Artificial Intelligence and Worker Power

Owen F. Davis*

June 2024

A growing body of economic research explores the possible consequences of artificial intelligence replacing or augmenting human labor, yet little is known about how AI might affect worker power. This is despite the fact that the same aspects of AI that make it uniquely disruptive as a tool of automation—carrying out non-routine tasks, assimilating tacit knowledge, refining predictions—also position AI for applications in management and human resources. This scoping review surveys and reframes the economic literature on AI and labor, detailing how new technologies might alter worker power with consequences for income distribution and job quality. To help clarify the theoretical effects of AI on workers and the nature of work, I draw a distinction between the *labor demand* impacts and the *worker power* impacts of new technologies (or, between AI's effects on job content versus job context). Labor demand impacts stem from AI that substitutes for tasks now carried out by humans. Worker power impacts stem from AI that sharpens managerial control, undermines workplace norms, or heightens information asymmetries. With this framework in mind, I formalize several avenues by which AI potentially shifts power dynamics in workplace and bargaining contexts, including monitoring and surveillance, predictive analytics in wage bargaining, and algorithmic management systems that diminish worker autonomy. This study links the evolving literature on the economics of AI to research strands concerned with the determinants of worker power and its long-term trends.

1 Introduction

Millions of workers have experienced artificial intelligence enter the workplace in recent years. Public-opinion surveys indicate that between one-fifth and one-half of U.S. workers have used ChatGPT or other generative AI tools at work. Rates of reported AI usage reach 75% among knowledge workers and 90% among software developers. A quarter of human resources professionals report using AI technologies. AI-powered tools that struggled with basic cognitive tasks a decade ago now rival humans in contexts ranging from image and speech recognition to medical diagnostics to business idea generation. Researchers project that large swaths of the labor force are exposed to AI, a prediction shared by workers themselves.[1]

While economic research has understandably centered on the possible labor-market consequences of AI replacing or augmenting human labor, comparatively little research focuses on how AI might affect worker

*Siegel Family Endowment: owen.davis@siegelendowment.org.

[1] Overall worker AI use: Conference Board (2023), Kreacic et al. (2024), He (2024), McClain (2024). Knowledge workers: Microsoft and LinkedIn (2024). Software developers: GitHub (2023). Human resources: SHRM (2024). Image and speech recognition: Kiela et al. (2023). Medical diagnostics: Alvarez-Valle and Lungren (2023). Business idea generation: Girotra et al. (2023). Researcher projections of labor force impacts: see studies reviewed in Section 5. Worker projections of AI impacts: Faveiro and Tyson (2023).

power. This is despite high-profile examples of AI-enabled technologies diminishing worker autonomy and altering the nature of work in settings ranging from warehouses to taxi service to call centers. It is widely accepted that AI-driven automation could have profound effects on the labor market and wider society, yet the potentially far-reaching implications of AI for power dynamics have been relatively neglected.

This paper surveys and reframes the economic literature on AI and labor, detailing how AI might alter worker power and outlining potential consequences for income distribution and job quality. To situate this project within the existing research landscape, I review the growing literature on AI and labor, which focuses almost exclusively on AI in production, or technology that substitutes for human labor in the tasks that make up a job. Yet to lay out a research agenda around AI and worker power, it is necessary to highlight a broader range of theoretical approaches and aspects of work other than production.

To help clarify the theoretical effects of AI on work and workers, I draw a distinction between the *labor demand* impacts and the *worker power* impacts of new technologies. This distinction can also be seen as delineating between AI's effects on job content versus job context.[2] AI tools that substitute for human labor in job tasks, or the content of work, affect workers primarily through changes to labor demand. AI tools that alter the management and organization of work, or job context, affect workers through their potential to shift the division of rents or surplus. A similar demarcation can be made between the effects of AI on bargaining position and on bargaining power. New technologies may affect the bargaining positions of firms and workers by raising the threat of automation or limiting workers' outside options. AI may affect worker power, for example, by disrupting norms or creating information asymmetries.

Given the multitude of AI's potential economic ramifications, why focus on worker power? The primary motivation stems from the observation that the same aspects of AI that make it uniquely disruptive as a tool of automation—carrying out nonroutine tasks, assimilating tacit knowledge, refining predictions—also position AI for applications in management and human resources that shift workplace power dynamics. Although largely overlooked by economic theorists, the view that AI confers a power advantage to managers has been articulated in sociology and management studies (Kellogg, Valentine and Christin, 2020; Jarrahi et al., 2021; Bernhardt, Kresge and Suleiman, 2023) as well as legal scholarship (Kim, 2022; Rogers, 2023). Pizzinelli et al. (2023) project that among occupational categories, managers face the highest potential for AI-enabled complementation.

This study lays out several specific avenues by which AI augmentation or automation of managerial functions may affect worker power. AI enhances monitoring, which in standard efficiency-wage models shifts rents from workers to firms. AI-driven predictive analytics in wage offer and bargaining provide more scope for discriminating monopsony, or paying each worker as close as possible to their reservation wage. Automating the direction and evaluation of workers shifts authority to management and reduces the need to incentivize worker initiative. In each of these contexts, I sketch simple analytical models to capture the effects of interest. Other worker-power impacts of AI include surveillance tools that chill organizing efforts; recruiting systems that allow firms to select for compliant workers; algorithmic management systems that sever channels of negotiation and the exercise of voice; business reorganizations that allow firms to further disintermediate the workforce through fissuring, eroding internal wage fairness norms; and targeted automation and deskilling that erode worker solidarity.[3] A theme that emerges across

[2]Previous scholarship and policy-related analyses have made similar distinctions. Lane and Saint-Martin (2021) delineate between AI impacts on "productivity, employment and wages" and impacts on "the work environment." Gmyrek, Berg and Bescond (2023) consider job quantity and job quality separately. The US Council of Economic Advisors (2022) report distinguishes between impacts on workers and impacts on the workplace.

[3]AI in the workplace or in job search may also benefit workers. Algorithmic evaluation could reduce managerial caprice

theoretical models is that AI may disadvantage workers by shifting information rents to employers.

This paper connects to several related economic literatures concerned with how worker power relates to distributional trends and the role of norms and institutions (Krueger, 2018). On one side is research that seeks to quantify the degree of monopsony in the labor market (Sokolova and Sorensen, 2021) as well as its effects on workplace dynamics and job quality (Dube, Naidu and Reich, 2022). Institutionalist researchers have explored the worker-power impacts of declining unionization (Levy and Temin, 2007) and evolving firm structure (Weil, 2014). A related research area explores how declines in worker power might help to explain wider distributional trends and related macroeconomic puzzles (Setterfield, 2006; Stansbury and Summers, 2020; Lombardi, Riggi and Viviano, 2020; Jacobo, 2023). Finally, an older radical strand, exemplified by Braverman (1998), Marglin (1974) and Edwards (1979), foregrounds the power dynamics inherent in technological change.

The burgeoning economic literature on AI and labor has grown independently of research lines concerned with worker power. In the task framework, automation might disadvantage workers, but considerations of power in bargaining and in the workplace are largely absent. This is not to say that the major contributors to this literature have overlooked the broader political economy of AI—to the contrary.[4] Yet the models that typically underlie theories of AI's impacts on workers could benefit from a sustained research focus on AI and worker power. Notably, worker power is not discussed in the survey of economic research on AI by Abrardi, Cambini and Rondi (2022).

This paper is structured as follows: Section 2 provides conceptual background on the definition of AI and distinctions between the labor demand and worker power impacts of AI. Section 3 reviews the literature around automation and AI in job tasks or the content of work, with consequences for labor demand. Section 4, the core of the paper, discusses and formalizes the ways AI might affect the context of work and worker power. Section 5 reviews empirical findings around AI in the workplace. Section 6 concludes and points towards future research needs.

2 Conceptual Background

2.1 What is AI?

Defining AI presents several challenges. One concern is whether and to what extent one should distinguish and enumerate subfields of AI (large language models, deep learning, computer vision, etc). Another issue is how strictly to tie AI to notions of genuine human-level intelligence and, however defined. There are workable, if lengthy, definitions from official bodies such as the OECD[5] that attempt to be both comprehensive and explicit about the types of technologies that fall under the AI umbrella.

I offer a more modest definition of AI, one that echoes past contributions in economics (Webb, 2019; Brynjolfsson, Li and Raymond, 2023). Digital systems have AI capabilities if they "learn" to complete tasks by identifying patterns in data rather than following explicitly programmed instructions. These tasks can include prediction capabilities (scoring resumes of job applicants), generative capabilities (pro-

and bias. Increases in productivity owing to better managerial allocation of resources and talent may be passed through to workers. Hiring algorithms could enable better worker-firm matches, benefiting both parties. See also Autor (2024).

[4] See, e.g., Agrawal, Gans and Goldfarb (2018); Brynjolfsson (2022); Acemoglu and Johnson (2023).

[5] "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment." (OECD, 2024).

ducing written output or code), and any number of other domains where sufficiently large amounts of data can be fed into a machine learning or similar algorithm. The key difference from other sorts of automation is that the task has not been broken down into a series of explicitly programmed steps, as might be the case, e.g., with an industrial machine or a simple regression-based prediction system.

The definition above points to a novel aspect of AI: its potential to automate non-routine tasks. Scholars of technological change distinguish between routine and non-routine tasks in identifying the sorts of jobs exposed to automation. A task is routine "if it can be accomplished by machines following explicit programmed rules," which exposes it to potential automation (Autor, Levy and Murnane, 2003, 1283). Jobs centered on non-routine tasks largely withstood technological disruption over the centuries and into the computer age. Yet AI may interrupt that trend. As Autor, Levy and Murnane wrote in their seminal 2003 paper, "[t]here are very few computer-based technologies that can draw inferences from models, solve novel problems, or form persuasive arguments," with additional non-routine tasks including "[n]avigating a car through city traffic or deciphering the scrawled handwriting on a personal check" (Autor, Levy and Murnane, 2003, 1283–1284). Clearly, times have changed.

Another way of distinguishing AI from its technological predecessors is its ability to assimilate *tacit* knowledge, or knowledge about which "we know more than we can tell," in Michael Polanyi's famous construction. Polanyi's paradox, as Autor (2014) called it, helps explain why certain kinds of jobs persist despite low skill requirements while others are complemented rather than automated by technology: an inability to lay out all the steps needed to complete certain tasks that humans do intuitively.

Tacit knowledge exists across the skill distribution. Physical tasks that children complete with ease stymy engineers. At the other end of the spectrum, delicate communicative and creative skills in office settings continue to require human labor (even as other office work, like adding tables of figures and copying memos, has been computerized). Automation runs up against our inability to enumerate the rules one should follow in order to, say, catch a ball or respond professionally to an email. Yet new AI tools show that many outputs of tacit knowledge can now be automated. Where many forms of tacit knowledge are concerned, we no longer need to tell computers what to do, only show them.

Recent years have seen a proliferation of AI methodologies and functionalities, the extent of which is too various and the distinctions too porous to fully outline here. Broadly, however, it is worthwhile to distinguish between predictive analytics, which includes recommendation algorithms and other forms of prediction, and generative AI, which refers to AI applications that analyze and generate text and images.

2.2 What is work?

Studying the potential impacts of AI on workers requires attention to both the content of work and the context of work. Conceptually, it is important to distinguish between the tasks that make up a job and the range of factors that influence the nature of work, from physical setting to social environment to organizational structures. While research concerned with AI and work typically focuses on job content, AI has the potential to dramatically change its context as well, with consequences for worker power.

Job Content

The content of a job refers simply to what a worker does: the tasks they perform. Crucially, jobs always consist of a bundle of tasks (Autor, Levy and Murnane, 2003). Assembly-line workers not only feed and convey materials but also monitor production and report issues. Radiologists interpret X-rays but also communicate with patients, confer with doctors, and decide whether more testing is needed.[6] Jobs may vary in the breadth of work, or the number of tasks that are involved, as well as the routineness and cognitive demands of their tasks, the way work activities fit together, and the centrality of the main task(s).

Another aspect of job content is to what degree workers take part in the planning and organization of what they do. In the schema of Braverman (1998), work entails both conception and execution: the planning and the doing. In agency theory, workers may be delegated more or less authority to choose which tasks or projects to execute (Aghion and Tirole, 1997). Different job designs may require varying degrees of conception or authority on the part of the worker, depending on factors including production technologies and organizational structure.

Job Context

While job content focuses on the specific activities workers perform, job context covers the broader circumstances and conditions that shape how that work is done. It encompasses the management practices that guide and oversee workers, the organizational structures that define their roles and relationships, and the social dynamics that influence their interactions. Workplace technologies impact job context by shaping the way workers communicate, collaborate, and produce. A vast array of contextual factors might be considered when analyzing job context. The occupational information database O*NET, for instance, records data on more than 50 aspects of job context, ranging from use of email to exposure to extreme temperatures to the freedom to make decisions (National Center for O*NET Development, 2024).

Since this paper concerns worker power, I will focus only on those aspects of job context that relate specifically to management. Management can be divided into four basic functions: staffing, direction, evaluation and monitoring.[7] Staffing involves recruitment, designing job offers, hiring workers and making retention decisions. Monitoring consists of observing workers at their tasks and assessing how much effort they expend. Direction refers to assigning responsibilities, scheduling tasks, and determining the scope of control workers have over their labor. Evaluation includes gauging the quantity or quality of work output, communicating performance metrics, and making judgments about discipline and incentive compensation. Each function of management has significant ramifications for the context of work. Each may also evolve to incorporate emerging AI tools.

Worker Power

The final piece of the conceptual framework to define is worker power, which has proved to be a slippery concept in the literature. Stansbury and Summers (2020, 3) delineate between monopoly and monopsony power on one side and worker power on the other, with the latter arising from "unionization or the threat of union organizing, from firms being run partly in the interests of workers as stakeholders, or

[6] The range of tasks radiologists perform helps explain why AI pioneer Geoffrey Hinton's infamous 2015 prediction that radiologists would be replaced in five years' time proved so inaccurate (Agrawal, Gans and Goldfarb, 2019*a*).

[7] This delineation borrows from a few alternative enumerations of the roles of management (with an particular focus on algorithmic management). Adopting the mechanisms of control laid out by Edwards (1979), Kellogg, Valentine and Christin (2020) divide management's role for workers into direction, evaluation and discipline. Nurski and Hoffman (2022) specifies the functions of management as goal specification, task specification, planning, incentivizing, and staffing. Parent-Rocheleau and Parker (2021) list monitoring, goal setting, performance management, scheduling, compensation, and job termination.

from efficiency wage effects." More broadly, worker power might be thought of as the sum of institutions, organizational structures, laws and norms that affect how employment rents are split between workers and employers.

As the above definition suggests, it can be unclear as to what falls under the heading of "power." What about globalization? Business cycles? For purely illustrative purposes, a clarifying model is the employment relationship represented in an asymmetric Nash bargaining framework. The bargaining process yields a wage w^* that solves

$$
\begin{aligned}
w^* &= \arg\max \overbrace{(w-b)}^{\substack{\text{worker}\\\text{surplus}}}{}^{\alpha}\,\overbrace{(p-w-d)}^{\substack{\text{firm}\\\text{surplus}}}{}^{1-\alpha} && (1)\\
&= \alpha(p-d)+(1-\alpha)b && (2)
\end{aligned}
$$

where w is the wage, b is the worker's outside option (e.g., the expected payoffs to continued job search), p is productivity, d is the value of the outside option for the employer, and α denotes the worker's bargaining power. This paper is largely concerned with how AI and workplace technologies may affect α.

In this setup, labor-substituting automation primarily affects the bargaining *position* of each party, which has to do essentially with the content of work. New (labor-displacing) technology may increase the value of the employer's outside option d. Alternatively, the aggregate displacement effects of AI could diminish the worker's outside option b since finding another job is harder when fewer skill-specific job openings exist. Both of these changes to bargaining positions push down the negotiated wage but leave bargaining *power* unaffected. Although such shifts in bargaining position are sometimes characterized as changes in bargaining power, I aim to retain the distinction.

The determinants of the bargaining power parameter α are somewhat more abstract. Since the Nash solution was first described and applied to labor contexts, the term has been described as "thoroughly vague," "imprecisely defined" and a "black box" (Cross, 1965; Binmore, Rubinstein and Wolinsky, 1986; Cahuc, Postel-Vinay and Robin, 2006). Theoretical models tend to treat bargaining power as exogenous, driven by factors such as cultural shifts or changes in the legal environment. I treat bargaining power as a function of the organizational structures, technologies and norms that pertain to bargaining and wage-setting decisions—the context of work.

AI-related organizational and technological shifts might have the effect of reducing α though a number of means, including: reorganization of the workplace that erodes fairness norms, solidarity, and the effective exercise of voice; enhanced monitoring which shifts rents to employers through efficiency-wage effects and chills organizing efforts; and heightened employer bargaining abilities.[8] Section 4 examines these possibilities and others in more detail.

3 AI, the Content of Work, and Labor Demand

The dominant approach in the economic theory of AI and labor focuses on technological changes to job content and the potential impacts on labor demand. In most cases, considerations of worker power—or bargaining power in contrast to bargaining position—play no role. Although in reality there are likely to

[8] The effect of AI-enabled wage-setting and wage-offer systems might be better thought of as heightening the information asymmetries in an imperfect information strategic bargaining game rather than a change in the power parameter in an axiomatic Nash setup, which assumes perfect information.

be overlaps between job content and context, or labor demand and worker power, these distinctions are useful when discussing the task model and related frameworks common in the economics of AI.

The foundation for the contemporary economics of technology and labor is the literature on skill-biased technical change (SBTC). Motivated by observations of rising educational wage inequality (Katz and Murphy, 1992), the first wave of SBTC literature investigated whether the 20th-century technological developments boosted demand for higher-skilled workers over others. The motivating theoretical model posited a production function in which factor-augmenting technological change differentially raises the productivity of certain groups of workers (i.e., the college-educated), boosting their relative wages (Autor, Katz and Krueger, 1998; Acemoglu, 1998). Empirical analyses for the mid- to late-20th century found support for this picture, though debate persisted over the degree to which distributional shifts were driven by institutional factors like deunionization and the declining real value of the minimum wage (Card and DiNardo, 2002; Autor, Katz and Kearney, 2008).

The successor to SBTC was the task model, still the dominant approach for theorizing the aggregate effects of technological change. The task model met the need for a model that could rationalize the post-1980s trend of labor market polarization: rising wages and employment at the high and low ends of the wage distribution with a flattening in the middle. Pioneered by Autor, Levy and Murnane (2003), the task model conceives of jobs as consisting of a bundle of tasks, differentiating between routine and non-routine tasks and across both cognitive and manual domains. Technology complements non-routine tasks—those requiring creativity, problem-solving, physical coordination and interpersonal skills—while substituting for routine ones.[9] Depending on the task makeup of jobs and where routine tasks are concentrated in the skill distribution, this can result in disparities in labor demand. Empirical work linking routine task intensity within occupations to labor market outcomes have had success in matching the predictions of the theory (Acemoglu and Autor 2011; Goos, Manning and Salomons 2009; Restrepo 2023; for a critical view, see Mishel, Schmitt and Shierholz 2013). In these studies, job polarization reflects substitution of routine middle-skill occupations by computers (e.g., clerical jobs) as jobs heavy on non-routine tasks enjoyed technological complementarities.

Refinements of the task model help explain why, after centuries of labor-saving automation, there are still so many jobs (Autor, 2015). Automating technologies displace workers, but they also have two important counteracting tendencies: a productivity effect, by which the increased efficiency of remaining workers raises labor demand, and a reinstatement effect, in which new technologies call forth entirely new tasks requiring human labor (Acemoglu and Restrepo, 2019*b*).[10] The rise of ATMs might have displaced some bank tellers, but it also prompted the expansion of bank branches due to increased productivity and gave rise to new lines of work such as ATM installation and repair (Bessen, 2015). As Autor et al. (2024) find, roughly 60% of workers in 2018 had job titles that did not yet exist in 1940. The aggregate effects of automation are thus ambiguous.

Much of the new scholarship on the rise of AI adopts some form of the task model (Autor 2022; see, e.g., the contributions in Agrawal, Gans and Goldfarb 2019*c*). While AI is novel in the types of tasks it can automate—non-routine and creative cognitive work among them—task-based models can represent this as a straightforward shift upward in the part of skill distribution exposed to automation (though

[9] An important difference between the task model and SBTC automation is possible in the task model; the latter only models differential increases in productivity across skill groups. The task model thus opens up the possibility for real wage *declines* for groups of workers (Acemoglu and Autor, 2011).

[10] There are other second-order effects: capital deepening raises demand for workers in the capital goods–producing sector; deepening of automation raises productivity without displacing any additional workers Acemoglu and Restrepo (2019*a*).

AI's effects might be represented in multiple other ways within the task model; see, e.g., Acemoglu 2024). Numerous contributors have adopted the view from the task model that the aggregate effects of AI on labor markets depend essentially on the balance between the number of jobs displaced and the counteracting productivity and reinstatement effects (e.g., Acemoglu and Restrepo, 2018; Dell, Nestoriak and Marlar, 2020; Klinova and Korinek, 2021).

Other theoretical models have been offered. Agrawal, Gans and Goldfarb (2019*b*) present AI as a form of prediction that can either complement or substitute for human judgment depending on the environment; the ultimate labor market effects are ambiguous (Agrawal, Gans and Goldfarb, 2019*a*). In a similar vein, Athey, Bryan and Gans (2020) use a principal-agent model to show how a principal's delegation of decision rights to AI versus human agents changes according to the (mis)alignment of the parties' interests or the accuracy of the AI system. Korinek and Stiglitz (2019) explore a range of possible political-economic consequences of AI, including the role of innovator rents and directed technical change.

The task framework and related conceptual approaches provide a number of insights about the potential effects of AI on workers. Acemoglu (2021) lays out several "harms of AI" to labor, including the possibility of excessive automation, which reduces the labor share but barely increases productivity, with particularly adverse results when there are existing labor market imperfections. Using simpler toy models, Acemoglu considers two other possibilities: first, that by pushing workers to specialize in a smaller range of tasks, AI could reduce economies of scope in on-the-job learning, with negative second-order effects on productivity. He also discusses the potential harms of excess monitoring, which I return to below.

There is some debate about the balance of potential harms and benefits to workers owing to AI task displacement. Working from the view that distributive outcomes depend whether AI primarily automates or augments work, Brynjolfsson (2022, 273) warns against a potential "Turing trap" in which businesses and developers seek AI applications that replace human labor rather than complementing it. This risks an equilibrium where economic and political power is concentrated in few hands and the disempowered majority lack means to improve their outcomes. Against this view, Agrawal, Gans and Goldfarb (2023) hypothesize a potential "Turing transformation" in which the labor-augmenting effects of AI enhance the job prospects of lower-skilled workers. Autor (2024) suggests that the informational benefits of AI could usher more workers into the ranks of high-paying jobs requiring decision-making and expertise.[11]

Although some contributions within the task-based framework leave room for considerations of worker power or monopsony power,[12] these aspects of the literature are typically secondary or absent altogether. Because the dominant approach to AI in labor economics centers on the content of work and effects that feed through the demand for labor, considerations of worker power take a backseat. This may not entirely incidental: as the SBTC literature matured, it was cast as an alternative to institutional and political explanations of rising inequality. The advance of AI prompts renewed consideration of worker power in the context of technological change.

[11] Both Autor (2024) and Agrawal, Gans and Goldfarb (2023) draw inspiration from studies that show a potential for AI tools to reduce skill gaps among workers (see Section 5). Acemoglu (2024) relies on the same studies but reaches less optimistic conclusions.

[12] Acemoglu (2021) addresses labor market imperfections and monitoring. See also the working paper of Azar et al. (2023), who build a task model in the style of Acemoglu and Restrepo (2018) with an imperfect labor market, showing that monopsony likely exacerbates the negative effects of automation for affected workers.

4 AI, the Context of Work, and Worker Power

To explore how AI affects the context of work and worker power, it is necessary to look beyond the task framework. The intertwining of technology and power is not a new notion in economics. Technological change played a key role for Marx, who saw automation feeding the reserve army of labor via displacement effects and located the source of capitalists' power in the "hidden abode of production" rather than in the idealized market for labor (Marx, 1992, 279). This emphasis on control in the workplace encouraged a literature on the *labor process*. Braverman (1998) explored how Taylorism and twentieth-century organizational strategies shaped workplace power relations. In Braverman's account, scientific management techniques degraded work by separating the conception and execution of production tasks via deskilling technologies. Edwards (1979) traced how managerial strategies to exert control within the "contested terrain" of the workplace evolved from the industrial revolution through contemporary capitalism. In both accounts, technical advance has as much to do with power relations as with efficiency.[13]

Key contributions from the burgeoning literature on AI and work in sociology and organizational studies draw from the labor process approach. Kellogg, Valentine and Christin (2020) adopt the typology of managerial control delineated by Edwards (1979)—direction, evaluation and discipline—to explore how algorithmic systems affect workers and workplace relations. Griesbach et al. (2019) examine platform work through the lens of "algorithmic control," a successor to "technical" and "bureaucratic" modes of control outlined by Edwards (1979). A broader multidisciplinary literature investigates algorithmic management (Lee et al., 2015), conceived as the use of data and predictive analytics to automate aspects of workforce management and coordination (Parent-Rocheleau and Parker, 2021).

A common thread running through research on algorithmic management and workplace technology is that AI presents novel opportunities for managers to exercise power over workers (Kellogg, Valentine and Christin, 2020; Jarrahi et al., 2021; Bernhardt, Kresge and Suleiman, 2023; Nurski and Hoffman, 2022). This perspective is echoed in the growing legal scholarship around AI and workplace technologies (Bales and Stone, 2020; Benkler, 2022; Kim, 2022; Rogers, 2023). The wide array of AI tools that could reshape power relations has been enumerated in white papers and policy reports that emphasize the potential for these tools to alter the nature of work and the balance of power in the workplace (Kresge, 2020; Mateescu and Nguyen, 2019; Negrón, 2021).

In economics there has been little corresponding theoretical or empirical work exploring connections between AI and worker power.[14] The rest of this section examines specific ways AI and emerging technologies may affect worker power through the functions of management: monitoring, staffing, direction and evaluation. In each case I present simple economic models to illustrate the phenomena of interest.

4.1 Monitoring

Among workplace AI tools relevant for worker power, monitoring is perhaps the easiest to address with existing economic theory. Although employee monitoring has been a historical constant in modern

[13] See also Marglin (1974), who explored the hypothesis that the industrial revolution's turn towards factories was motivated by entrepreneurs' desire to exert more control over workers (and see Landes [1986] for a response). Additional perspectives on the labor process literature can be found in Spenner (1983) and Smith and Thompson (1998).

[14] An apparent exception is Acemoglu and Johnson (2023)'s *Power and Progress*, which explores how social and political power dynamics shaped (and were shaped by) technological change over the past millennium. Their emphasis on power, however, centers on the question of whose vision drives the course of technological development. It is less concerned with workplace power dynamics, though monitoring is discussed in some detail (see pgs 320–324). While the emphasis on vision and directed technological change provides an important perspective, it differs from the primary focus advanced here.

workplaces, advances in computer vision and machine learning make virtually constant observation a viable management technique. Roughly half of U.S. workers report being monitored by technology at work (American Psychological Association, 2023). Hundreds of workplace tech providers now offer tools to monitor inputs ranging from email text to GPS position to heart rate (Negrón, 2021; Nguyen, 2021; Ajunwa, 2023). AI-enabled systems monitor delivery workers' speed and driving habits, including how often they buckle their seat belts, glance away from the road, and take sips of coffee (Royle, 2023). Apps track remote workers at home, logging their keystrokes, reporting on their web and app usage, and in some cases capturing continuous webcam feeds (Thier, 2023). Mood analytics tools infer emotional states from tone of voice and facial expressions (Andalibi, 2024). Beyond the workplace, employers trawl social media and corporate wellness programs track home fitness (Ajunwa, Crawford and Schultz, 2017). Even the generalist ChatGPT4 model, suitably prompted, can take in photos of workers and provide apparently accurate judgments as to whether they seem to be working or resting.[15]

Workplace monitoring is typically understood through efficiency wage theory. The shirking model, a cornerstone of the efficiency wage literature, considers a situation in which employers desire a certain level of effort from workers yet cannot perfectly monitor effort. Employers dismiss workers who supply less effort than desired (those who "shirk"), but not all workers can be monitored all the time. A straightforward way for employers to address this agency problem is to pay higher wages, increasing the cost of job loss to workers and compelling them to exert desired effort levels. The imperfections of workplace monitoring thus confer employment rents upon workers. In the resulting equilibrium, unemployment can be thought of as a component of worker discipline, since higher unemployment devalues the workers' outside option and increases the cost of job loss (Shapiro and Stiglitz, 1984; Bowles, 1985). It is straightforward to model AI enhancements to monitoring systems as a reduction in the unit cost of monitoring inputs.

Efficiency wage models have direct implications for distributive outcomes. In the standard shirking model, where exerting effort is an all-or-nothing decision, wages and monitoring technologies emerge as substitutes. When the effectiveness of monitoring rises, employers effectively capture employment rents from workers (e.g., Acemoglu and Newman, 2002). Yet this result relies on a particular set of formal assumptions, chief among them being that workers face a binary work-or-shirk decision and that the level of monitoring is exogenous. As Allgulin and Ellingsen (2002) show, relaxing these assumptions in a static setting produces equilibria in which wages and monitoring costs are typically complements rather than substitutes. When standard functional forms are posited for the disutility of effort and the firm's production function, a fall in monitoring costs counter-intuitively raises wages.[16] The efficiency-wage effects of increased monitoring technologies rest on softer foundations than is often presumed.

I extend Allgulin and Ellingsen (2002) to incorporate the notion that higher effort levels reduce the probability of being dismissed in a continuous fashion rather than via the step function typically assumed in shirking models. In the standard approach, the interaction of firm and workers strategies produces an incentive compatibility constraint according to which the probability of dismissal is 0 when a worker supplies effort of at least $\hat{e}$. Just below that effort level, the probability of dismissal is p, the probability that a worker's effort is observed. It may be more realistic to assume that the probability of dismissal varies smoothly with effort. This is the case, for instance, in the labor discipline model of Bowles (1985),

[15] Author's correspondence with ChatGPT4 (January, 2024).

[16] The reason is that workers must be compensated for additional effort. A fall in monitoring costs may make it profitable for firms to raise effort and production levels together via greater monitoring, leading to higher output as well as higher wages.

who treats effort e as a share of a worker's time on the job spent actively working, as well as in Ritter and Taylor (2011), in whose model firms receive a noisy signal of worker effort.

The basic setup, described fully in Appendix A, is sketched here. Workers optimize according to the utility function

$$U(e) = -C(e) + (1 - q(e,p))w + q(e,p)\bar{w} \tag{3}$$

where $C(e) > 0$, with $C'(e) > 0$, represents effort costs. The dismissal function $q(e,p)$ denotes the probability of being dismissed, where $q_e < 0$ and $q_p > 0$. The wage paid if the job is retained at the end of the period is w and $\bar{w}$ is income received if let go. The first-order condition can be rearranged to yield a wage function $w(e,p)$, which can be interpreted as the wage required to bring forth the effort e at monitoring level p.

Firms maximize static profits given by

$$\Pi(e,p) = G(e) - \mu M(p) - w(e,p) \tag{4}$$

where $G(e)$ is production and $\mu M(p)$ is the cost of monitoring at level p, with $M'(p) > 0$. The parameter μ, denoting the unit cost of monitoring, may be reduced through exogenous technological change. From the firm's first- and second-order conditions optimizing with respect to e and p, it can be show that the change in wages for a change in the cost of monitoring $\frac{dw}{d\mu}$ takes the general form

$$\frac{dw}{d\mu} = \frac{M'(p)}{D}\left[\frac{\partial w}{\partial e} \cdot \frac{\partial^2 w}{\partial e \partial p} + \frac{\partial w}{\partial p}\left(G''(e) - \frac{\partial^2 w}{\partial e^2}\right)\right] \tag{5}$$

where $D := \Pi_{ee}\Pi_{pp} - \Pi_{ep}^2 > 0$ (from the second-order conditions on firm optimization).

The sign of $\frac{dw}{d\mu}$ depends on specifications for the underlying functions: the cost of effort $C(e)$ (which enters into the partial derivatives of the wage function), the probability of dismissal $q(e,p)$, and production $G(e)$. In some cases wages and monitoring costs are unambiguously complements or substitutes, while in others the sign of $\frac{dw}{d\mu}$ varies. The two cases outlined in Appendix A demonstrate straightforward settings in which wages and monitoring costs are substitutes and thus wages fall when monitoring becomes cheaper. The production function $G(e)$ plays a key role in each case. Production techniques that are linear or concave in effort inputs are more likely to make wages and monitoring substitutes. The convexity of worker effort costs also affects the solution. When, following Bowles (1985), $e \in [0,1]$ is interpreted as the share of time workers spend in production and the dismissal function is $q(e,p) = p(1-e)$, the condition for $\frac{dw}{d\mu} > 0$ is $C''(e)^2 < C'(e)C'''(e)$.[17] This condition is not met with a simple power function $C(e) = ce^d$ but it does hold when the cost of effort is asymptotic at $e = 1$, such as with the function $C(e) = \frac{c}{(1-e)^d}$.[18]

Although the model presented above somewhat complicates the basic shirking model, its conclusions align broadly with those of the existing literature. Imperfections in monitoring boost the pay of workers above their outside options, conferring a form of employment rent. By the same token, improvements in monitoring technologies shift these rents away from workers. It is worth noting that even when wages and monitoring costs are complements, a fall in monitoring costs may increase functional income inequality. From the envelope theorem, the change in profits for a reduction in monitoring costs is $M(p)$.

[17] This assumes a linear or concave production function.

[18] This type of function meets the conditions specified for effort functions in canonical agency theory models, such as that of Aghion and Tirole (1997).

If $M(p) > \frac{dw}{d\mu}$, improved monitoring will increase profits more than it will increase wages. Similarly, in the case that $\frac{dw}{d\mu} = 0$, wages may not decrease in absolute terms when monitoring improves, but worker welfare will fall since increased monitoring prompts higher (uncompensated) effort levels.

The model offered here is of course only a rough approximation. Further insights may be gleaned by making dynamic or by introducing firm competition. One intriguing extension would be to posit different worker types, sectors, or occupations. In the work of Bulow and Summers (1986), among others, dual-labor-market dynamics are shown to emerge from differences in monitoring capabilities between otherwise identical job types. Thus harder-to-monitor jobs garner higher wages. As AI lowers the costs of monitoring more knowledge-intensive work activities, occupations previously immune to invasive oversight may see shifts in wages and the nature of the job. Speculatively, if monitoring intensity changes more for high-skill than low-skill jobs, the result could be a counterintuitive decline in wage inequality.

AI-driven advances in employee monitoring have additional implications that go beyond the efficiency-wage framework. First, the experience of surveillance may harm workers' well-being by itself, leaving aside reduced efficiency wages or higher effort requirements. Reviews of the management literature find that electronic monitoring tends to increase stress and diminish job satisfaction (Ball, 2021; Siegel, König and Lazar, 2022; Ravid et al., 2023). Second, AI monitoring tools could play a role in helping firms forestall union organizing (Bernhardt, Kresge and Suleiman, 2023). Large companies already contract with businesses that use predictive analytics to identify establishments at risk of organizing activity (Vogel, 2021). Amazon reportedly uses in-house software systems to monitor union organizing threats (Rey and Ghaffary, 2020). These sorts of tactics have raised legal questions and attracted the scrutiny of the National Labor Relations Board (Garden, 2018; Rogers, 2023; National Labor Relations Board, 2022). More broadly, the acceptable scope of AI-enhanced monitoring in and around the workplace remains an open question among legal scholars (Ajunwa, Crawford and Schultz, 2017; Bales and Stone, 2020).

Smarter technological monitoring may also hold benefits for workers. One upside is a potential reduction in managerial caprice and favoritism. A "tough but fair" algorithmic monitor may be preferable to an office tyrant. Workers may also capture a portion of the productivity gains that arise from shifting labor away from non-productive monitoring activities, particularly in competitive labor markets or settings where workers have institutional levers of power.

4.2 Staffing

Staffing represents the second broad area where AI could alter worker power. There are numerous managerial processes related to hiring and firing where AI might be used: recruitment, resume screening, wage offer, contract design, and retention decisions. In a 2024 survey of HR professionals by the Society for Human Resource Management (SHRM), 26% of respondents reported using some form of AI in HR-related activities; among these, 64% used AI in recruitment, interviewing, or hiring (SHRM, 2024).[19] Some AI systems are already well-established in the hiring process. Job boards such as LinkedIn employ machine learning algorithms to match job seekers to openings and suggest candidates to recruiters (Guo, 2019). AI-enabled tools that screen resumes based on factors like keywords and gaps in one's work history are widely used despite broad public concern (Chen, 2024; Rainie et al., 2023; Starace and Horn, 2024).

AI is less prevalent in other staffing functions, such as pay-setting and retention. Although there have

[19] Among respondents, the most common use cases of AI in staffing functions were generating job descriptions, targeting job ads, and reviewing resumes.

been reports of AI used in layoff decisions (Verma, 2023), just 1% of HR managers in the SHRM survey claimed to use AI software in these situations. AI also appears to play a minimal role in wage-setting , with the notable exceptions of the major platform work companies (see below). Another exception is AI pioneer IBM, which has made its HR department "client zero" for deploying AI internally; use cases include directing workers to career opportunities and answering questions around company policy (Armstrong et al., 2024). IBM also utilizes a compensation advisor system that offers managers recommendations as to whether employees should receive raises and, if so, the size of the pay bump. A company report notes that "managers tend to follow the recommendations the AI provides, and this has helped ensure employees are not overpaid or underpaid at IBM" (Guenole and Feinzig, 2022, 8).

The use of AI in staffing functions is perhaps most advanced among app-based delivery and ride-hail platforms. Gig-work companies automate most of the hiring process and pay is determined in large part algorithmically (Lee et al., 2015). The opacity of the algorithms and lack of human managerial interaction creates steep information asymmetries between gig workers and their employers (Rosenblat and Stark, 2016). When ridehail workers receive pickup requests, they typically see the proposed payment and pickup location, but not the destination. The determinants of fluctuations in fares can be difficult for workers to intuit. According to California gig workers interviewed by Dubal (2023) patterns in fare offers make platform work seem "like you are being manipulated" and "like gambling," impressions that point towards a broader trend towards gamification of rewards at work (Kresge, 2020). Another worker perception is that personalized wage setting in gig-work settings provides a means of mapping out workers' reservation wages in order to minimize wage costs. As Griesbach et al. (2019, 6) found: "Many Instacart workers with whom we spoke believe that the algorithm 'learns' the lowest rate it can successfully offer for an order within a particular region at a particular time and day of the week."

Bargaining and negotiation

Better targeting of job seekers' reservation wages provides an instructive example for the way AI tools in hiring and pay-setting are poised to affect worker power. Roughly one-third of jobs involve wage bargaining (Hall and Krueger, 2012). In wage negotiations, the worker's outside option is private information. It is a standard result in bargaining theory that this form of information asymmetry strengthens the bargaining power of the selling party.[20] In making a wage offer, a prospective employer must work from a prediction of the job seeker's current reservation wage. By aiding in prediction tasks (Agrawal, Gans and Goldfarb, 2019*b*), AI hiring tools have the potential to weaken workers' information advantage and shift employment rents to the firm.

A simple partial-equilibrium model can illustrate the effect of improved firm predictions of worker reservation wages. Consider a labor market in which a firm meets job seekers and makes take-it-or-leave-it wage offers. A filled job produces p for a single period while an unfilled job produces 0. Workers are identical in all ways save their reservation wage b, which is distributed according to $F_b(b)$. Firms observe workers' reservation wages with some noise so that the observed reservation wage is $z = b + \varepsilon$, with ε drawn from a distribution $F_\varepsilon(\varepsilon)$ that is symmetric around $\varepsilon = 0$. A firm that always offered a wage equal to z would fail to hire half of the workers it met, incurring vacancy costs (foregone production). It may be more profitable to offer a wage $w(z, \phi) = z + \phi$ in order to increase expected production.

[20] See, e.g., Chatterjee and Samuelson (1983) or Muthoo (1999, ch. 9); some exceptions exist. As noted by Manning (2013, 136), the prevalence of this kind of private information on the worker's part likely encourages employers to make a policy of wage posting rather than negotiation.

As detailed more fully in Appendix B, the firm chooses a wage policy parameterized by ϕ^* that maximizes expected one-period profits:

$$\Pi(\phi) = \int_{-\infty}^{p-\phi} (p - \phi - z)(1 - F_{\varepsilon|z}(-\phi))f_z(z)dz \tag{6}$$

The first term in parentheses denotes revenue less wages. The second term, $(1 - F_{\varepsilon|z}(-\phi))$, is the probability that the wage offer $z + \phi$ exceeds the worker's reservation wage, conditional on the value of z that is observed (since $z = b + \varepsilon$, z contains information about the possible value of ε). The expectation is taken over z since this variable is what the firm observes.[21] Although the first-order condition is simple to characterize, it consists of an integral that is generally intractable. Figure 1 in Appendix B plots a range of numerical solutions for ϕ^* assuming $b \sim N(\mu_b, \sigma_b)$ and $\varepsilon \sim N(0, \sigma_\varepsilon)$. In all cases, ϕ^* rises monotonically with σ_ε. An improvement in the firm's ability to predict the job seeker's reservation wage reduces the average wage offer.

This simple model abstracts from second-order effects that might mitigate the losses workers face when their information rents are eroded. One consideration is that when firms better target their wage offers, a greater share of those offers are accepted. Appendix Figure 1 (second row) indicates that the share of worker-firm encounters that lead to consummated matches rises as the noise in reservation wage observations falls. If this effect held across firms, the result would be tighter labor markets and greater upward pressure on wages generally, potentially offsetting the reduction in wages from workers' diminished information advantage in bargaining. Another possibility is that AI-powered compensation-setting tools help firms target salaries more effectively, reducing costly attrition by preemptively raising wages. HR technology providers now offer salary benchmarking services based on machine learning algorithms and large data sets containing salary and position data. Cullen, Li and Perez-Truglia (2022) study the rollout of such a tool for customers of the largest U.S. payroll processing company, finding that it reduced salary dispersion and raised starting salaries among lower-skilled positions.[22]

Worker screening

A related set of concerns surrounding AI in hiring has to do with firm screening on non-productive characteristics of workers. Job application systems now assess workers' personalities via AI-driven analysis of job interviews, social media footprints, and a range of other inputs (Kresge, 2020). For instance, Paradox.ai, which counts employers such as FedEx and McDonald's as customers, measures prospective employees' Big Five personality traits using image-based screening tests featuring a cast of blue-skinned humanoids.[23] These sorts of screening devices have implications for workplace power dynamics. While employers may understandably value hiring dependable and intelligent employees, they may also prefer acquiescent workers to assertive types (Kim, 2022). Personality traits and other apparently non-productive

[21] The lower bound of $-\infty$ of course implies that some workers will accept negative wages. The lower bound could instead be set at some minimum value w_0 without substantially changing the results. In the parameterizations presented in the appendix, $P(z < 0)$ is vanishingly small.

[22] In a theoretical model building on auction theory, Cullen, Li and Perez-Truglia (2022) show that the availability of salary benchmarks raises the expected wage in equilibrium, an effect driven by the marginal firm gaining a better understanding of the competitiveness of the market and consequently boosting its wage offer. It differs from the model considered here in that firms competitively bid for talent in a setting where individual firms lack information about the values other firms place on similar workers; reservation wages play no role.

[23] www.paradox.ai/products/assessments. Accessed 19 April, 2024. Maiberg (2024) reports on bemused social-media reactions to the tests.

characteristics reliably affect earnings and labor market outcomes (Bowles, Gintis and Osborne, 2001). Among the Big Five personality traits—openness, conscientiousness, extraversion, agreeableness, and neuroticism—the latter two are consistently associated with lower earnings (Alderotti, Rapallini and Traverso, 2023). The negative relationship between agreeableness and earnings can be explained most readily by the diminished appetite for bargaining and conflict among agreeable types (Judge, Livingston and Hurst, 2012; Flinn, Todd and Zhang, 2020).

Although the personality-earnings relationship is typically theorized at the level of individual workers, workplace personality composition likely has spillover effects across workers in a firm. HR departments invest heavily in understanding employee social dynamics to avoid outcomes like low morale and "turnover contagion" (Felps et al., 2009). Sapia.ai, another maker of AI-powered personality assessments, addresses this relationship explicitly in a blog post on turnover contagion: "Are you hiring people with the personality and behavioural traits that make them more likely to stay and perform in your company?"[24] Reducing inefficient turnover is an understandable goal for a company to pursue. Yet selecting for workers less likely to exercise the exit option may have consequences for worker pay across the workforce.

Consider a partial-equilibrium two-period model in which a firm hires a fixed number N of workers and commits to paying them a wage w bounded by a wage floor or commonly held reservation wage w_0. Workers come in one of two types, exiters and loyalists, with the share of exiters in the labor market denoted x_0 and the share at the firm x. We may think of loyalists as those who score highly on agreeableness in Big Five personality tests. The firm can pay to select for workers by paying a per-worker cost $c(x - x_0)^2$. Production takes two periods to complete. Between the first and second period, exiters sample from the exogenous outside job distribution $F(\tilde{w})$ and accept other jobs when $\tilde{w} > w$ (a simple extension would allow both types of workers to search on the job, only at different rates). In the second period, production is completed by the remaining workers, who number $N(1 - x\bar{F}(w))$, where $\bar{F}(\cdot) \equiv 1 - F(\cdot)$. Although all workers receive the wage w, output is determined by the number of workers remaining after the job search. Profit per worker is

$$\Pi(w, x) = A(1 - x\bar{F}(w)) - w - c(x - x_0)^2 \tag{7}$$

The comparative statics, presented in Appendix C, are straightforward: for interior solutions, a fall in the cost c of screening workers leads to reductions in both the equilibrium wage w^* and the equilibrium share of exiters at the firm x^*.[25] More cost-efficient screening allows the firm to shift the composition of its workforce towards workers less likely to expose it to wage competition. At the extreme, a firm that can costlessly filter out all exiters pays the minimum wage w_0 to all workers. By enhancing the ability of firms to predict job applicants' personality types, AI screening tools may diminish worker pay and dampen labor market dynamism.

This model is only partial, however, and may be extended in ways that alter its results. Most importantly, it concerns only a single firm and takes the outside wage distribution as exogenous. A fully specified search and matching model with two worker types and a continuum of firms could give rise to general equilibrium dynamics not contemplated here.[26] Exiters face two opposing forces: screening

[24] https://sapia.ai/resources/blog/how-to-diagnose-cure-and-prevent-turnover-contagion. Accessed 23 April, 2024.

[25] $F(\tilde{w})$ is specified as an exponential distribution truncated from below at w_0. This specification is not essential to the results.

[26] One might also specify a search model that replaces the wage posting assumption with a model where outside job

by employers but also greater mobility once employed. If the former outweighs the latter, employers hoping to hire only loyalists might end up paying greater vacancy costs, since loyalists have already been hired elsewhere. On the other hand, if exiters climb the job ladder more quickly, firms lower down the job distribution will face a labor pool dominated by loyalists. In either case, one can imagine dual-labor-market patterns emerging. The aggregate implications for workers are unclear a priori. Another possibility not modeled above is that better screening could improve worker-firm matches in productivity-enhancing and thus mutually beneficial ways. This may the case, for instance, with cultural fit (which of course raises separate concerns over equity and fairness).

Issues of AI bias and discrimination in hiring have received ample public attention and policy analysis. Machine learning tools trained on historical data can pick up patterns of discrimination which may then be inadvertently perpetuated under the guise of neutral algorithms. In a demonstrative example, Amazon abandoned an AI recruiting tool in 2017 that was found internally to disadvantage women's resumes due to the gender bias of the training data set, penalizing terms like "women" in "women's chess club captain" and the names of all-women's colleges (Dastin, 2018). With AI algorithms used at various stages of the hiring process, from sourcing to recruitment to resume review, bias has numerous opportunities to manifest (Bogen and Rieke, 2018).[27] Although vendors typically advertise the steps they take to mitigate bias, algorithmic auditing has yet to be standardized (Raghavan et al., 2020). There is a silver lining here: as researchers in this field often note, properly tested algorithmic screening and hiring tools have the potential to reduce bias relative to the counterfactual, since "biased algorithms are easier to fix than biased people" (Mullainathan, 2019).

On the supply side of the labor market, AI tools may give workers a leg up. Online job boards like LinkedIn and Indeed provide personalized recommendations to job seekers and surface a wider range of open positions than workers might see otherwise. Access to such tools may improve worker-firm matches and reduce unemployment spells (Bhuller et al., 2023). Worker-focused AI tools may also aid job seekers in negotiation. The site Glassdoor, a platform for workers to rate and review employers, uses its vast store of user-supplied data to offer personalized salary calculators.[28] Given the poor understanding workers typically have about the wage distribution they face in the labor market (Jäger et al., 2024), this information provides an advantage to workers' bargaining abilities and expectation formation. The information workers can glean from AI-powered job boards and review sites may help to counteract the effects of enhanced prediction and selection on the employer side.

4.3 Direction and evaluation

The final area where AI could affect worker power is in the direction and evaluation of work.[29] This set of managerial processes comprises how firms allocate tasks, delegate decisions, instruct workers, and assess outputs. Direction and evaluation form the core of what organizational psychologists call job design, "the content and organization of one's work tasks, activities, relationships, and responsibilities"

offers allow workers to extract raises from their current employers, as in Cahuc, Postel-Vinay and Robin (2006); different personality types here may entail different bargaining abilities.

[27] For example, the algorithmic audit of job advertisements by Imana, Korolova and Heidemann (2021) surfaced evidence of gender skew in Facebook ad delivery.

[28] https://www.glassdoor.com/Salaries/know-your-worth.htm. Accessed 23 April, 2024.

[29] Although direction and evaluation are two separate management processes according to the taxonomy laid out in Section 2.2, their relation is tight enough that it makes sense to group them into one for expository reasons. This also accords with the contract design literature following Aghion and Tirole (1997), in which delegation arrangements (direction) depend on how work projects are evaluated.

(Parent-Rocheleau and Parker, 2021, para. 2). The centrality of direction and evaluation to everyday work experience means that AI impacts in this area may be among the most keenly felt by workers. For Kellogg, Valentine and Christin (2020), technologically mediated direction and evaluation make up two of the three elements that define the "new contested terrain of control" within the workplace (the third being discipline).

Although some of the potential effects of AI on work and worker power in this section are quite speculative, firms have already introduced a range of AI-powered tools that direct or instruct workers in real-time and incorporate continuous, fine-grained feedback. The clearest examples of this trend can be found in the warehousing and logistics sector (Mehta and Levy, 2020). Algorithmic systems determine the optimal paths human workers should take in picking, packing, and loading goods. Barcode scanners, human-worn sensors, and handheld devices interact to provide instantaneous instructions and feedback so that workers' output can be continually evaluated and their next tasks assigned. These integrated algorithmic management systems tend to intensify work effort, giving rise to a struggle among employees to "make rate" (Gutelius and Theodore, 2019) and a sense of being treated like robots (Bell, 2022). As an operations manager at a distribution center informed Mehta and Levy (2020, 22), "What I would really like is software that keeps track of every person and every robot on the floor and tells each of them what it should do next."

Logistics is not the only industry where managers have begun to direct and evaluate workers via algorithm. Manufacturers have introduced human-worn augmented-reality lenses to provide real-time instructions on tasks (Moore, 2020). In hands-on field work employed on the shop floor of a German electronics manufacturer, Schaupp (2022) found his instructions placed on an electronic screen that updated with new tasks. The goal, according to a manager, was "either to make things go faster or to enable people with less qualifications to do it" (Schaupp, 2022, 17). In hospitality, task management apps direct custodians to hotel rooms in an effort to maximize their time spent cleaning (Mateescu and Nguyen, 2019). Other algorithmic management tools have been introduced in settings ranging from call centers to retail trade to marketing agencies (US Council of Economic Advisors, 2022; Wood, 2021).

Several challenges arise in mapping how AI innovations in direction and evaluation might affect worker power. The primary difficulty lies in identifying the effects of algorithmic management techniques on power dynamics while holding constant the task content and skill requirements of the job. As the managers interviewed by Schaupp (2022) indicated, a major appeal of dynamic instructional tools on the factory floor is that they allow workers with lower skill and experience levels to contribute to production.[30] AI-enabled tools that dynamically assign tasks to workers effectively automate a set of cognitive tasks centered on planning and time allocation. In the schema of Braverman (1998), these systems separate the conception of work from its execution. By subdividing work processes into discrete steps, providing directions and instructions in smaller chunks, and evaluating intermediate outputs rather than the final product, AI-powered direction-and-evaluation systems reduce standards for expertise and experience by diminishing the importance of allocative and planning skills among workers.

While undoubtedly a pressing concern for many workers, the deskilling brought about by task automation lies outside the framework of worker power defined at the outset. Diminished standards undermine

[30]The desire to arrange production so as to minimize skill and experience requirements goes back to the era of Taylorist scientific management, if not well before. An engineer at Ford's 1910 Highland Plant once remarked that the company's production process "desires and prefers machine-tool operators who have nothing to unlearn . . . and will simply do what they are told, over and over again, from bell-time to bell-time" (quoted in Montgomery, 1980, 119).

the bargaining *position*, though not necessarily the bargaining *power*, of incumbent and experienced workers. Although it is worth exploring the interplay of worker power and deskilling, the present discussion centers on the worker-power consequences of AI-powered advances in direction and evaluation holding constant the skill and task bundles associated with a given job. These effects fall into two broad buckets: the interplay of delegation and authority with monetary incentives; and the organization of the firm, with consequences for pay norms and worker voice.

Delegation and authority

The first way AI-powered direction and evaluation might affect worker power stems from the potential of algorithmic management tools to shift the balance of delegation and authority within the firm. Aghion, Jones and Jones (2019) posit that AI tools could allow firms to delegate more decisionmaking farther down the organizational hierarchy since AI makes it easier to reverse suboptimal decisions and allows for better output monitoring.[31] Athey, Bryan and Gans (2020) consider the delegation of authority between human and AI agents who operate in parallel (and, possibly, in competition). These are among a small but growing strand within the economics of AI drawing on contract theory, particularly the seminal Aghion and Tirole (1997).

The Aghion-Tirole framework is a principal-agent setup where each party can independently research the payoffs associated with a menu of projects whose benefits may differ between the principal and agent. The optimal contract depends on the relative difficulty each party faces in discovering payoffs and on the extent to which interests align or diverge regarding project preferences. The principal may retain formal authority over project selection, delegate it fully to the agent, or occupy a middle-ground where the agent has "real" authority and the uninformed principal rubber-stamps the agent's decisions.

Within the setup of Aghion and Tirole (1997), workplace AI technologies may be thought of as reducing the cost of acquiring information for the principal (manager) in a setting where production entails non-routine tasks such as planning and time management. Consider a fulfillment center that shifts from largely analog processes to a digital warehouse management system. In the prior setting, management lacks ready access to information that might inform the worker which "project" or work task to take on next: which aisle to walk down, which items to pick or pack and in what order, etc. Thus the real authority to choose and execute the next task rests with the worker, who is more familiar with what's happening at the ground level. Given the information asymmetry, management may structure pay such that workers are incentivized to carry out tasks in ways that maximize profit, even though those management-preferred tasks may also impose various costs on the worker relative to his preferred tasks.

In Appendix D, I present a straightforward application of the monetary incentive model of Aghion and Tirole (1997, section V.B) in which lower information acquisition costs for management reduces wage incentives offered to workers. The intuition is that when managerial information costs fall, management can rely more on its own effort, reducing the need for incentive payments to induce worker initiative.[32] Although the worker still carries out the same tasks, incentive pay falls as a result of changes in the

[31] To distinguish between the effort monitoring considered previously and the evaluation considered here, it is useful to see monitoring as observation of worker inputs into production (or effort) while evaluation consists of observing outputs, final or intermediate.

[32] It is interesting to note that the information-effort relationship described here reverses the conclusions of the efficiency-wage monitoring models, in which improved monitoring increases worker effort. The difference can be reconciled by noting that the effort or initiative considered here differs from that of the shirking class of models: In agency models, worker effort contributes information rather than direct output.

informational environment. Continuing with the example of the warehouse given above, new technologies allow management to more easily choose what tasks should be pursued. The inventory management system "knows" which items are where, or a machine learning algorithm has predicted the best route the worker should take to pick up all the contents of a container. By increasing the likelihood that management can choose the right task without worker input, the algorithmic management system reduces the need for monetary incentives that align worker-firm interests.

Some caveats and limitations should be noted. First, although the model posits an unambiguously negative effect of improved managerial information on the per-project incentive wage, it is less clear about the expected total compensation of the worker. If better managerial information makes management's preferred projects more probable, workers can expect to receive incentive pay more often. Ultimately it is ambiguous whether the worker's expected compensation rises of falls for a given change in managerial information costs. Second, in order to simply the analysis, the presentation abstracts from shifts in the information environment significant enough to eliminate incentive pay altogether. For certain regions of the parameter space—particularly (though not exclusively) when management's information costs fall below those of the worker—the firm optimally eschews incentive pay and makes worker initiative unimportant. Based on case studies from technologically advanced warehousing and logistics operations, this setting may be closer to reality. Workers in these highly structured environments receive precise directions throughout the work day and lack the autonomy of their counterparts in more analog warehouses.

There remains the question of whether the effects of technological change within the contract theory framework laid out above reflect changes in worker power per se. In the Aghion-Tirole framework, the "real" authority of workers constitutes command over decision-making in the workplace, a responsibility that certainly aligns with common-sense notions of power. Yet as noted above, the situation can be recast as a form of automation in which the firm substitutes machine intelligence for part of the worker's task bundle, reducing the demand for skills and experience related to decision-making and planning on the job. The result is a form of deskilling, frequent throughout the history of technology, whose primary effect is to reduce the leverage and bargaining position of incumbent workers.

Additional models might be brought to bear on the relationship between delegation, authority and worker power under AI. Agency theory offers a multitude of theoretical frameworks beyond that of Aghion and Tirole (1997) that may shed light on AI-driven management (e.g., Dessein, 2002). Another option, following Becker and Murphy (1992), is to model the range of tasks that make up a job together with the coordination costs required to combine them. In a short research note, Isztin (2023) draws on the Becker-Murphy model to predict how AI will affect the distribution of "specialists" and "generalists" in the labor market. In a similar fashion, one might see AI as reducing the costs associated with coordinating a worker's various tasks, allowing for more granularity in the direction and evaluation of work.

The organization of the firm

New patterns of delegation and authority brought about by AI can be expected to alter the organization of the firm, much as advances in information and communications technologies (ICT) have long done. Aghion, Jones and Jones (2019) posit that AI-powered technologies should lead to more decentralized and less hierarchical organizations because evaluation or output monitoring tasks previously done by mid-level managers can be automated, a prediction echoed by Abrardi, Cambini and Rondi (2022). Looking at the broader "workforce ecosystems" that surround firms, which include contract and temporary workers together with full-time employees, Kiron, Altman and Riedl (2023) associate AI with larger and flatter teams due to reductions in coordination costs. Ide and Talamàs (2024) use a version of a task assignment model to show how AI that replaces managers leads to larger, more decentralized organizations in which remaining human managers have a wider span of control, or number of direct reports.

An alternate view highlights the tendency of AI to centralize power (Brynjolfsson, 2022; Brynjolfsson and Ng, 2023). Drawing on the Grossman-Hart-Moore framework of decision rights, Brynjolfsson ponders the consequences when useful knowledge is made alienable from individual human minds by large-scale AI systems that can easily codify and digitize tacit understandings. Under certain assumptions, it may be optimal to centralize decision rights and asset ownership at the top of the hierarchy. This tendency poses political as well as economic risks: "When useful knowledge is inalienably locked in human brains, so too is the power it confers. But when it is made alienable, it enables greater concentration of decision-making and power" Brynjolfsson (2022, 277).

How do we reconcile the conflicting predictions of AI's effects on firm organization? First, it is worthwhile to separate two predictions that are sometimes implicitly equated: one about the hierarchical organization of the firm and the other about the delegation of authority. When a firm becomes "flatter" by reducing the ranks of middle management, decision-making authority is often devolved down into the lower ranks of the leadership structure. Yet this need not be the case. In highly automated warehouse settings, for instance, monitoring and evaluation technologies displace middle managers who oversee floor-level operations. Such a flattening does not mean authority has been delegated downward.

Being more precise about how AI is used also helps clarify predictions about its effects on firm organization. A valuable framework is the one outlined in Bloom et al. (2014), building on Garicano (2000), which distinguishes between the information and communications components of ICT. Consider a firm with a base layer of production workers and a higher layer of managers. Production workers encounter problems that range in complexity. Being generalists, workers can solve simple problems. But more complex challenges force a choice: They can either acquire the information independently at some cost (e.g., find and read a manual), or engage in costly communication with specialists in management who provide the appropriate solution. Low-information-cost settings promote decentralization, since workers can cheaply access problem-solving information on their own, while low-communication-cost settings encourage centralization of authority.

The theory advanced in Bloom et al. (2014) makes divergent predictions regarding the way advances in information or communications technologies affect firm structure. AI tools that reduce workers' information barriers promote decentralization of decision making (e.g., a nurse practitioner who gains access to diagnostic software). AI tools that reduce communications barriers promote centralization of decision-making (e.g., a warehouse manager who outfits her workers with sensors that continually communicate their activities). It is not hard to imagine these opposite shifts occurring simultaneously at different levels of the same firm. For instance, AI-enabled informational tools might help corporate headquarters decentralize management structures and delegate authority to plant managers even as the plant managers centralize authority within each plant thanks to advances in AI-powered direction and evaluation.

A prediction that remains relatively constant across these various theoretical viewpoints is that AI will expand managers' span of control, or number of subordinates.[33] This has implications for worker power. A

[33] This prediction is explicit in Aghion, Jones and Jones (2019), Kiron, Altman and Riedl (2023) and Ide and Talamàs (2024). Bloom et al. (2014) link reductions in information costs to greater span of control, though the relationship between communication costs and span of control is both theoretically and empirically ambiguous.

wider managerial span of control likely weakens workers' ability to exert interpersonal influence and build social capital. According to (Duggan et al., 2020, 7), this is a key function of algorithmic management: "Through its very purpose, algorithmic management eliminates the more interpersonal and empathetic aspects of people management." Although it might be tempting to write off interpersonal relationships as more of a job amenity than a component of worker power, an extensive literature in management studies connects workers' ability to build social capital and engage in "upward influence" to their compensation and career outcomes (Wayne et al., 1997; Higgins, Judge and Ferris, 2003; Ng et al., 2005).

Another AI-driven change in firm organization relevant to worker power is increasing subcontracting or fissuring of the firm (Bernhardt, Kresge and Suleiman, 2023; Kiron, Altman and Riedl, 2023). By reducing coordination costs, AI-powered systems may make subcontracting more feasible for firms that would otherwise keep services in-house. Fissured job arrangements sever key aspects of the employment relationship, chief among them the fairness and vertical equity norms that help to equalize wages within and between classes of workers (Weil, 2014). Outsourced workers also face higher barriers to exercising voice, either through formal mechanisms such as unions or informal channels within the workplace.

A final worker-power implication of changes to firm organization relates to workers' ability to bargain for a share of the surplus. When authority is delegated upstream or worker-manager contacts are mediated by algorithm, workers lose insights into their own productivity contributions to the firm. In interviews with employees in AI-mediated warehouse settings, Bell (2022) describes workers feeling a loss of a system-level understanding of their workplace. Jarrahi et al. (2021) characterize this situation as one of "technical opacity." In addition to potentially disempowering and alienating workers, technical opacity could hamper workers' bargaining abilities owing to loss of information. The situation is analogous to the that described in Section 4.2 on algorithmic wage-setting. Now, instead of firms gaining a more accurate picture of workers' outside options, AI management tools obscure workers' view of how the firm values the match (which is a function of worker productivity). This opacity shifts bargaining power to the firm.

5 Empirical Evidence

A large and growing body of empirical research examines the extent of AI in the economy and its impacts on work and workers. These studies fall into three buckets: projections of exposure to AI technologies across the labor force; measures of AI in current use; and experiments (natural or otherwise) examining the application of AI tools to job tasks or workplace settings. Across these categories, research tends to focus on the task content of jobs and labor demand impacts rather than the context of work and worker power. For that reason, this section provides only a brief and selective survey. It is intended to give a high-level overview of the research already conducted into AI in the workplace, emphasizing those contributions that have implicit or explicit work power implications.

5.1 Projections

Numerous studies attempt to quantify AI exposure across occupations, geographies, and demographic groups. These necessarily imperfect exercises have two goals: to project AI's potential future impacts across the labor force and to provide measures of past and current AI exposure to be used in empirical studies of AI labor-market impacts.

Projections of AI impacts begin with detailed occupation-level descriptions of job and worker characteristics, most often those drawn from the O*NET database. Researchers then construct occupational indices of AI exposure by assessing to what degree aspects of a job might be at risk of AI substitution. Key research decisions center on what component of the job to examine (tasks, skills, or abilities) and how to assess AI exposure within that component. Brynjolfsson, Mitchell and Rock (2018) crowdsource judgments of task-level exposure to create a measure of suitability for machine learning. Felten, Raj and Seamans (2018) focus on occupation-level abilities rather than tasks, relying on expert judgments to assess AI potential within each ability. Webb (2019) draws from the text of AI-related patents to quantify the AI exposure of O*NET job tasks. All three of these influential studies predict widespread exposure to AI, though they remain agnostic as to the balance of augmentation or displacement implied by exposure. They also find that AI exposure tends to rise with occupational income or education levels.

The late-2022 release of ChatGPT helped spur a second wave of AI projections, some of which focus exclusively on generative AI (Felten, Raj and Seamans, 2023; Eloundou et al., 2023).[34] Svanberg et al. (2024) zoom in on a specific ability, computer vision, to produce a detailed, end-to-end examination of both the ability and the cost-effectiveness of AI taking on vision tasks. They find that nearly a quarter of workers' wages in vision-related tasks may be attractive to automate. Researchers have also begun to nuance the projections by adding additional dimensions of job characteristics to the analysis. Pizzinelli et al. (2023) and Cazzaniga et al. (2024) consider not just the content but also the context of work, seeking to distinguish occupations where AI is likely to complement rather than replace human labor. These studies confirm prior findings that AI exposure rises with educational attainment while also noting that highly educated occupations generally score high on AI complementarity. Pizzinelli et al. (2023) find that, among broad occupation groups, managers exhibit the highest degree of complementary AI exposure. This is consistent with the intuition that AI tools will tilt power to managers.

Although existing AI-projection studies have implications for workers' task mix and thus their bargaining position, they say little about potential consequences for worker power. With the exception of Pizzinelli et al. (2023) and Cazzaniga et al. (2024), this literature focuses exclusively on job content rather than job context. None examines how workers are managed, measured, or monitored. Future research could use O*NET or similar databases to assess occupational exposure to AI-powered management and oversight tools in the workplace. Are certain job characteristics common across occupations noted for widespread algorithmic management practices?

5.2 Measures of AI use

Another group of studies documents the prevalence of AI tools among firms. Over the past decade government data collection agencies have introduced various questions related to AI in nationally representative surveys. Both Acemoglu et al. (2022*b*) and McElheran et al. (2023) draw from newly available modules of the U.S. Census Bureau's Annual Business Survey (ABS) to document the prevalence of AI used in the production of goods and services in U.S. businesses. In the 2018 ABS, McElheran et al. (2023) find AI use to be "low and skewed," with AI adoption rising with firm size. Overall, 5.8% of firms used AI in reference year 2017, with 18.2% of employees working at AI-using companies. The comparable figures for Acemoglu et al. (2022*b*)—who use the 2019 ABS, reference period 2016–2018—are 3.2% of

[34] A variety of nonacademic researchers have also projected AI exposure across the labor force, including Goldman Sachs (2023); Goldstein et al. (2023); Kimbrough and Carpanelli (2023); Chui et al. (2023) and Jobs for the Future (2023).

firms and 12.6% of employees. More recent evidence comes from Bonney et al. (2024), who use the Census Bureau's Business Trends and Outlook Survey (BTOS) to track business uptake of AI technologies between September 2023 and April 2024. Over that period, the share of firms using some AI technology in production rose from 3.7% to 5.4%, or 4.5% to 9% weighted by employment.[35]

Survey questions capturing AI used in personnel management and human resources, which would be particularly relevant to issues of worker power, have not yet been added to U.S. Census data products. ABS and BTOS survey questions around AI pertain only to the production of goods and services. The ABS does ask respondents to quantify changes in the ratio of supervisory to nonsupervisory workers brought about by the introduction of AI, and from these questions researchers may gain indirect insights into AI-driven shifts to management practices. Yet the data do not yet provide a strong foundation for inquiries into AI's effects on management and the organization of work. This is in contrast to AI-related surveys administered by agencies elsewhere around the world, including the European Commission and the Israel Central Bureau of Statistics, which cover a variety of contexts for AI use, including in recruitment and human resource management (Montagnier and Ek, 2021).

Researchers have also tracked the prevalence of AI among employers using job postings, patent applications, worker resumes, and other non-government data sources. Utilizing job listings data, Alekseeva et al. (2021) find that demand for AI skills spans industries, with the highest demand in information technology, architecture and engineering, science, and management occupations. They estimate that AI skills premiums are largest in management occupations, suggesting that "AI's value may rest in its combination with a broader organizational change." Goldfarb, Taska and Teodoridis (2023) use job postings to argue that the cluster of technologies around machine learning is widely used enough to qualify as a general purpose technology (Trajtenberg, 2019). Bessen et al. (2018) surveyed AI-related startups on the nature and expected labor-market effects of their products. In the aggregate, respondents saw their technologies spurring net job growth for higher-paying occupations (professionals, managers, sales and marketing) and net job loss among lower-paying occupations (front-line service, clerical, and manual workers).

5.3 AI in job tasks

Extensive empirical research examines the impact of AI adoption on workers, firms, and labor markets. The largest and fastest-growing branch of this literature studies the productivity consequences for individual workers of AI tools or copilots that automate or assist with work tasks. Researchers have explored AI assistance in a wide range of tasks and work settings: coding (Campero et al., 2022; Kreitmeir and Raschky, 2023; Peng et al., 2023; Vaithilingam, Zhang and Glassman, 2022), office and knowledge work (Cambon et al., 2023; Dell'Acqua et al., 2023; Noy and Zhang, 2023), legal work (Choi and Schwarcz, 2023), medical diagnostics (McDuff et al., 2023), recruiting (Dell'Acqua, 2023), creative work (Doshi and Hauser, 2023; Jia et al., 2023), forecasting (Schoenegger et al., 2024), and taxi driving (Kanazawa et al., 2022). Studies in this area tend to share two broad findings: increased productivity for the average worker and reduced skill gaps between the workers varying in skill and experience.[36] Yet researchers have also

[35] Differences in AI usage statistics between Acemoglu et al. (2022*b*) and McElheran et al. (2023) likely come from changes in the AI-related questions between the survey years. AI usage appears to be far lower in Bonney et al. (2024) since they diverge from the other two papers by including responses of "Do not know" in the denominators.

[36] A related though distinct body of literature tests the performance of AI outputs against human outputs in tasks such as medical diagnostics or business idea generation. These benchmarking exercises differ from those studies, summarized above, that compare human-AI collaboration to human-only performance.

found exceptions to these broad conclusions, documenting AI applications that diminish productivity for certain workers or exacerbate skill gaps. (Dell'Acqua, 2023; Dell'Acqua et al., 2023; Otis et al., 2023; Haslberger, Gingrich and Bhatia, 2023; Yu et al., 2024).

One of the very few studies to document individual-level AI tool impacts in real-world business settings—beyond the confines of artificial experiments and laboratory tests—is Brynjolfsson, Li and Raymond (2023). The authors examine the staggered rollout of an AI assistance tool in a customer service department of a Fortune 500 software firm, finding that agents equipped with AI assistance resolved more customer issues in less time, with the largest productivity gains for the least-experienced workers. AI adoption reduced attrition, especially at lower tenures, and the text-based sentiment of customer interactions improved, suggesting a better work experience. The AI system succeeded due to its "ability to embody the best practices of high-skill workers in our firm and make it accessible to other workers" Brynjolfsson, Li and Raymond (2023, 27). Cui et al. (2024) study the rollout of the GitHub Copilot coding assistant at Microsoft and Accenture, finding positive though imprecisely estimated productivity effects.

Another set of papers examines the firm-level effects of AI uptake. Acemoglu et al. (2022*a*) use job posting data to examine the impact of AI advances on hiring in AI and non-AI roles. Using the occupational AI exposure index of Felten, Raj and Seamans (2018), they document a negative relationship between AI adoption and firm-level non-AI job postings, as well as an increase in new skills in firm job postings. Most but not all of these results are robust to using alternate indices of AI exposure (Brynjolfsson, Mitchell and Rock, 2018; Webb, 2019). Babina et al. (2023) use matched data from resumes, job postings and financial filings to explore firm investments in AI skills. They document that AI workforce investment is associated with a flatter organizational structure with less emphasis on middle and senior management. Several studies have found a positive relationship between firm-level AI adoption and either productivity or revenue growth (Damioli, Van Roy and Vertesy, 2021; Bäck et al., 2022; Czarnitzki, Fernández and Rammer, 2023; Lee et al., 2022).

Finally, a number of studies have estimated AI impacts at the level of labor markets. In addition to assessing firm-level impacts, Acemoglu et al. (2022*a*) examine the effects of AI adoption (proxied using the indices listed above) on industry- and occupation-level wages and employment. The results are null across specifications. Bonfiglioli et al. (2024), using a novel O*NET-based AI measure, find a robustly negative association of AI adoption and aggregate employment at the commuting-zone level, with the largest effects for lower-skilled and production workers. These results echo those of Cornelli, Frost and Mishra (2023), who document a positive association between AI adoption and income inequality in a panel of country-level data. In contrast to these studies, Albanesi et al. (2023) show a positive relationship between occupational AI exposure and employment in a pooled sample of European countries, with effects limited to high-skilled occupations.

6 Conclusion

This paper laid out a research agenda around the potential impacts of AI technologies on worker power. By distinguishing between the labor demand impacts and worker power impacts of AI, I provided a broad framework to help conceptualize AI's varied implications for workers and labor markets. The analysis identified and formalized several key mechanisms through which AI may reshape power dynamics in the workplace, including enhanced monitoring, predictive analytics in wage bargaining, advanced employee screening, and algorithmic management systems. The unifying theme across these mechanisms is that, by enhancing prediction and information gathering in managerial functions, AI may shift information rents from workers to firms. These insights seek to extend the economic literature on AI and labor by offering a novel perspective on AI's potential to influence income distribution and the nature of work.

Future empirical research is needed to validate the theoretical predictions made in this paper. The first and most pressing priority is to collect high-quality data capable of capturing the extent of AI technologies used in management, HR, and related business functions. As noted in Section 5, current U.S. Census surveys on business AI adoption focus almost exclusively on AI used in the production of goods and services. These surveys overlook the many applications of AI to recruiting, hiring, wage-setting, scheduling, monitoring, task allocation, and performance evaluation—applications with significant worker-power implications. Attention to this kind of AI tool would also be valuable in the context of projections of AI exposure and measures of AI adoption based on alternative data types, such as worker resumes and job postings. Further theoretical contributions regarding technology and worker power are also needed.

This paper left unaddressed some pressing questions related to AI and political economy. Namely, how might *existing* structures of worker power influence the development and deployment of automating and job-altering AI technologies? One set of answers points to directed technological change and the role of labor-centered institutions in shaping the vision of the future (Noble, 2011; Acemoglu and Johnson, 2023). A much older literature, dating at least to the industrial revolution (Babbage, 1832), highlights the use of automation to purposefully dislodge nodes of worker solidarity. Thus Marx saw labor-saving machinery as "the most powerful weapon for suppressing strikes" (Marx, 1992, 562).[37] Whether due to worker influence over technological development or targeted deskilling, existing power relations and institutional structures are likely to influence the course of AI adoption in ways that exceed the scope of this study.

[37] A particularly notable example can be found in late-19th-century capitalist Cyrus McCormick, Jr., who introduced new machinery to his mechanical reaper plant to "weed out the bad element among the men," i.e., skilled workers who had organized a union (quoted in Winner, 1980, 124–125).

References

Abrardi, Laura, Carlo Cambini, and Laura Rondi. 2022. "Artificial Intelligence, Firms and Consumer Behavior: A Survey." *Journal of Economic Surveys*, 36(4): 969–991.

Acemoglu, Daron. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *The Quarterly Journal of Economics*, 113(4): 1055–1089.

Acemoglu, Daron. 2021. "Harms of AI." *National Bureau of Economic Research Working Paper Series*, 29247.

Acemoglu, Daron. 2024. "The Simple Macroeconomics of AI." National Bureau of Economic Research. Working Paper 32487.

Acemoglu, Daron, and Andrew F. Newman. 2002. "The labor market and corporate structure." *European Economic Review*, 46(10): 1733–1756.

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." *Handbook of Labor Economics*, 4: 1043–1171.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review*, 108(6): 1488–1542.

Acemoglu, Daron, and Pascual Restrepo. 2019*a*. "Artificial Intelligence, Automation, and Work." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay K. Agrawal, Joshua Gans and Avi Goldfarb. Chicago: University of Chicago Press.

Acemoglu, Daron, and Pascual Restrepo. 2019*b*. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives*, 33(2): 3–30.

Acemoglu, Daron, and Simon Johnson. 2023. *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity.* Cambridge, MA: PublicAffairs.

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo. 2022*a*. "Artificial Intelligence and Jobs: Evidence from Online Vacancies." *Journal of Labor Economics*, 40(S1): S293–S340.

Acemoglu, Daron, Gary W. Anderson, David N. Beede, Catherine Buffington, Eric E. Childress, Emin Dinlersoz, Lucia S. Foster, Nathan Goldschlag, John C. Haltiwanger, Zachary Kroff, Pascual Restrepo, and Nikolas Zolas. 2022*b*. "Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey." *Technology, Productivity, and Economic Growth*, , ed. Susanto Basu, Lucy Eldridge, John Haltiwanger and Erich Strassner, Chapter 7. University of Chicago Press.

Aghion, Philippe, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy*, 105(1): 1–29.

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones. 2019. "Artificial Intelligence and Economic Growth." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb. Chicago: University of Chicago Press.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence.* Boston, MA: Harvard Business Review Press.

Agrawal, Ajay K., Joshua S. Gans, and Avi Goldfarb. 2019*a*. "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction." *Journal of Economic Perspectives*, 33(2): 31–50.

Agrawal, Ajay K., Joshua S Gans, and Avi Goldfarb. 2019*b*. "Exploring the Impact of Artificial Intelligence: Prediction Versus Judgment." *Information Economics and Policy*, 47: 1–6.

Agrawal, Ajay K., Joshua S. Gans, and Avi Goldfarb. 2019*c*. "Introduction to "The Economics of Artificial Intelligence: An Agenda"." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb. Chicago: University of Chicago Press.

Agrawal, Ajay K., Joshua S. Gans, and Avi Goldfarb. 2023. "The Turing Transformation: Artificial Intelligence, Intelligence Augmentation, and Skill Premiums." Brookings Institution Policy Brief.

Ajunwa, Ifeoma. 2023. *The Quantified Worker: Law and Technology in the Modern Workplace.* Cambridge: Cambridge University Press.

Ajunwa, Ifeoma, Kate Crawford, and Jason Schultz. 2017. "Limitless Worker Surveillance." *California Law Review*, 105(3): 735–776.

Albanesi, Stefania, António Dias da Silva, Juan F. Jimeno, Ana Lamo, and Alena Wabitsch. 2023. "New Technologies and Jobs in Europe." National Bureau of Economic Research Working Paper 31357.

Alderotti, Giammarco, Chiara Rapallini, and Silvio Traverso. 2023. "The Big Five Personality Traits and Earnings: A Meta-Analysis." *Journal of Economic Psychology*, 94: 102570.

Alekseeva, Liudmila, José Azar, Mireia Giné, Sampsa Samila, and Bledi Taska. 2021. "The Demand for AI Skills in the Labor Market." *Labour Economics*, 71.

Allgulin, Magnus, and Tore Ellingsen. 2002. "Monitoring and Pay." *Journal of Labor Economics*, 20(2): 201–216.

Alvarez-Valle, Javier, and Matthew Lungren. 2023. "GPT-4's Potential in Shaping the Future of Radiology." Microsoft. Project MAIRA working paper.

American Psychological Association. 2023. "2023 Work in America Survey: Artificial intelligence, monitoring technology, and psychological well-being." American Psychological Association.

Andalibi, Nazanin. 2024. "Emotion-Tracking AI On The Job: Workers Fear Being Watched – And Misunderstood." *The Conversation.* https://theconversation.com/emotion-tracking-ai-on-the-job-workers-fear-being-watched-and-misunderstood-222592.

Armstrong, Ben, Adam Landman, Sal Companieh, and Jonathan DeBusk. 2024. "Generative AI and Work of the Future Industry Panel." 2024 MIT AI Conference: Tech, Business, and Ethics (panel conversation). 29 February, 2024.

Athey, Susan C., Kevin A. Bryan, and Joshua S. Gans. 2020. "The Allocation of Decision Authority to Human and Artificial Intelligence." *AEA Papers and Proceedings*, 110: 80–84.

Autor, David. 2014. "Polanyi's Paradox and the Shape of Employment Growth." National Bureau of Economic Research Working Paper 20485.

Autor, David. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *The Journal of Economic Perspectives*, 29(3): 3–30.

Autor, David. 2022. "The Labor Market Impacts of Technological Change: From Unbridled Enthusiasm to Qualified Optimism to Vast Uncertainty." National Bureau of Economic Research Working Paper 30074.

Autor, David. 2024. "Applying AI to Rebuild Middle Class Jobs." National Bureau of Economic Research Working Paper 32140.

Autor, David, Caroline Chin, Anna Salomons, and Bryan Seegmiller. 2024. "New Frontiers: The Origins and Content of New Work, 1940–2018." *The Quarterly Journal of Economics*.

Autor, David, Frank Levy, and Richard J Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics*, 118(4): 1279–1333.

Autor, David, Lawrence F Katz, and Alan B Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, 113(4): 1169–1213.

Autor, David, Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90(2): 300–323.

Azar, José, Marina Chugunova, Klaus Keller, and Sampsa Samila. 2023. "Monopsony and Automation." Max Planck Institute for Innovation & Competition Research Paper 23-21.

Babbage, Charles. 1832. *On the Economy of Machinery and Manufactures.* London: C. Knight.

Babina, Tania, Anastassia Fedyk, Alex X. He, and James Hodson. 2023. "Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition." National Bureau of Economic Research 31325.

Bäck, Asta, Arash Hajikhani, Angela Jäger, Torben Schubert, and Arho Suominen. 2022. "Return of the Solow-paradox in AI?: AI-adoption and Firm Productivity." Centre for Innovation Research, Lund University. Papers in Innovation Studies no. 2022/01.

Bales, Richard A, and Katherine VW Stone. 2020. "The Invisible Web at Work: Artificial Intelligence and Electronic Surveillance in the Workplace." *Berkeley Journal of Employment and Labor Law*, 41: 1.

Ball, Kirstie. 2021. "Electronic Monitoring and Surveillance in the Workplace: Literature Review and Policy Recommendations." Joint Research Centre, European Commission, Luxembourg.

Becker, Gary S, and Kevin M Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge." *The Quarterly Journal of Economics*, 107(4): 1137–1160.

Bell, Stephanie A. 2022. "AI and Job Quality: Insights from Frontline Workers." Partnership on AI, San Francisco, CA.

Benkler, Yochai. 2022. "Power and Productivity: Institutions, Ideology, and Technology in Political Economy." In *A Political Economy of Justice.* , ed. Danielle Allen, Yochai Benkler, Leah Downey, Rebecca Henderson and Josh Simons. Chicago: University of Chicago Press.

Bernhardt, Annette, Lisa Kresge, and Reem Suleiman. 2023. "The Data-Driven Workplace and the Case for Worker Technology Rights." *ILR Review*, 76(1): 3–29.

Bessen, James. 2015. "Toil and Technology: Innovative Technology Is Displacing Workers to New Jobs Rather Than Replacing Them Entirely." *Finance & Development*, 52(001): 16.

Bessen, James E., Stephen Michael Impink, Lydia Reichensperger, and Robert Seamans. 2018. "The Business of AI Startups." Boston University School of Law. Law and Economics Research Paper 11-2018.

Bhuller, Manudeep, Domenico Ferraro, Andreas R Kostøl, and Trond C Vigtel. 2023. "The Internet, Search Frictions and Aggregate Unemployment." National Bureau of Economic Research. Working Paper 30911.

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky. 1986. "The Nash Bargaining Solution in Economic Modelling." *The RAND Journal of Economics*, 17(2): 176–188.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. 2014. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization." *Management Science*, 60.

Bogen, Miranda, and Aaron Rieke. 2018. "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias." Upturn, Washington, DC.

Bonfiglioli, Alessandra, Rosario Crinò, Gino Gancia, and Ioannis Papadakis. 2024. "Artificial Intelligence and Jobs: Evidence from US Commuting Zones." Centre for Economic Policy Research DP 18495.

Bonney, Kathryn, Cory Breaux, Cathy Buffington, Emin Dinlersoz, Lucia S Foster, Nathan Goldschlag, John C Haltiwanger, Zachary Kroff, and Keith Savage. 2024. "Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey." National Bureau of Economic Research Working Paper 32319.

Bowles, Samuel. 1985. "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models." *American Economic Review*, 75(1): 16–36.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*, 39(4): 1137–1176.

Braverman, Harry. 1998. *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century.* New York: Monthly Review Press.

Brynjolfsson, Erik. 2022. "The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence." *Daedalus*, 151(2): 272–287.

Brynjolfsson, Erik, and Andrew Ng. 2023. "Big AI Can Centralize Decision-Making and Power, and That's a Problem." In *Missing Links in AI Governance.* , ed. Benjamin Prud'homme, Catherine Régis, Golnoosh Farnadi, Vanessa Dreier, Sasha Rubel and Charline d'Oultremont, 65–87. UNESCO.

Brynjolfsson, Erik, Danielle Li, and Lindsey R Raymond. 2023. "Generative AI at Work." National Bureau of Economic Research. Working Paper 31161.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock. 2018. "What Can Machines Learn and What Does It Mean for Occupations and the Economy?" *AEA Papers and Proceedings*, 108: 43–47.

Bulow, Jeremy I, and Lawrence Summers. 1986. "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment." *Journal of Labor Economics*, 4(3): 376–414.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica*, 74(2): 323–364.

Cambon, Alexia, Brent Hecht, Benjamin Edelman, Donald Ngwe, Sonia Jaffe, Amy Heger, Mihaela Vorvoreanu, Sida Peng, Jake Hofman, Alex Farach, Margarita Bermejo-Cano, Eric Knudsen, James Bono, Hardik Sanghavi, Sofia Spatharioti, David Rothschild, Daniel G. Goldstein, Eirini Kalliamvakou, Peter Cihon, Mert Demirer, Michael Schwarz, and Jaime Teevan. 2023. "Early LLM-based Tools for Enterprise Information Workers Likely Provide Meaningful Boosts to Productivity." Microsoft Research. MSR-TR-2023-43.

Campero, Andres, Michelle Vaccaro, Jaeyoon Song, Haoran Wen, Abdullah Almaatouq, and Thomas Malone. 2022. "A Test for Evaluating Performance in Human-Computer Systems." arXiv preprint. arXiv:2206.12390.

Card, David, and John E DiNardo. 2002. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*, 20(4): 733–783.

Cazzaniga, Mauro, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus J Panton, Carlo Pizzinelli, Emma J Rockall, and Marina Mendes Tavares. 2024. "Gen-AI: Artificial Intelligence and the Future of Work." International Monetary Fund. Staff Discussion Notes No. 2024/001.

Chatterjee, Kalyan, and William Samuelson. 1983. "Bargaining under Incomplete Information." *Operations Research*, 31(5): 835–851.

Chen, Te-Ping. 2024). "Your Résumé Might Be Getting Tossed by AI. How to Push Back." *Wall Street Journal*. https://www.wsj.com/lifestyle/careers/if-humans-wont-read-your-resume-should-you-let-the-robot-72bb641c.

Choi, Jonathan H., and Daniel Schwarcz. 2023. "AI Assistance in Legal Analysis: An Empirical Study." University of Minnesota Law School. Minnesota Legal Studies Research Paper No. 23-22.

Chui, Michael, Roger Roberts, Lareina Yee, Eric Hazan, Alex Singla, Kate Smaje, Alex Sukharevsky, and Rodney Zemmel. 2023. "The Economic Potential of Generative AI: The Next Productivity Frontier." McKinsey & Company.

Conference Board. 2023. "Majority of US Workers Are Already Using Generative AI Tools—But Company Policies Trail Behind." https://www.conference-board.org/press/us-workers-and-generative-ai.

Cornelli, Giulio, Jon Frost, and Saurabh Mishra. 2023. "Artificial Intelligence, Services Globalisation, and Income Inequality." Bank for International Settlements Working Paper 1135.

Cross, John G. 1965. "A Theory of the Bargaining Process." *The American Economic Review*, 55(1/2): 67–94.

Cui, Kevin Zheyuan, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. 2024. "The Productivity Effects of Generative AI: Evidence from a Field Experiment with GitHub Copilot." MIT. In "An MIT Exploration of Generative AI,".

Cullen, Zoe B, Shengwu Li, and Ricardo Perez-Truglia. 2022. "What's My Employee Worth? The Effects of Salary Benchmarking." National Bureau of Economic Research. Working Paper 30570.

Czarnitzki, Dirk, Gastón P Fernández, and Christian Rammer. 2023. "Artificial Intelligence and Firm-Level Productivity." *Journal of Economic Behavior & Organization*, 211: 188–205.

Damioli, Giacomo, Vincent Van Roy, and Daniel Vertesy. 2021. "The Impact of Artificial Intelligence on Labor Productivity." *Eurasian Business Review*, 11: 1–25.

Dastin, Jeffrey. 2018. "Insight—Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women." *Reuters*. https://www.reuters.com/article/idUSKCN1MK0AG/.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine C. Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani. 2023. "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality." Harvard Business School. Technology & Operations Mgt. Unit Working Paper No. 24-013.

Dell, Kristin, Nicole Nestoriak, and Jenny Marlar. 2020. "Assessing the Impact of New Technologies on the Labor Market: Key Constructs, Gaps, and Data Collection Strategies for the Bureau of Labor Statistics." Bureau of Labor Statistics. Submitted by Gallup.

Dell'Acqua, Fabrizio. 2023. "Falling Asleep at the Wheel: Human/AI Collaboration in a Field Experiment on HR Recruiters." Laboratory for Innovation Science, Harvard Business School. working paper.

Dessein, Wouter. 2002. "Authority and Communication in Organizations." *The Review of Economic Studies*, 69(4): 811–838.

Doshi, Anil R., and Oliver Hauser. 2023. "Generative Artificial Intelligence Enhances Creativity but Reduces the Diversity of Novel Content." arXiv preprint. arXiv:2312.00506.

Dubal, Veena. 2023. "On Algorithmic Wage Discrimination." *Columbia Law Review*, 123(7): 1929–1992.

Dube, Arindrajit, Suresh Naidu, and Adam D Reich. 2022. "Power and Dignity in the Low-Wage Labor Market: Theory and Evidence from Walmart Workers." National Bureau of Economic Research. Working Paper 30441.

Duggan, James, Ultan Sherman, Ronan Carbery, and Anthony McDonnell. 2020. "Algorithmic Management and App-Work in the Gig Economy: A Research Agenda for Employment Relations and HRM." *Human Resource Management Journal*, 30(1): 114–132.

Edwards, Richard. 1979. *Contested Terrain: The Transformation of the Workplace in the Twentieth Century.* New York: Basic Books.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. "GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models." arXiv preprint. arXiv:2303.10130.

Faveiro, Michelle, and Alec Tyson. 2023. "What the Data Says About Americans' Views of Artificial Intelligence." Pew Research Center, Washington, DC. https://www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence/.

Felps, Will, Terence R Mitchell, David R Hekman, Thomas W Lee, Brooks C Holtom, and Wendy S Harman. 2009. "Turnover Contagion: How Coworkers' Job Embeddedness and Job Search Behaviors Influence Quitting." *Academy of Management Journal*, 52(3): 545–561.

Felten, Edward W., Manav Raj, and Robert Seamans. 2018. "A Method to Link Advances in Artificial Intelligence to Occupational Abilities." *AEA Papers and Proceedings*, 108: 54–57.

Felten, Edward W., Manav Raj, and Robert Seamans. 2023. "Occupational Heterogeneity in Exposure to Generative AI." Available at SSRN 4414065.

Flinn, Christopher, Petra Todd, and Weilong Zhang. 2020. "Personality Traits, Job Search and the Gender Wage Gap." University of Cambridge. Cambridge Working Papers in Economics: 2053.

Garden, Charlotte. 2018. "Labor Organizing in the Age of Surveillance." *Saint Louis University Law Journal*, 63: 55.

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy*, 108(5): 874–904.

Girotra, Karan, Lennart Meincke, Christian Terwiesch, and Karl T. Ulrich. 2023. "Ideas are Dimes a Dozen: Large Language Models for Idea Generation in Innovation." SSRN working paper.

GitHub. 2023. "Survey Reveals AI's Impact on the Developer Experience." https://github.blog/2023-06-13-survey-reveals-ais-impact-on-the-developer-experience/.

Gmyrek, Pawel, Janine Berg, and David Bescond. 2023. "Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality." International Labour Organization. ILO Working Paper 96.

Goldfarb, Avi, Bledi Taska, and Florenta Teodoridis. 2023. "Could Machine Learning Be a General Purpose Technology? A Comparison of Emerging Technologies Using Data from Online Job Postings." *Research Policy*, 52(1): 104653.

Goldman Sachs. 2023. "Generative AI: Hype, or Truly Transformative?" Goldman Sachs, New York, NY.

Goldstein, Jill, Bill Lobig, Cathy Fillare, and Christopher Nowak. 2023. "Augmented Work for an Automated, AI-Driven World." IBM Institute for Business Value, Armonk, NY.

Goos, Maarten, Alan Manning, and Anna Salomons. 2009. "Job Polarization in Europe." *American Economic Review*, 99(2): 58–63.

Griesbach, Kathleen, Adam Reich, Luke Elliott-Negri, and Ruth Milkman. 2019. "Algorithmic Control in Platform Food Delivery Work." *Socius*, 5: 1–15.

Guenole, Nigel, and Sheri Feinzig. 2022. "The Business Case for AI in HR." IBM. https://www.ibm.com/downloads/cas/A5YLEPBR.

Guo, Qi. 2019. "The AI Behind LinkedIn Recruiter Search and Recommendation Systems." LinkedIn Engineering Blog. https://www.linkedin.com/blog/engineering/recommendations/ai-behind-linkedin-recruiter-search-and-recommendation-systems.

Gutelius, Beth, and Nik Theodore. 2019. "The Future of Warehouse Work: Technological Change in the U.S. Logistics Industry." UC Berkeley Labor Center, Berkeley, CA.

Hall, Robert E., and Alan B. Krueger. 2012. "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search." *American Economic Journal: Macroeconomics*, 4(4): 56–67.

Haslberger, Matthias, Jane Gingrich, and Jasmine Bhatia. 2023. "No Great Equalizer: Experimental Evidence on AI in the UK Labor Market." SSRN working paper.

He, Amy. 2024. "State of Workers 2024 Report." Morning Consult. pro.morningconsult.com/analyst-reports/state-of-workers-2024.

Higgins, Chad A., Timothy A. Judge, and Gerald R. Ferris. 2003. "Influence Tactics and Work Outcomes: A Meta-Analysis." *Journal of Organizational Behavior*, 24(1): 89–106.

Ide, Enrique, and Eduard Talamàs. 2024. "Artificial Intelligence in the Knowledge Economy." arXiv preprint. arXiv:2312.05481.

Imana, Basileal, Aleksandra Korolova, and John Heidemann. 2021. "Auditing for Discrimination in Algorithms Delivering Job Ads." In *WWW '21: Proceedings of the Web Conference 2021*. 3767–3778. Association for Computing Machinery.

Isztin, Peter. 2023. "A Note on Human Capital, the Division of Labor, and Artificial Intelligence." SSRN Working Paper 4581225.

Jacobo, Juan. 2023. "There is Power in General Equilibrium." arXiv preprint. arXiv:2309.00909.

Jarrahi, Mohammad Hossein, Gemma Newlands, Min Kyung Lee, Christine T Wolf, Eliscia Kinder, and Will Sutherland. 2021. "Algorithmic Management in a Work Context." *Big Data & Society*, 8(2).

Jia, Nan, Xueming Luo, Zheng Fang, and Chengcheng Liao. 2023. "When and How Artificial Intelligence Augments Employee Creativity." *Academy of Management Journal*, 67(1).

Jobs for the Future. 2023. "The AI-Ready Workforce: How Leaders and Workers Can Prepare for a Reshaped Future of Work." The Center for Artificial Intelligence and the Future of Work, Jobs for the Future.

Judge, Timothy A, Beth A Livingston, and Charlice Hurst. 2012. "Do Nice Guys—and Gals—Really Finish Last? The Joint Effects of Sex and Agreeableness on Income." *Journal of Personality and Social Psychology*, 102(2): 390–407.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer. 2024. "Worker Beliefs About Outside Options." *The Quarterly Journal of Economics*.

Kanazawa, Kyogo, Daiji Kawaguchi, Hitoshi Shigeoka, and Yasutora Watanabe. 2022. "AI, Skill, and Productivity: The Case of Taxi Drivers." National Bureau of Economic Research. Working Paper 30612.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, 107(1): 35–78.

Kellogg, Katherine C, Melissa A Valentine, and Angele Christin. 2020. "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals*, 14(1): 366–410.

Kiela, Douwe, Tristan Thrush, Kawin Ethayarajh, and Amanpreet Singh. 2023. "Plotting Progress in AI." Contextual AI Blog. https://contextual.ai/blog/plotting-progress.

Kimbrough, Karin, and Mar Carpanelli. 2023. "Preparing the Workforce for Generative AI: Insights and Implications." LinkedIn.

Kim, Pauline. 2022. "AI and Inequality." In *The Cambridge Handbook on Artificial Intelligence & the Law.* , ed. Kristin Johnson and Carla Reyes. Cambridge: Cambridge University Press.

Kiron, David, Elizabeth J. Altman, and Christoph Riedl. 2023. "Workforce Ecosystems and AI." Brookings Institution, Washington, DC.

Klinova, Katya, and Anton Korinek. 2021. "AI and Shared Prosperity." *AIES '21*, 645–651. Association for Computing Machinery.

Korinek, Anton, and Joseph E. Stiglitz. 2019. "Artificial Intelligence and Its Implications for Income Distribution and Unemployment." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb. Chicago: University of Chicago Press.

Kreacic, Ana, Amy Lasater-Wille, Lucia Uribe, Ravin Jesuthasan, John Romeo, and Simon Luong. 2024. "How Generative AI is Transforming Business and Society." Oliver Wyman.

Kreitmeir, David, and Paul A. Raschky. 2023. "The Unintended Consequences of Censoring Digital Technology—Evidence from Italy's ChatGPT Ban." arXiv preprint. arXiv:2304.09339.

Kresge, Lisa. 2020. "Data and Algorithms in the Workplace: A Primer on New Technologies." UC Berkeley Labor Center, Berkeley, CA.

Krueger, Alan B. 2018. "Luncheon Address: Reflections on Dwindling Workers Bargaining Power and Monetary Policy." Proceedings of the Jackson Hole Economic Policy Symposium, "Changing Market Structures and Implications for Monetary Policy".

Landes, David S. 1986. "What Do Bosses Really Do?" *The Journal of Economic History*, 46(3): 585–623.

Lane, Marguerita, and Anne Saint-Martin. 2021. "The Impact of Artificial Intelligence on the Labour Market: What Do We Know So Far?" OECD. Social, Employment and Migration Working Papers No. 256.

Lee, Min Kyung, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. "Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers." *CHI '15*, 1603–1612. New York: Association for Computing Machinery.

Lee, Yong Suk, Taekyun Kim, Sukwoong Choi, and Wonjoon Kim. 2022. "When Does AI Pay Off? AI-Adoption Intensity, Complementary Investments, and R&D Strategy." *Technovation*, 118: 102590.

Levy, Frank, and Peter Temin. 2007. "Inequality and Institutions in 20th Century America." National Bureau of Economic Research. Working Paper 13106.

Lombardi, Marco J, Marianna Riggi, and Eliana Viviano. 2020. "Bargaining Power and the Phillips Curve: A Micro-Macro Analysis." Bank of Italy. Temi di Discussione (Working Paper) No. 1302.

Maiberg, Emanuel. 2024. "Service Jobs Now Require Bizarre Personality Test From AI Company." *404 Media*. https://www.404media.co/low-paying-jobs-require-bizarre-personality-evaluation-from-ai-company.

Manning, Alan. 2013. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

Marglin, Stephen A. 1974. "What Do Bosses Do?: The Origins and Functions of Hierarchy in Capitalist Production." *Review of Radical Political Economics*, 6(2): 60–112.

Marx, Karl. 1992. *Capital, Volume 1: A Critique of Political Economy.* London: Penguin Classics.

Mateescu, Alexandra, and Aiha Nguyen. 2019. "Explainer: Algorithmic Management in the Workplace." Data and Society.

McClain, Colleen. 2024. "Americans' Use of Chatgpt Is Ticking Up, But Few Trust its Election Information." Pew Research Center. https://www.pewresearch.org/short-reads/2024/03/26/americans-use-of-chatgpt-is-ticking-up-but-few-trust-its-election-information/.

McDuff, Daniel, Mike Schaekermann, Tao Tu, Anil Palepu, Amy Wang, Jake Garrison, Karan Singhal, Yash Sharma, Shekoofeh Azizi, Kavita Kulkarni, Le Hou, Yong Cheng, Yun Liu, S Sara Mahdavi, Sushant Prakash, Anupam Pathak, Christopher Semturs, Shwetak Patel, Dale R Webster, Ewa Dominowska, Juraj Gottweis, Joelle Barral, Katherine Chou, Greg S Corrado, Yossi Matias, Jake Sunshine, Alan Karthikesalingam, and Vivek Natarajan. 2023. "Towards Accurate Differential Diagnosis with Large Language Models." arXiv preprint. arXiv:2312.00164v1.

McElheran, Kristina, J Frank Li, Erik Brynjolfsson, Zachary Kroff, Emin Dinlersoz, Lucia S Foster, and Nikolas Zolas. 2023. "AI Adoption in America: Who, What, and Where." National Bureau of Economic Research. Working Paper 31788.

Mehta, Arshia, and Frank Levy. 2020. "Warehousing, Trucking, and Technology: The Future of Work in Logistics." MIT Work of the Future Research Brief RB04-2020.

Microsoft, and LinkedIn. 2024. "AI at Work Is Here. Now Comes the Hard Part." Microsoft. https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part.

Mishel, Lawrence, John Schmitt, and Heidi Shierholz. 2013. "Assessing the Job Polarization Explanation of Growing Wage Inequality." Economic Policy Institute. Working Paper 295.

Montagnier, Pierre, and Irene Ek. 2021. "AI Measurement in ICT Usage Surveys." OECD Digital Economy Papers 308.

Montgomery, David. 1980. *Workers' Control in America: Studies in the History of Work, Technology, and Labor Struggles.* New York and London: Cambridge University Press.

Moore, Phoebe V. 2020. "OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces." European Agency for Safety and Health at Work discussion paper.

Mullainathan, Sendhil. 2019. "Biased Algorithms Are Easier to Fix Than Biased People." *New York Times*. https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html. Published 6 December, 2019.

Muthoo, Abhinay. 1999. *Bargaining Theory with Applications.* Cambridge University Press.

National Center for O*NET Development. 2024. "O*NET OnLine." www.onetonline.org.

National Labor Relations Board. 2022. "NLRB General Counsel Issues Memo on Unlawful Electronic Surveillance and Automated Management Practices." Press release.

Negrón, Wilneida. 2021. "Little Tech is Coming for Workers: A Framework for Reclaiming and Building Worker Power." Coworker.org.

Ng, Thomas W. H., Lillian T. Eby, Kelly L. Sorensen, and Daniel C. Feldman. 2005. "Predictors of Objective and Subjective Career Success: A Meta-Analysis." *Personnel Psychology*, 58(2): 367–408.

Nguyen, Aiha. 2021. "The Constant Boss: Labor Under Digital Surveillance." Data and Society.

Noble, David F. 2011. *Forces of Production: A Social History of Industrial Automation.* New Brunswick, NJ: Transaction Publishers.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence." *Science*, 381(6654): 187–192.

Nurski, Laura, and Mia Hoffman. 2022. "The Impact of Artificial Intelligence on the Nature and Quality of Jobs." Bruegel. Working Paper 14/2022.

OECD. 2024. "Explanatory Memorandum on the Updated OECD Definition of an AI System." OECD Artificial Intelligence Papers 8.

Otis, Nicholas G., Rowan Clarke, Solène Delecourt, David Holtz, and Rembrand Koning. 2023. "The Uneven Impact of Generative AI on Entrepreneurial Performance." Harvard Business School 24-042.

Parent-Rocheleau, Xavier, and Sharon K. Parker. 2021. "Algorithms as Work Designers: How Algorithmic Management Influences the Design of Jobs." *Human Resource Management Review*, 32(3): 1–17.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot." arXiv preprint. arXiv:2302.06590.

Pizzinelli, Carlo, Augustus J. Panton, Marina Mendes Tavares, Mauro Cazzaniga, and Longji Li. 2023. "Labor Market Exposure to AI: Cross-country Differences and Distributional Implications." International Monetary Fund. Working Paper No. 2023/216.

Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices." *FAT* '20*, 469–481. New York, NY, USA: Association for Computing Machinery.

Rainie, Lee, Monica Anderson, Colleen Mcclain, and Emily A. Vogels nd Risa Gelles-Watnick. 2023. "AI in Hiring and Evaluating Workers: What Americans Think." Pew Research Center.

Ravid, Daniel M., Jerod C. White, David L. Tomczak, Ahleah F. Miles, and Tara S. Behrend. 2023. "A Meta-Analysis of the Effects of Electronic Performance Monitoring on Work Outcomes." *Personnel Psychology*, 76(1): 5–40.

Restrepo, Pascual. 2023. "Automation: Theory, Evidence, and Outlook." National Bureau of Economic Research. Working Paper 31910.

Rey, Jason Del, and Shirin Ghaffary. 2020. "Leaked: Confidential Amazon Memo Reveals New Software to Track Unions." *Vox*. https://www.vox.com/recode/2020/10/6/21502639/amazon-union-busting-tracking-memo-spoc.

Ritter, Joseph A., and Lowell J. Taylor. 2011. "Racial Disparity in Unemployment." *The Review of Economics and Statistics*, 93(1): 30–42.

Rogers, Brishen. 2023. *Data and Democracy at Work: Advanced Information Technologies, Labor Law, and the New Working Class.* Cambridge, MA: MIT Press.

Rosenblat, Alex, and Luke Stark. 2016. "Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers." *International Journal of Communication*, 10: 27.

Royle, Orianna Rose. 2023. "Amazon Driver Breaks Down The A.I. System Watching Workers For Safety Violations Like Drinking Coffee While Driving And Counting The Times They Buckle Their Seatbelt." *Fortune*. https://fortune.com/2023/02/27/amazon-delivery-driver-breaks-down-ai-surveillance-in-vans/.

Schaupp, Simon. 2022. "Cybernetic Proletarianization: Spirals of Devaluation and Conflict in Digitalized Production." *Capital and Class*, 46(1): 11–31.

Schoenegger, Philipp, Peter S. Park, Ezra Karger, and Philip E. Tetlock. 2024. "AI-Augmented Predictions: LLM Assistants Improve Human Forecasting Accuracy." arXiv preprint. arXiv:2402.07862.

Setterfield, Mark. 2006. "Balancing the Macroeconomic Books on the Backs of Workers: A Simple Analytical Political Economy Model of Contemporary US Capitalism." *International Journal of Political Economy*, 35(3): 46–63.

Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *The American Economic Review*, 74(3): 433–444.

SHRM. 2024. "2024 Talent Trends: Artificial Intelligence in HR." https://shrm-res.cloudinary.com/image/upload/AI/2024-Talent-Trends-Survey_Artificial-Intelligence-Findings.pdf.

Siegel, Rudolf, Cornelius J. König, and Veronika Lazar. 2022. "The Impact of Electronic Monitoring on Employees' Job Satisfaction, Stress, Performance, and Counterproductive Work Behavior: A Meta-Analysis." *Computers in Human Behavior Reports*, 8: 100227.

Smith, Chris, and Paul Thompson. 1998. "Re-Evaluating the Labour Process Debate." *Economic and Industrial Democracy*, 19(4): 551–577.

Sokolova, Anna, and Todd A. Sorensen. 2021. "Monopsony in Labor Markets: A Meta-Analysis." *ILR Review*, 74(1): 27–55.

Spenner, Kenneth I. 1983. "Deciphering Prometheus: Temporal Change in the Skill Level of Work." *American Sociological Review*, 48(6): 824–837.

Stansbury, Anna, and Lawrence H. Summers. 2020. "The Declining Worker Power Hypothesis: An Explanation for the Recent Evolution of the American Economy." *Brookings Papers on Economic Activity*, 2020(1): 1–96.

Starace, Jessica, and Carl Van Horn. 2024. "U.S. Workers Assess the Impacts of Artificial Intelligence on Jobs." Heldrich Center for Workforce Development, Rutgers University.

Svanberg, Maja, Wensu Li, Martin Fleming, Brian Goehring, and Neil Thompson. 2024. "Beyond AI Exposure: Which Tasks are Cost-Effective to Automate with Computer Vision?" SSRN working paper 4700751.

Thier, Jane. 2023. "More Companies Are Watching Their Remote Workers WFH on Camera—But It's Costing Them Big Time." *Fortune*. https://fortune.com/2023/03/29/companies-watching-remote-workers-wfh-camera-video/.

Trajtenberg, Manuel. 2019. "Artificial Intelligence as the Next GPT: A Political-Economy Perspective." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb. Chicago: University of Chicago Press.

US Council of Economic Advisors. 2022. "The Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America." The White House.

Vaithilingam, Priyan, Tianyi Zhang, and Elena L. Glassman. 2022. "Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models." *CHI EA '22*. New York, NY, USA: Association for Computing Machinery.

Verma, Pranshu. 2023. "AI is Starting to Pick Who Gets Laid Off." *Washington Post*. https://www.washingtonpost.com/technology/2023/02/20/layoff-algorithms/.

Vogel, Susanna. 2021. "How Companies Use Predictive Analytics to Get Ahead of Union Drives." *HR Brew*. https://www.hr-brew.com/stories/2021/11/03/how-companies-use-predictive-analytics-to-get-ahead-of-union-drives.

Wayne, Sandy J., Robert C. Liden, Isabel K. Graf, and Gerald R. Ferris. 1997. "The Role of Upward Influence Tactics in Human Resource Decisions." *Personnel Psychology*, 50(4): 979–1006.

Webb, Michael. 2019. "The Impact of Artificial Intelligence on the Labor Market." SSRN working paper 3482150.

Weil, David. 2014. *The Fissured Workplace.* Harvard University Press.

Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus*, 109(1): 121–136.

Wood, Alex J. 2021. "Algorithmic Management: Consequences for Work Organisation and Working Conditions." European Commission, Joint Research Centre JRC Working Papers Series on Labour, Education and Technology No. 2021/07.

Yu, Feiyang, Alex Moehring, Oishi Banerjee, Tobias Salz, Nikhil Agarwal, and Pranav Rajpurkar. 2024. "Heterogeneity and Predictors of the Effects of AI Assistance on Radiologists." *Nature Medicine*, 30(3): 837–849.

Appendix

A Monitoring

This appendix presents a general framework for a static shirking model in which effort and monitoring levels are variable and endogenous, drawing loosely on Allgulin and Ellingsen (2002). It proceeds to develop two specific examples to explore how the model's conclusions vary under different functional forms for the key relationships.

The model consists of a single worker meeting a single firm to contract for labor. The worker chooses an effort level $e > 0$ to maximize static single-period utility:

$$U(e) = -C(e) + (1 - q(e, p))w + q(e, p)\bar{w} \tag{A.1}$$

where $C(e) > 0$, with $C'(e) > 0$, represents the cost to the worker of expending effort, w is the wage paid if the job is retained at the end of the period, and $\bar{w}$ is the utility of non-work, i.e., unemployment benefits and the value of job search.[38] Workers are dismissed with a probability $q(e, p)$ where p is a monitoring parameter chosen by the employer. I assume that q is continuously differentiable with $q_e < 0$ and $q_p > 0$. The assumption of a smooth dismissal function deviates from the typical assumption in shirking models of a single effort level in the incentive compatibility constraint at which the probability of dismissal steps down from p to 0. The probability of dismissal is more likely to be a smooth function of effort when, for instance, effort levels are observed with noise (Ritter and Taylor, 2011) or when effort e denotes the share of a worker's time spent in production, as assumed in Case 1 below following Bowles (1985).

The first-order condition can be rearranged to yield the wage function:

$$w(e, p) = \bar{w} - \frac{C'(e)}{q_e(e, p)} \tag{A.2}$$

which can be interpreted as the wage required to bring forth the effort e at monitoring level p.

The firm chooses a target effort level and monitoring technique to maximize static profits, which consists of revenue net of monitoring costs and wages:

$$\Pi(e, p) = G(e) - \mu M(p) - w(e, p) \tag{A.3}$$

where $G(e)$ denotes production and $\mu M(p)$ is the cost of monitoring at level p. Assuming an interior solution, the first-order conditions yield

$$\Pi_e = G'(e) - \frac{\partial w}{\partial e} = 0 \tag{A.4}$$

$$\Pi_p = -\mu M'(p) - \frac{\partial w}{\partial p} = 0 \tag{A.5}$$

A local optimum requires that at the values e^* and p^*, $\Pi_{ee} <$ and $\Pi_{pp} < 0$ and $\Pi_{ee}\Pi_{pp} - \Pi_{ep}^2 > 0$, where subscripts denote derivatives.

[38] In a more fully specified dynamic setting, $\bar{w}$ may not be independent of monitoring technologies. If firms are assumed to be homogeneous, a change in monitoring technologies that requires greater effort from workers will be reflected in the value of the outside option.

The comparative statics of this system are straightforward. To determine how wages change with a reduction in the cost of monitoring, $\frac{dw}{d\mu}$, we first need expressions for $\frac{de}{d\mu}$ and $\frac{dp}{d\mu}$. These can be found by differentiating the first order conditions:

$$\Pi_{ee} de + \Pi_{ep} dp = 0 d\mu \tag{A.6}$$

$$\Pi_{ep} de + \Pi_{pp} dp = M'(p) d\mu \tag{A.7}$$

Using Cramer's rule, we have

$$\frac{de}{d\mu} = \frac{-M'(p)\Pi_{ep}}{D} \tag{A.8}$$

$$\frac{dp}{d\mu} = \frac{M'(p)\Pi_{ee}}{D} \tag{A.9}$$

where $D := \Pi_{ee}\Pi_{pp} - \Pi_{ep}^2$. Expressions for the second derivatives are

$$\Pi_{ee} = G''(e) - \frac{\partial^2 w}{\partial e^2} \tag{A.10}$$

$$\Pi_{ep} = -\frac{\partial^2 w}{\partial e \partial p} \tag{A.11}$$

$$\Pi_{pp} = -\mu M''(p) - \frac{\partial^2 w}{\partial p^2} \tag{A.12}$$

The change in wages for a change in monitoring costs μ is thus

$$\frac{dw}{d\mu} = \frac{\partial w}{\partial e}\frac{de}{d\mu} + \frac{\partial w}{\partial p}\frac{dp}{d\mu} \tag{A.13}$$

$$= -\frac{\partial w}{\partial e}\frac{M'(p)\Pi_{ep}}{D} + \frac{\partial w}{\partial p}\frac{M'(p)\Pi_{ee}}{D} \tag{A.14}$$

$$= \frac{M'(p)}{D}\left[\frac{\partial w}{\partial e} \cdot \frac{\partial^2 w}{\partial e \partial p} + \frac{\partial w}{\partial p}\left(G''(e) - \frac{\partial^2 w}{\partial e^2}\right)\right] \tag{A.15}$$

When $\frac{dw}{d\mu} > 0$, wages and monitoring are substitutes; a fall in the cost of monitoring reduces wages. The sign of equation A.15 depends on the nature of of the wage function, particularly the form of $C(e)$ and $q(e,p)$, as well as the convexity of the firm's output $G(e)$. In the two examples outlined below, certain forms of the functions listed above give rise to wage-monitoring relationships that are uniformly negative, uniformly positive, or ambiguous.

A.1 Case 1: Convex effort costs and simple monitoring

In this setup I follow Bowles (1985) in interpreting e as the portion of a worker's work time spent actively engaged in production and p as the probability of being observed at any given time. A worker is dismissed if observed while not engaged in production. The probability of dismissal is thus $q(e;p) = p(1-e)$ and the probability of retention is $1 - p(1-e)$. This yields the utility function

$$U(e) = -C(e) + (1 - p(1-e))w + p(1-e)\bar{w} \tag{A.16}$$

and, from the worker's first-order conditions, the associated wage function

$$w(e, p) = \bar{w} + \frac{C'(e)}{p} \tag{A.17}$$

The wage is decreasing in p: better monitoring reduces the efficiency wage required to compel effort. The second-order condition requires that $C''(e) > 0$, i.e., the cost of exertion is convex in effort $e \in [0, 1]$. From equations A.15 and A.17 we can characterize the change in the wage for a change in monitoring costs (provided the solution is interior):

$$\frac{dw}{d\mu} = \frac{M'(p)}{D}\left[-\frac{C''(e)^2}{p^3} - \frac{C'(e)}{p^2}\left(G''(e) - \frac{C'''(e)}{p}\right)\right] \tag{A.18}$$

$$= -\frac{M'(p)}{p^3 D}\left[C''(e)^2 + C'(e)\Big(pG''(e) - C'''(e)\Big)\right] \tag{A.19}$$

Wages and monitoring costs are substitutes when the expression in square brackets is negative.[39] The overall sign depends both on the nature of production $G(e)$ and the curvature of $C(e)$. Assuming a non-convex production technology, $G''(e) \leq 0$, equation A.19 will be positive when $C''(e)^2 < C'(e)C'''(e)$. Consider the power function $C(e) = cx^d$, with $d > 1$. In this case wages and monitoring costs are complements ($\frac{dw}{d\mu} < 0$) for any $G''(e) \geq 0$. When firms instead face diminishing returns to production, $G''(e) < 0$, the cost function above gives rise to an expression for $\frac{dw}{d\mu}$ that may be positive at some points in the (e, p) space and negative at others.

Sticking with the assumption of a concave production function $G''(e) < 0$, one effort function that makes wages and monitoring costs substitutes is $C(e) = c(1-x)^{-d}$, $d > 0$. Since by assumption $e \in [0, 1]$, this effort function reflects a situation where effort costs rise asymptotically as workers approach non-stop exertion. Finally, with exponential effort costs, $C(e) = c \cdot \exp(de)$, wages and monitoring costs are substitutes when $G(e)$ is concave and $\frac{dw}{d\mu} = 0$ when $G(e)$ is linear. The example below approaches this same result from a separate set of assumptions.

A.2 Case 2: Linear effort costs and Poisson monitoring

Consider a situation in which discrete signals of worker effort accumulate at random intervals, as do monitoring reports sent to managers. Specifically, suppose that the fruits of worker effort or "effort signals" generated by the workers are Poisson distributed, accumulating at a rate governed by worker effort e. Meanwhile, reports on worker performance accumulate according to the rate parameter p, which is determined by the firm. Workers are dismissed when at least one report arrives over some fixed time period (say, one day), during which the count of effort signals is 0. This can be represented by the joint probability

$$\Pr(dismiss) = \Pr(signal = 0) \cdot \Pr(report \geq 1) \tag{A.20}$$

$$= \frac{e^0 \exp(-e)}{0!} \cdot \left(1 - \frac{p^0 \exp(-p)}{0!}\right) \tag{A.21}$$

$$= \exp(-e)\left(1 - \exp(-p)\right) \tag{A.22}$$

[39] Note that the term in parentheses, $G''(e) - \frac{C'''(e)}{p} = \Pi_{ee}$, is necessarily negative following the second-order conditions.

Higher effort reduces the probability of dismissal and greater monitoring increases it. Using equation A.22 as the dismissal function $q(e, p)$ and assuming linear effort costs $C(e) = e$, workers optimize according to

$$U(e) = -e + \big[1 - \exp(-e)\delta(p)\big]w + \exp(-e)\delta(p)\bar{w} \tag{A.23}$$

where $\delta(p) := 1 - \exp(-p)$. The first-order conditions yield the wage function

$$w(e, p) = \bar{w} + \frac{\exp(e)}{\delta(p)} \tag{A.24}$$

Taking the appropriate derivatives of equation A.24 and using equation A.18 yields the following expression:

$$\frac{dw}{d\mu} = \frac{M'(p)}{D}\left[\frac{\partial w}{\partial e}\cdot\frac{\partial^2 w}{\partial e \partial p} + \frac{\partial w}{\partial p}\left(G''(e) - \frac{\partial^2 w}{\partial e^2}\right)\right] \tag{A.25}$$

$$= \frac{M'(p)}{D}\left[\frac{\exp(e)}{\delta(p)}\left(-\frac{\exp(e-p)}{\delta(p)^2}\right) - \frac{\exp(e-p)}{\delta(p)^2}\left(G''(e) - \frac{\exp(e)}{\delta(p)}\right)\right] \tag{A.26}$$

$$= -\frac{M'(p)}{D}\cdot\frac{\exp(e-p)}{(1-\exp(-p))^2}\cdot G''(e) \tag{A.27}$$

Now the direction of the relationship between monitoring costs and wages is determined entirely by the production technique. When the firm faces diminishing returns to worker effort, $G''(e) < 0$, wages and monitoring costs are substitutes. With a linear production function, wages do not change with monitoring costs. Yet since a fall in μ prompts firms to invest more in monitoring, workers must supply more effort. The gains from this increased effort are captured entirely by the firm as worker welfare falls. Efficiency-wage effects may manifest as an uncompensated increase in effort rather than an absolute fall in wages.

B Wage offer

This section considers the effects of improved information on the part of a hiring manager in a setting where the buyer of labor is uncertain as to the reservation wage of the seller. As in a broad class of search models, job seekers encounter firms and firms make take-it-or-leave-it job offers. Workers vary only in their reservation wages, which are distributed according to $F_b(b)$. When an encounter occurs, the firm observes the worker's reservation wage with some noise ε distributed according to $F_\varepsilon(\varepsilon)$. The observed reservation wage is $z = b + \varepsilon$, where $E(\varepsilon) = 0$. Although the parameters governing the two distributions are known to the firm, for any given observation z the firm cannot distinguish between the noise and the signal. Yet since the firm knows the underlying distributions, it is able to formulate the conditional distribution $F_{\varepsilon|z}(\varepsilon|z)$. The observed value of z carries information about the value of ε, helping the firm to target the wage rule.

The firm chooses a wage offer $w(z, \phi)$ according to a consistent rule parameterized by ϕ. Workers accept the offer if $w(z, \phi) \geq b$ and begin producing immediately at productivity p. Otherwise they reject the offer and the two parties go separate ways.[40] Consider a firm whose wage rule is $w(z) = z$. When ε is symmetric around 0, the firm can be sure that half of its wage offers will be rejected and expected

[40] A more involved version of the model could formalize a bargaining process.

production will also be reduced by half from the maximum possible per job opening. Thus it might be profitable to offer a wage above z to reduce the share of rejected offers. To simplify the analysis, the wage rule is assumed to be one in which the firm adds a constant or "buffer" ϕ on top of the observed reservation wage, $w(z, \phi) = z + \phi$, though more elaborate rules might be explored.

The firm's objective is to choose a wage buffer ϕ that maximizes expected single-period profits:

$$\Pi(\phi) = \mathbb{E}\Big[(p - w(z, \phi)) \Pr(w(z, \phi) \geq b)\Big] \tag{B.1}$$

Given the specification for the wage rule, the expression for the probability of job acceptance becomes $\Pr(\varepsilon + \phi \geq 0) = 1 - F_{\varepsilon|z}(-\phi)$. Since the firm observes only z, the expectation operator is taken over z. For simplicity, assume that when the firm observes a signal z for which the wage rule would specify an unprofitable wage $w(z, \phi) > p$, the firm chooses not to make a wage offer.[41] The objective function is thus

$$\Pi(\phi) = \int_{-\infty}^{p-\phi} (p - \phi - z)(1 - F_{\varepsilon|z}(-\phi)) f_z(z) dz \tag{B.2}$$

The first-order condition yields

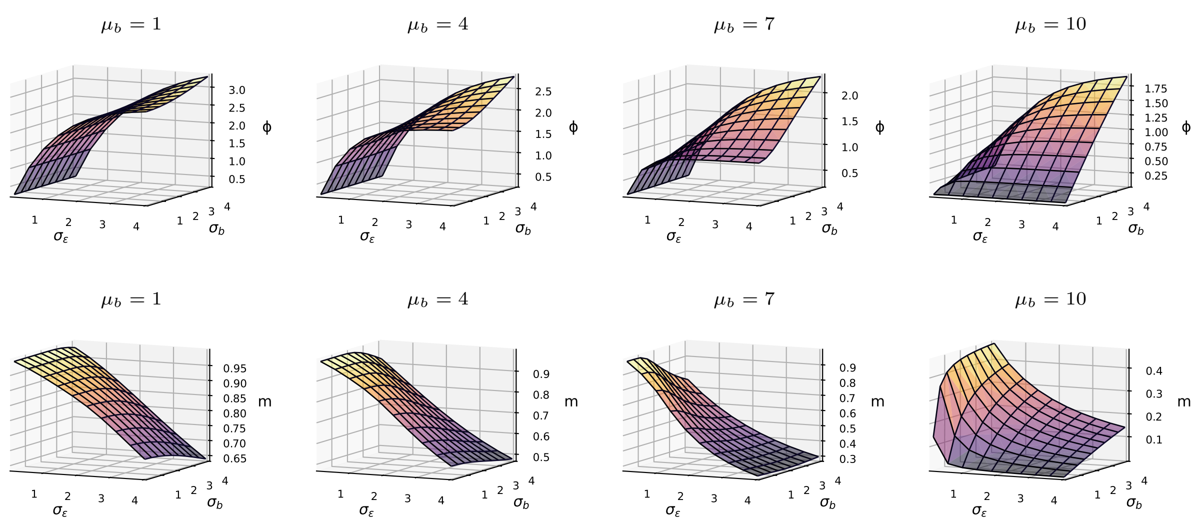
$$\Pi'(\phi) = \int_{-\infty}^{p-\phi} \Big[-1 + F_{\varepsilon|z}(-\phi) + (p - \phi - z) f_{\varepsilon|z}(-\phi)\Big] f_z(z) dz = 0 \tag{B.3}$$

which defines implicitly the firm's optimal wage buffer ϕ^*.

When distributions F_z and F_ε are chosen to yield nontrivial solutions, equation B.3 is generally intractable, as are the comparative statics involving ϕ. In what follows, I present numerical solutions of the optimal wage buffer ϕ^* for a range of values in the parameter space in which $b \sim N(\mu_b, \sigma_b)$ and $\varepsilon \sim N(0, \sigma_\varepsilon)$. For this numerical demonstration, I set $p = 10$. Figure 1 depicts the solutions of ϕ^* over the $(\sigma_b, \sigma_\varepsilon)$ space for four different values of μ_b. The wage buffer ϕ rises with observation noise σ_ε across all combinations of μ_b and σ_b. The second row in the figure illustrates the expected match rate between worker and firm, denoted m, at each optimal ϕ value. As might be expected, the match rate falls as the noise rises in the firm's observation of worker reservation wages.

[41] This assumption could be relaxed at the expense of considerably more complexity.

Figure 1: Numerical demonstration of wage offer with noisy reservation wage observation



Note: Figures show results of numerical solutions to Equation B.3 for productivity value $p = 10$, $b \sim N(\mu_b, \sigma_b)$, and $\varepsilon \sim N(0, \sigma_\varepsilon)$. Each column reflects solutions for a different value of the reservation wage mean μ_b. In the first row: The variable on the vertical axis, ϕ, is the firm-optimal value of the wage buffer. In the second row: The variable on the vertical axis, m, is the expected worker-firm match rate at the optimal value of ϕ corresponding to σ_b and σ_ε.

C Employee screening

The setup is a two-period model in which a single firm hires a fixed number N of workers and commits to paying them a wage w bounded by a wage floor or commonly held reservation wage w_0. Workers come in one of two types, exiters and loyalists. The share of exiters in the labor market, denoted x_0, is fixed. The share at the firm is denoted by x. The firm can screen workers by paying a per-worker cost $c(x-x_0)^2$.

Production takes two periods to complete. Between the first and second period, exiters sample from the exogenous outside job distribution $F(\tilde{w})$ and accept other jobs when $\tilde{w} > w$. In period two, production is completed by the remaining workers, who number $N(1 - x\bar{F}(w))$, where $\bar{F}(\cdot) \equiv 1 - F(\cdot)$. Although all workers receive the wage w, output is determined by the number of remaining workers after the job search occurs. Profit per worker is

$$\Pi(w, x) = A(1 - x\bar{F}(w)) - w - c(x - x_0)^2 \tag{C.1}$$

The first-order conditions are

$$\Pi_w(w, x) = Axf(w) - 1 = 0 \tag{C.2}$$

$$\Pi_x(w, x) = -A\bar{F}(w) - 2c(x - x_0) = 0 \tag{C.3}$$

Equation C.3 yields $x = x_0 - \frac{A}{2c}\bar{F}(w)$: the firm optimally screens out exiters so that the share of exiters among its hires falls below their share in the workforce.

Assuming a continuously differentiable outside wage distribution, the comparative statics of the system for a change in the cost of screening c are straightforward to characterize. By the implicit function theorem:

$$\begin{bmatrix} \frac{dw}{dc} \\ \frac{dx}{dc} \end{bmatrix} = -\frac{1}{D} \begin{bmatrix} -\Pi_{wx}\Pi_{xc} \\ \Pi_{ww}\Pi_{xc} \end{bmatrix} \tag{C.4}$$

$$= -\frac{1}{D} \begin{bmatrix} 2Af(w)(x - x_0) \\ -2Af'(w)x(x - x_0) \end{bmatrix} \tag{C.5}$$

where $D \equiv \Pi_{ww}\Pi_{xx} - \Pi_{wx}^2$ and by the second-order conditions $D > 0$. It then follows that

$$\frac{dw}{dc} = -\frac{2Af(w)(x - x_0)}{D} > 0 \tag{C.6}$$

since $x < x_0$. For interior solutions, a fall in the cost of screening workers leads to a reduction in wages. The intuition is straightforward: By screening out exiters, the firm insulates itself from wage competition.

D Delegation and authority

This section builds on the model of authority and incentives described in Section V.B of Aghion and Tirole (1997), a special case of their more general model. A principal or manager ("she") contracts with an agent or worker ("he") to select a project from a menu of options $k \in \{1, \ldots, N\}$.[42] The manager's preferred project delivers profits B while the worker's preferred project yields private benefit b. For the manager, there are two "relevant" projects: one with profits $B > 0$ and the other yielding zero. The rest of the projects yield negative profits, at least one of which is "sufficiently negative" that an uninformed manager would choose inaction over engaging in a project with unknown returns. The same two "relevant" projects yield nonnegative benefits to the worker, one of which produces $b > 0$ and the other 0.

The ex ante probability that the manager and worker prefer the same project is governed by the congruence parameter $\alpha \in (0, 1]$. Thus the expected profit for the manager when the worker picks his highest-benefit project is αB; likewise, the worker's expected benefit when the manager picks is αb. Both manager and worker may invest costly effort into learning project payoffs. By expending effort $E \in [0, 1]$ the manager learns all project payoffs with probability E while learning nothing with probability $1 - E$. The worker likewise expends effort $e \in [0, 1]$ to learn about the projects with probability e. Effort costs are $G(E)$ and $g(e)$ for the manager and worker, respectively, functions that are strictly increasing and strictly convex. It is reasonable to expect that the worker has a lower cost of acquiring information since he is closer to the action.

Although it may be the case that the manager delegates decision-making entirely to the worker—that is, she promises not to reverse the decisions that the worker makes—the relevant situation here is one in which the manager retains formal authority. The manager provisionally accepts the worker's suggestion, but, having simultaneously expended effort $G(E)$ and discovered project payoffs with probability E, may reverse that decision to enact her preferred project. In order to spur initiative on the part of the worker, the manager offers an incentive payment of w when the project yielding B is selected and 0 otherwise. Assuming a risk-neutral worker, a payment of $w \geq b$ ensures that an informed worker will always report the manager's preferred decision. In what follows, I will pursue this case of aligned incentives.

Expected utility functions for the manager (U) and worker (V) are

$$U(E, e, w) = E(B - w) + (1 - E)e(B - w) - G(E) \tag{D.1}$$

$$V(E, e, w) = E(w + \alpha b) + (1 - E)e(w + \alpha b) - g(e) \tag{D.2}$$

With probability E, the principal learns which project yields B, this project is enacted, and the agent receives expected utility $w + \alpha b$. With probability e the agent learns the payoffs and, since incentives are aligned, the principal's preferred project is again selected. In the case that neither learns the payoffs, no project is selected and the worker receives no incentive pay w.

Assuming that the worker takes the principal's actions as given, the first-order condition of equation D.2 can be inverted to produce a worker-effort function $\hat{e}(w, E)$. Suppose that the functional form of effort-cost function is $g(e) \equiv \frac{\gamma}{1-e}$ (and the manager's effort-cost function has the same form, $G(E) \equiv \frac{\Gamma}{1-E}$, with $\Gamma > \gamma$). Then, after taking the worker's first-order condition with respect to e, we have

$$\hat{e}(w, E) = 1 - \sqrt{\frac{\gamma}{(1 - E)(w + \alpha b)}} \tag{D.3}$$

[42] I adopt the pronoun conventions of the literature. Here it is helpful to think of the projects as tasks that might be carried out by the agent as part of the production process, the timing and order of which are not set in stone. This setting diverges somewhat from the situations typically referenced in agency theory, which more often have to do with delegating authority between units within the organizational hierarchy of a firm.

which has the intuitive properties that a higher wage and greater congruence of project preferences brings forth more initiative on the part of the worker, while greater effort on the part of the manager reduces the worker's initiative (since the manager is better able to identify her preferred project herself, rendering worker effort redundant). The manager's utility function is now

$$U(E,w) = E(B-w) + (1-E)(B-w)\hat{e}(w,E) - G(E) \tag{D.4}$$

$$= (B-w)\left(1-\sqrt{\frac{\gamma(1-E)}{w+\alpha b}}\right) - \frac{\Gamma}{1-E} \tag{D.5}$$

where the second line uses the assumed functional form of $G(E)$. We now have a single objective function in two variables, making comparative statics simple to express. The object of interest is $\frac{dw}{d\Gamma}$, or the change in the equilibrium wage for a change in the cost of monitoring projects for the manager. By the implicit function theorem, the comparative statics can be expressed as

$$\begin{bmatrix} \frac{dE}{d\Gamma} \\ \frac{dw}{d\Gamma} \end{bmatrix} = -\frac{1}{D}\begin{bmatrix} U_{ww}U_{E\Gamma} \\ -U_{Ew}U_{E\Gamma} \end{bmatrix} \tag{D.6}$$

where $D \equiv U_{ww}U_{xx} - U_{wx}^2$ and $D > 0$ by the second-order conditions. Taking the appropriate second derivatives yields

$$\frac{dE}{d\Gamma} = -\frac{3B+w+4\alpha b}{4D}\left[\frac{\gamma}{(1-E)^3(w+\alpha b)^5}\right]^{1/2} < 0 \tag{D.7}$$

$$\frac{dw}{d\Gamma} = \frac{B+w+2\alpha b}{4D}\left[\frac{\gamma}{(1-E)^5(w+\alpha b)^3}\right]^{1/2} > 0 \tag{D.8}$$

A reduction in project monitoring costs encourages the manager to exert more of her own effort and pay the worker less. This result reflects a straightforward substitution of managerial effort for worker initiative.