncodified Knowledge (Occupation Fixed Effects)



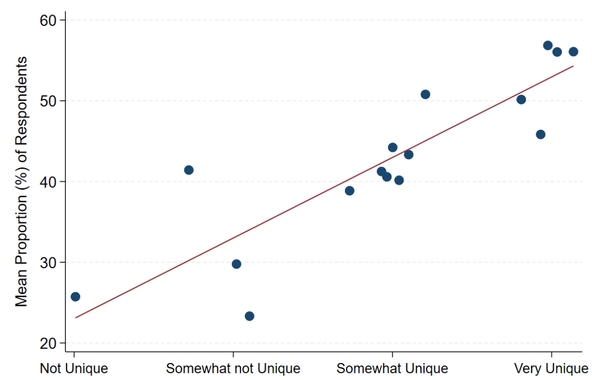
A. Education



B. Annual Income ($000s)



C. Cohort Size

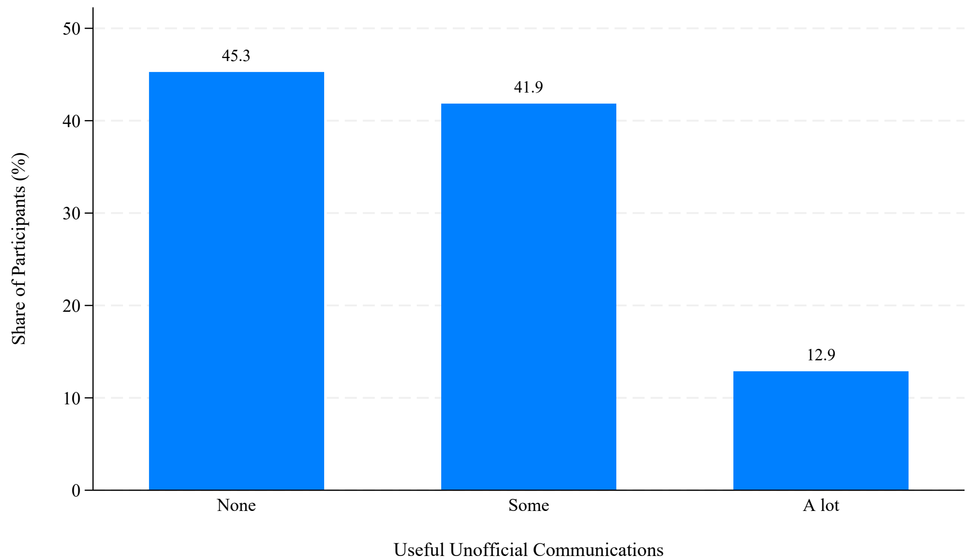


D. Perceived Uniqueness

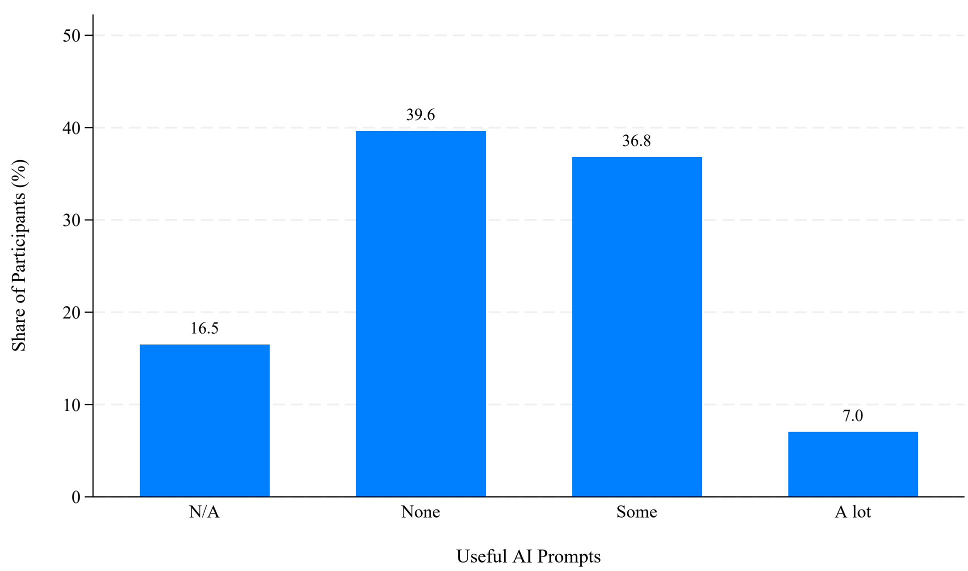
Notes: Respondents indicated whether their knowledge extended little, some, or a lot beyond recorded company documentation for six knowledge categories (as shown in Figure 1). The dependent variable is the mean share of categories rated "a lot" across respondents, expressed as a percentage. We control for occupation fixed effects using SOC categories, with occupations comprising less than 5% of the sample grouped into an "Other" category. Note that all hourly wages are converted to annual salaries. "Perceived Uniqueness" measures a respondent's self reported belief about whether their employer could find another person who would perform their job in the same way and with the same quality.

Figure A4: Work-relevant Data in Unofficial Channels

A. Unofficial Communications



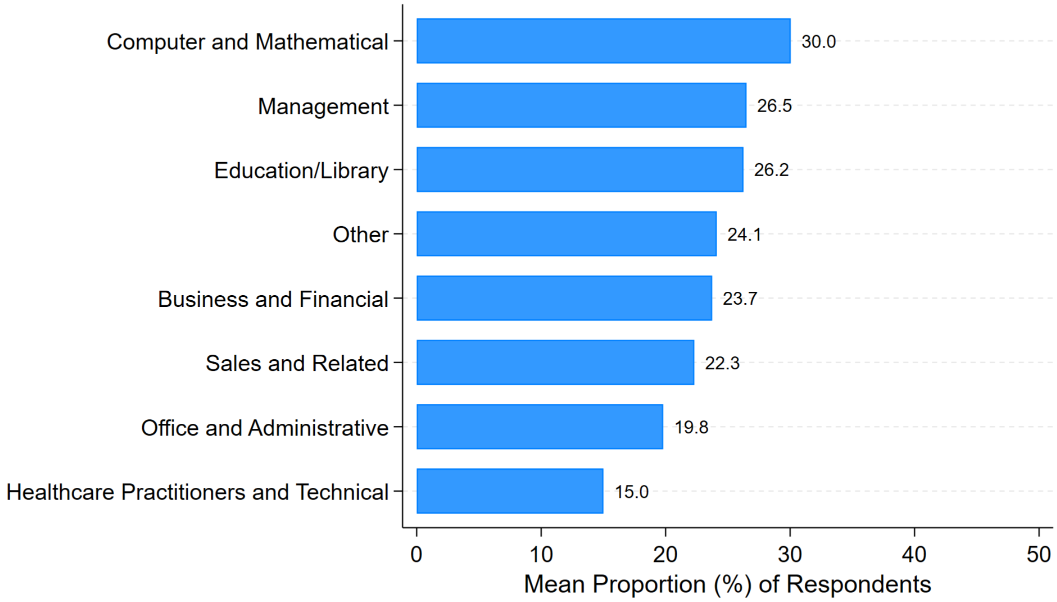
B. AI Prompts



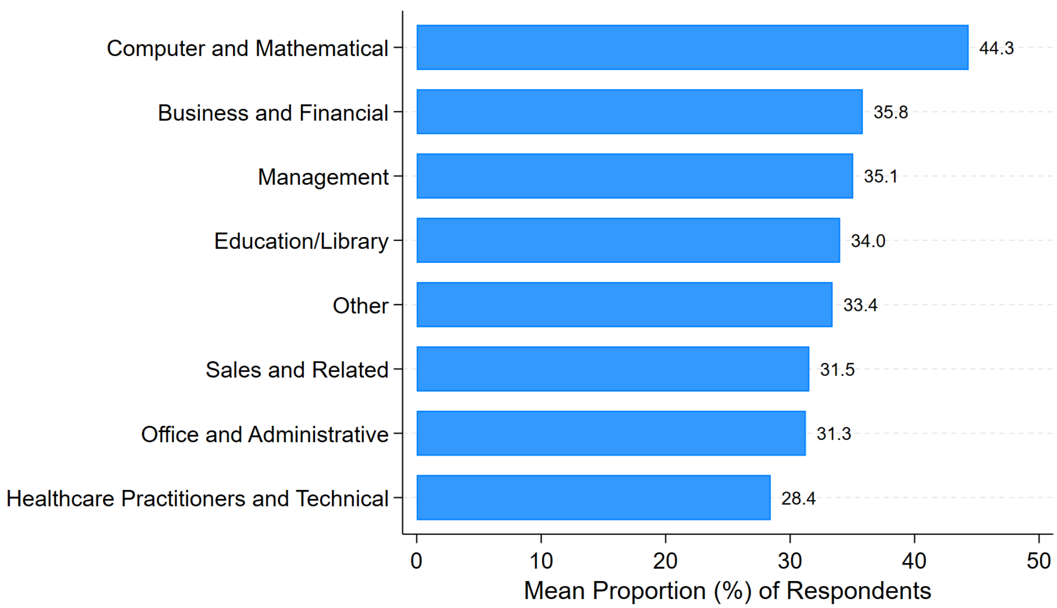
Notes: This figure describes the presence of work-relevant data in respondents' unofficial communications and personal AI prompts. To construct panel A, respondents were asked "To what extent does your unofficial communications (e.g., personal email, messaging apps) include work-related information that would be useful to your employer?" For panel B respondents were asked: "Consider your use of non-official AI tools for work purposes, such as using a personal ChatGPT account. To what extent does your AI prompt history include work-related information that would be useful to your employer?"

Figure A5: Worker Knowledge Supply by Occupation

A. Mean Proportion with "A Lot" of Knowledge Withholding Actions

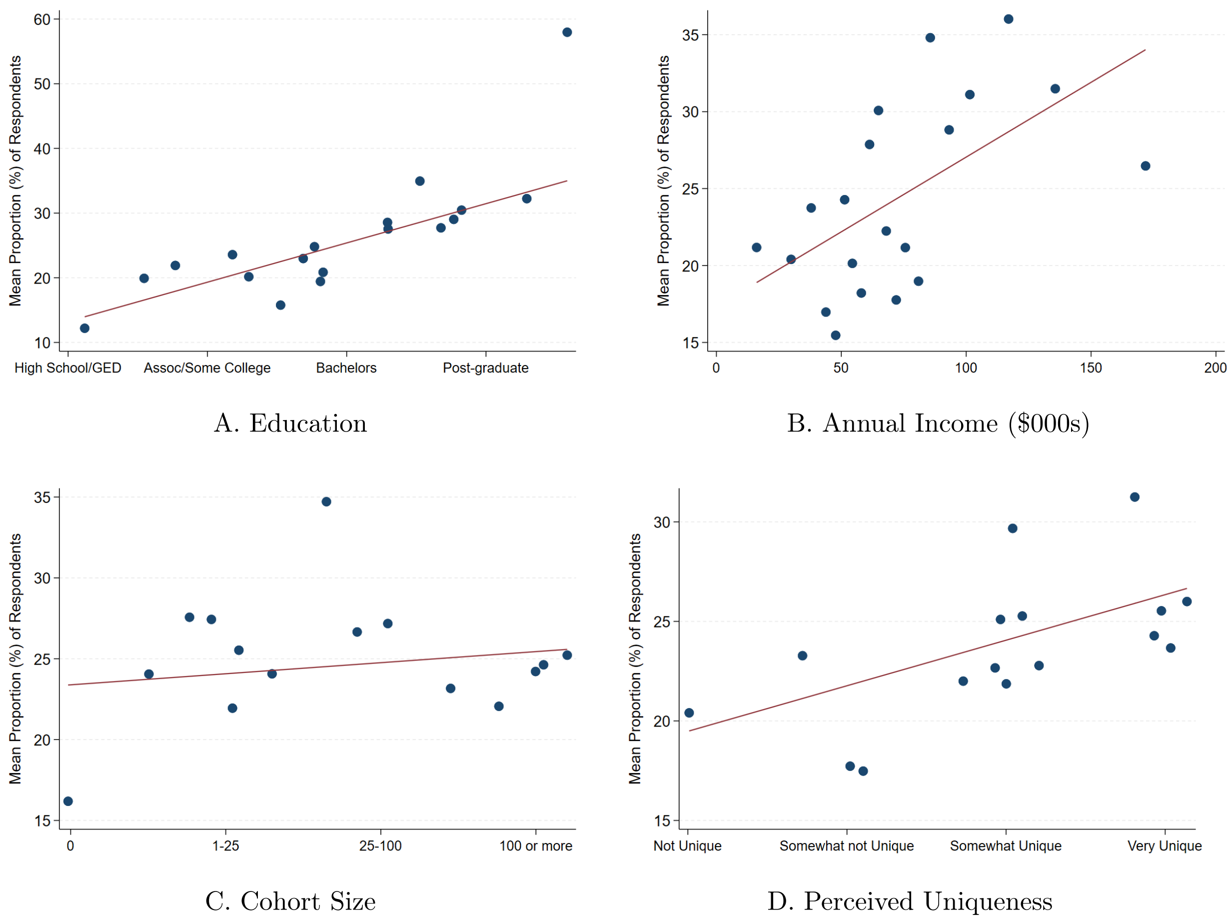


B. Mean Proportion with "A Lot" of Knowledge Sharing Actions



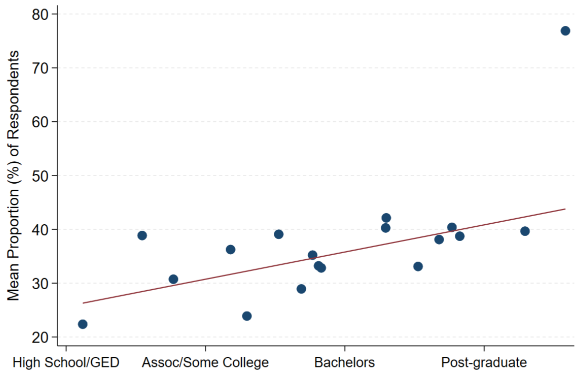
Notes: Occupation categories correspond to the 2018 Standard Occupational Classification (SOC) system. Occupations representing less than 5% of observations were consolidated into an "Other" category, including Life, Physical, and Social Science; Healthcare Support; and Arts, Design, Entertainment, Sports, and Media occupations.

Figure A6: Characteristics of Workers with “A Lot” of Knowledge Withholding Actions

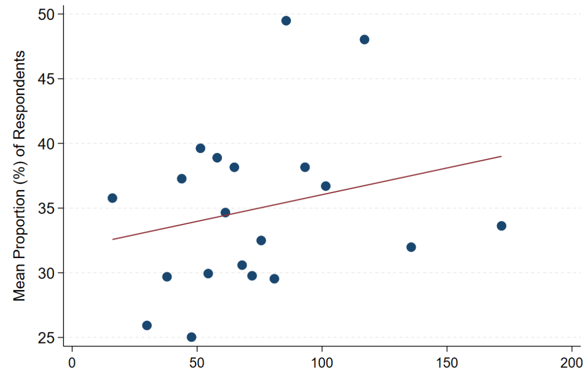


Notes: Respondents indicated whether they had little, some, or a lot of ability to take each of four types of withholding actions (as shown in Figure 4). The dependent variable is the mean share of action types rated “a lot” across respondents, expressed as a percentage. We control for occupation fixed effects using SOC categories, with occupations comprising less than 5% of the sample grouped into an "Other" category. Note that all hourly wages are converted to annual salaries. “Perceived Uniqueness” measures a respondent’s self reported belief about whether their employer could find another person who would perform their job in the same way and with the same quality.

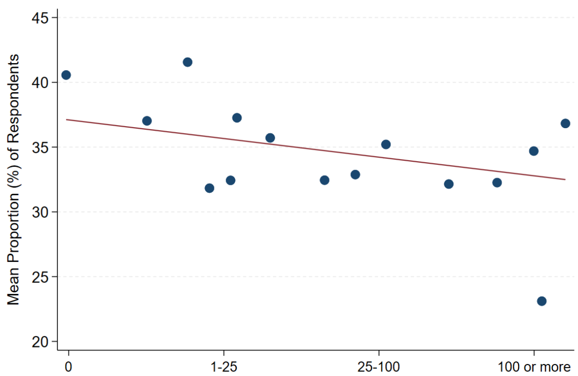
Figure A7: Characteristics of Workers with “A Lot” of Knowledge Sharing Actions



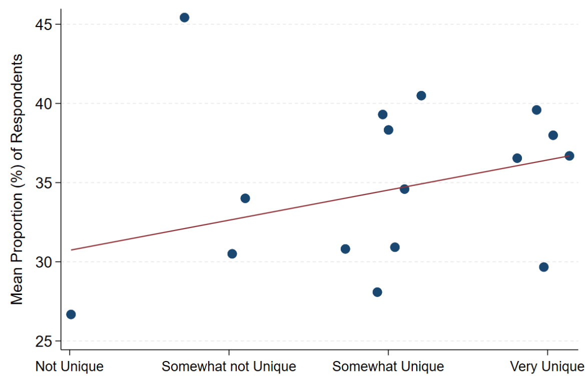
A. Education



B. Annual Income ($000s)



C. Cohort Size



D. Perceived Uniqueness

Notes: Respondents indicated whether they had little, some, or a lot of ability to take each of four types of sharing actions (as shown in Figure 4). The dependent variable is the mean share of action types rated “a lot” across respondents, expressed as a percentage. We control for occupation fixed effects using SOC categories, with occupations comprising less than 5% of the sample grouped into an "Other" category. Note that all hourly wages are converted to annual salaries. “Perceived Uniqueness” measures a respondent’s self reported belief about whether their employer could find another person who would perform their job in the same way and with the same quality.

Table A1: Comparison of Survey and CPS Sample Characteristics

	Survey	CPS	Difference
Age	37.77	41.02	-3.25
	(10.61)	(12.19)	
Male	0.50	0.53	-0.04
White	0.74	0.75	-0.00
For-Profit Firm	0.73	0.83	-0.10
Fully Remote	0.18		
5+ Days in the Office	0.52	0.63	-0.11
Unionized	0.14	0.11	0.03
Manager	0.47	0.19	0.28
Salaried	0.57	0.51	0.06
Annual Salary ($)	91,581.63	108,127.59	-16,545.96
	(40,995.08)	(102,302.74)	
Hourly	0.43	0.49	-0.06
Hourly Wage ($)	25.33	36.01	-10.68
	(9.77)	(20.33)	
N	497	45,185	

Notes: This table compares summary statistics between our survey sample and a cross section of the U.S. full-time workforce, with data obtained from the 2025 Annual Social and Economic Supplement (ASEC) of the Current Population Survey. The CPS dataset is restricted to employed full-time workers, aged 18-65 with at least a high school education. Standard deviations for continuous variables are shown in parentheses. "5+ Days in the Office" in the CPS is approximated using the telework supplement (respondents who did not telework or work from home for pay in the prior week). "Manager" in the CPS includes management occupations and first-line supervisors, which may not be fully comparable to the survey's self-reported manager measure (where we ask whther the respondent is *anyone's* manager or supervisor). This likely partially explains the overrepresentation of managers in our sample. "Fully Remote" is not available in the CPS. CPS demographic statistics are weighted using ASEC person weights. Union status, salaried/hourly classification, annual salary, and hourly wage are derived from the Outgoing Rotation Group (ORG) supplement, weighted using ORG earnings weights, and are available for approximately one-quarter of the CPS sample.

B Policy Vignette Language

B.1 No Monitoring of Work Activity

Imagine a policy that forbids your employer from monitoring or storing any data about your individual work activities except when strictly required for safety or legal compliance. This policy would not prevent you from using workplace tools, but it would prevent your employer from recording, archiving, or analyzing your workplace activities.

For example:

- *Your employer would have no right to intercept, archive, or analyze your workplace communications.*
- *Your employer could not track your physical location, or make video, audio, or screen recordings of your work.*

B.2 No AI Automation of Core Job Tasks

Imagine a policy that forbids your employer from developing or adopting AI models to automate core parts of your current job.

For example:

- *If you work in customer service, it would bar your employer from deploying AI chatbots or virtual assistants to handle customer calls.*
- *If you're an office administrator, it would stop your employer from automating tasks like filing expense reports or scheduling meetings.*

This restriction applies only to what your employer can doâit would not prevent you from choosing to use AI tools on your own.

B.3 No Use of Work Data for AI Development

Imagine a policy that bans the use of "work data" for AI model development. By "work data," we mean:

- *The materials you produce on the job—reports, presentations, code, designs, project plans, marketing assets, and so on.*

- *Any record of how you work—emails and chat logs, screen or video recordings, and logs of your digital or physical activities.*

Under this policy:

1. *Employers could not use work data to develop AI models, or sell work data to other firms to develop AI models.*
2. *Those seeking to develop AI models would have to hire workers to specifically produce AI training data; they could not use data that was produced as the byproduct of everyday work tasks.*

C Proofs

C.1 Proof of Theorem 4.9

Proof. The worker payoffs (7) from time-1 contributions induced by each technology are

$$w_1^j - c^j(k_1^j) + \psi\gamma \left\lfloor \theta_j - \alpha^j \left(k_1^{\overline{J}_1 \cup \{j\}} \right) \right\rfloor_+ , \tag{35}$$

$$w_1^j - c^j(k_1^j) + \psi\gamma \left\lfloor \theta_j - \tilde{\alpha}^j \left(k_1^{\overline{J}_1 \cup \{j\}} \right) \right\rfloor_+ . \tag{36}$$

By the same argument as in Lemma 4.4, these define a pair of supermodular games. Let us index these games by parameter τ, with $\tau = 0$ corresponding to $\left(\alpha^j\right)_{j\in J}$ and $\tau = 1$ corresponding to $\left(\tilde{\alpha}^j\right)_{j\in J}$. By (25), (26), and Lemma 4.3, it follows that each worker j's utility has increasing differences in (k_1^j, τ). Thus, the comparison of extremal equilibria follows by Theorem 6 of Milgrom and Roberts (1990).

Finally, observe that worker j's time-2 wage under the $\overline{k}_1$ and α is

$$\left\lfloor \theta_j - \alpha^j \left(\overline{k}_1\right) \right\rfloor_+ \geqslant \left\lfloor \theta_j - \alpha^j \left(\overline{\tilde{k}}_1\right) \right\rfloor_+ \geqslant \left\lfloor \theta_j - \tilde{\alpha}^j \left(\overline{\tilde{k}}_1\right) \right\rfloor_+ , \tag{37}$$

where the first inequality is by α non-decreasing and the second inequality by (25). The right-hand side of (37) is worker j's time-2 wage under the $\overline{\tilde{k}}_1$ and $\tilde{\alpha}$. The same argument holds comparing wages under $\underline{\tilde{k}}_1$ and $\underline{k}_1$. □

C.2 Proof of Theorem 4.12

Proof. By (26) we have

$$\alpha^l(\theta) - \alpha^l(0, \theta^{-j}) \geqslant \tilde{\alpha}^l(\theta) - \tilde{\alpha}^l(0, \theta^{-j}) \text{ for any workers } j \text{ and } l. \tag{38}$$

By (31) and (38), we have

$$-\alpha^l(0, \theta^{-j}) \geqslant -\tilde{\alpha}^l(0, \theta^{-j}) \text{ for any workers } j \text{ and } l. \tag{39}$$

Worker j's payoff under $\left(\alpha^l\right)_{l\in J}$ is

$$\begin{aligned} &\gamma\theta^j + \psi\gamma\left(\max\left\{\theta^j, \alpha^j\left(\theta\right)\right\} - \alpha^j\left(0, \theta^{-j}\right) + \sum_{l\neq j}\left(\max\left\{\theta^l, \alpha^l\left(\theta\right)\right\} - \max\left\{\theta^l, \alpha^l\left(0, \theta^{-j}\right)\right\}\right)\right) \\ &\geqslant \gamma\theta^j + \psi\gamma\left(\max\left\{\theta^j, \tilde{\alpha}^j\left(\theta\right)\right\} - \tilde{\alpha}^j\left(0, \theta^{-j}\right) + \sum_{l\neq j}\left(\max\left\{\theta^l, \tilde{\alpha}^l\left(\theta\right)\right\} - \max\left\{\theta^l, \tilde{\alpha}^l\left(0, \theta^{-j}\right)\right\}\right)\right), \end{aligned} \tag{40}$$

where the inequality is by (31) and (39). The right-hand side is equal to worker j's payoff under $\left(\tilde{\alpha}^l\right)_{l\in J}$ □

C.3 Proof of Theorem 4.13

Proof. We are comparing full-contribution equilibria, so workers incur no withholding costs. Thus, it suffices to show that workers' wages are lower under individual data ownership than under no surveillance.

There are at least two workers, they have identical maximum contributions, and their contributions are perfect substitutes, so expression (27) for pairwise surplus at time 2 under individual ownership reduces to

$$\max\left\{\theta^j, \alpha\left(k_1^J\right)\right\} - \alpha\left(k_1^{J\setminus\{j\}}\right) = \left\lfloor\theta^j - f(\theta^j)\right\rfloor_+, \tag{41}$$

which is strictly less than θ^j by $f(0) = 0$, $\theta^j > 0$, and f increasing. Thus wages at time 2 under individual data ownership are strictly lower than $\gamma\theta^j$ by $\gamma \in (0, 1]$, which is the wage under no surveillance.

Next we consider time-1 wages. There are at least three workers, so deviating to not hire a single worker at time-1 has no effect on time-2 output or on time-2 wages. Thus, the pairwise surplus of hiring worker j at time-1 is equal to θ^j, and time-1 wages are the same under individual ownership and under no surveillance.

We have established that individual ownership results in the same wages at time 1 and strictly lower wages at time 2, which by $\psi > 0$ implies that workers are strictly worse off. $\square$

C.4 Proof of Theorem 4.16

Proof. Observe that even under costly contributions, the pairwise surplus from employing an additional worker at time 2 is at least zero, so full employment at time 2 is consistent with equilibrium. Given full employment at time 2, worker j's payoff is still as in (28). By $\left|K^j\right|$ finite and the existence of Nash equilibrium in finite games, for any employed set $\overline{J}_1$, there exists a (possibly mixed) profile of time-1 contributions in which each worker's choice of k_1^j maximizes their expected payoff (that is, the expectation of (28)) given $\overline{J}_1$ and the other workers' mixtures over contributions.

It remains to show that full employment at time 1 is consistent with the equilibrium we are constructing. Observe that, under full employment at time 1, the pairwise surplus of the firm and worker j at time 1 (their time 1 payoffs plus their continuation payoffs) is

$$\begin{aligned}
&\underbrace{k_1^j - c^j(k_1^j)}_{\text{time 1 output and cost}} + \underbrace{\psi \max\left\{\theta^j, \alpha^j\left(k_1^{J_1}\right) - \alpha^j\left(k_1^{J_1\setminus\{j\}}\right)\right\}}_{\text{change in continuation payoff for role } j} \\
&\quad + \underbrace{\psi \sum_{l\neq j}\left((1-\gamma)\left\lfloor \max\left\{\alpha^l\left(k_1^{J_1}\right), \theta^l\right\} - \alpha^l\left(k_1^{J_1\setminus\{l\}}\right)\right\rfloor_+ + \alpha^l\left(k_1^{J_1\setminus\{l\}}\right)\right)}_{\text{time-2 profit from other roles with } j\text{'s data}} \\
&\quad - \underbrace{\psi \sum_{l\neq j}\left((1-\gamma)\left\lfloor \max\left\{\alpha^l\left(k_1^{J_1\setminus\{j\}}\right), \theta^l\right\} - \alpha^l\left(k_1^{J_1\setminus\{j,l\}}\right)\right\rfloor_+ + \alpha^l\left(k_1^{J_1\setminus\{j,l\}}\right)\right)}_{\text{time-2 profit from other roles without } j\text{'s data}}. \qquad (42)
\end{aligned}$$

The above expression is at least zero, by (34) and α^j, α^l non-decreasing, which completes the proof. $\square$

C.5 Proof of Theorem 4.17

Proof. Let κ_1^{-j} be the random variable representing other workers' time-1 contributions under some equilibrium. Observe that for k_1^j played with positive probability, we have

$$k_1^j \in \arg\max_{\hat{k}_1^j} E\left[w_1^j - c^j(\hat{k}_1^j) + \psi\gamma\lambda_j\left(\hat{k}_1^j, \kappa_1^{-j}, \overline{J}_1\right)\right], \tag{43}$$

where λ_j is as defined in (27). Moreover we have

$$\left\{\theta^j\right\} = \arg\max_{\hat{k}_1^j} E\left[w_1^j - c^j(k_1^j)\right], \tag{44}$$

because θ^j is the unique minimizer of c^j. Since $E\left[\lambda_j\left(\hat{k}_1^j, \kappa_1^{-j}, \overline{J}_1\right)\right]$ is non-decreasing in $\hat{k}_1^j$, the result follows by Topkis's theorem. $\square$

C.6 Proof of Theorem 4.18

Proof. Let $\mathbf{k}_1$ be the random variable representing the time-1 contribution vector in an equilibrium with collective ownership. For any worker j, they cannot profitably deviate to contribute θ^j at time 1, so for any k_1^j in the support of $\mathbf{k}_1$ we have

$$E\left[-c^j(k_1^j) + \gamma\psi\frac{\sum_l \max\left\{\theta^l, \alpha\left(k_1^j, \mathbf{k}_1^{-j}\right)\right\}}{|J|}\right] \geqslant E\left[\gamma\psi\frac{\sum_l \max\left\{\theta^l, \alpha\left(\theta^j, \mathbf{k}_1^{-j}\right)\right\}}{|J|}\right] \geqslant \gamma\psi\frac{\sum_l \theta^l}{|J|}. \tag{45}$$

Summing (45) across j and by $\gamma \in [0,1]$ we have

$$\sum_j E\left[-c^j(\mathbf{k}_1^j) + \psi\max\left\{\theta^l, \alpha(\mathbf{k}_1)\right\}\right] \geqslant \sum_j \psi\theta^j. \tag{46}$$

By the same argument as in Appendix C.5, if k_1^j is in the support of $\mathbf{k}_1$, then we have $k_1^j \geqslant \theta^j$. Thus by (46) we have

$$\sum_j E\left[\mathbf{k}_1^j - c^j(k_1^j) + \psi\max\left\{\theta^l, \alpha(\mathbf{k}_1)\right\}\right] \geqslant \sum_j \left(\theta^j + \psi\theta^j\right). \tag{47}$$

The left-hand side of (47) is the expected total surplus under collective ownership and the right-hand side is the total surplus in the no-AI case. $\square$